Reply to reviewer 2

We would like to thank the reviewer for the constructive comments. In the following we recall the comments and we reply to each of the comments in turn (in green).

The manuscript addresses a highly relevant topic for the natural hazards and coastal risk community. It makes use of data-driven metamodels as an effort to support operational early warnings for marine flooding. The comparison of analog (M1) and meta-model approaches (M2, M3) is scientifically interesting and potentially valuable for researchers and coastal managers. Furthermore, the study is generally well structured and the outcomes are sufficiently described.

However, in its current form, the manuscript requires further clarifications and editing improvements before considering publication. My main concerns arise from the following aspects:

The authors refer to “overflowing” processes, but the widely accepted term of wave run-up is not introduced or discussed. This aspect should be clarified and integrated into the terminology of the manuscript.

=> Please see answers to comments on page 5 line 8-20

(2) Several aspects of the modeling strategy require clearer justifications, including a brief presentation of the constructed pseudo-historical events. In particular, the construction of pseudo-historical events should be better described, including the ranges of forcing parameters. Furthermore,

=> Please see answers to comments on page 7 line 11

(3) The extraction of the surge level (SPM) parameter took place at the location of the buoy located offshore (at 50 m deoth) and not at the tide gauge located at the nearshore.

=> Please see answer to comment on page 6 line 5

(4) M3 method should be more clearly framed as a refinement or extension of M2, mainly improving the spatial reconstruction of flood depths rather than the prediction of flood extent or impact category.

=> Please see answer to comment on page 15 line 20

(5) The inclusion of 1000-year return period storm scenarios raise concerns, given the limited observational basis and uncertainties of the hydrodynamic parameters, and thus should be reconsidered.

=> Please see answer to comment on page 11 line 1

(6) Use of terminology, technical specifications and text editing

Detailed Comments

Page 2

Line 12: Think that this chapter should be renamed to “Mehodology.

⇨ We would prefer to keep the name of the chapter because the method is described in two chapters (chapter 2 for modeling and chapter 3 for metamodeling and performance assessment).

Line 13: This sub-chapter could be renames to “Study Area” => OK

Lines 27-28: Modify from “High Performance Computing, reduced process complexity with generalized overflowing, models or empirical formula, statistical analysis of pre-calculated flooding scenarios” to “(1) High Performance Computing (HPC); (2) reduced process complexity with generalized overflowing, models or empirical formulas; and (3) statistical analysis of pre-calculated flooding scenarios”. => OK

Line 31: Exclude “finally”. => OK

Page 3

Lines 3-4 and Line 32: Use of the widely acknowledged term of “Total Water Level (TWL)”

instead of “total sea levels”. Replace also elsewhere in the document. => OK

Lines 5-6: This sentence shall be replaced/modified to make more sense to the reader “If databases are often used to work by analogy”.

⇨ This sentence will be replaced by “Databases are often used ……., but this solution also provides possibilities ….”

Line 9: Add a comma or “and to” before (2). => OK

Lines 12-13 and Line 20: References listed here shall be updated. There is significant work achieved in this sector since 2020.

⇨ This part of the introduction will be updated as follows:

“ML-based metamodeling techniques have made great progresses for flood predictions (storm surge: e.g., Macdonald et al., 2025; hurricane: e.g., Irwin et al., 2025; tsunami: e.g., Ragu Ramalingam et al., 2025; estuarine: e.g., Wang et al., 2025a; river flood: e.g., Wang et al., 2025b) and opens up encouraging perspectives for marine flooding forecasting. Through a statistical analysis of pre-calculated training databases, it can predict key flooding indicators (surge, discharge, water height, etc.) at a given spatial location of interest within reasonable time and computing resources while preserving the accuracy of full process models. Yet, some issues remain to push this metamodel-based approach toward operational applications, and more specifically the production of spatialized

indicators, which is still a matter of active research (among others, Ma et al., 2022; Spiller et al., 2023; Fraehr et al., 2024; Rohmer et al., 2024a; Jung et al., 2025)."

Line 15: Replace "it can predict" with "they can predict". => OK

Line 22: Rephrase from "...forcing conditions and flood maps" to "...forcing conditions and the generation of flood maps". => OK

Line 23: Add a space before "s" (seconds). => OK

Line 24: Rephrase from "The objective is then to investigate" to "The objective of this work is to investigate". => OK

Line 27: Replace "...with needs of crisis managers" with "...according to the needs of crisis managers". => OK

Page 4

Lines 1-5: It is not clear who applied these simulations and if this is an on-line available database. Please justify. => The reference of the study (report Lecacheux et al. 2023) that enabled us to build the database will be added.

Line 6: Rephrase "In the following" with "In the following chapters" or something relevant. => OK

Lines 6-11: Replace "Sect." with "Section". => OK

Line 22: "NGF", the abbreviation is not defined. => the definition of NGF "the official vertical reference system to measure elevations in France" will be added

Page 5

Figure 1A: The depth scale is not well defined and shows just the 0 and 10 m values. Also, areas close to the shoreline appear to have depth values that are close to 10 m, while offshore values (close to the buoy) appear to have depths close to 0 m (instead of 50 m, as stated later at the text). Suspect that the depth scale here is wrong, please check this aspect carefully. Also, suggest including the north symbol.

=> The depth scale will be corrected (0 replaced by 50) and the north symbol added

Lines 8-20: Wave run-up is not mentioned here, a crucial component. Suggest revising this entire section, making the distinctions between wave-wind-river components of coastal flooding clearer to the reader. Suggest providing a distinction between - marine (waves, wind) and - riverine flooding. In alliance with the terminology used in your text.

=> The study excludes river contribution and wave overtopping (which is negligeable on the pilot side). It focuses on overflowing caused by the static sea water level (including tide, atmospheric surge and wave set-up). That is why the study does not mention the wave run-up that corresponds to total dynamic level reached by waves (including swash).

A paragraph will be added to precise the physical components taken into account in the study : “Thus, the study focuses on the simulation of the total steel water level (including tide, atmospheric surge caused by winds and pressure drop and the wave set-up caused by the breaking of waves in entry of the lagoon) and the resulting coastal flooding. The dynamic water level (including the swash component) is not taken into account.”

Line 19: Rephrase from “in particular in” to “particularly at”. => OK

Page 6

Line 5: “surge conditions”. I suspect that you refer to SPM (surge skewness) here. However, how is it possible to extract this parameter from a buoy located at 50 m depth? As surge involves the meteorological contribution to sea level (tide in this case), wouldn’t it be safer to estimate this value by using the tide gauge data instead?

=> We extract the SPM at 50 m depth because this position corresponds to the boundary of the hydrodynamic model that propagates the surge inside the domain and computes the additional surge generated by waves and local winds. This position was chosen as boundary because the storm surge is independent of sea level (tide) at this point which is not the case inside the lagoon. We acknowledge that this aspect deserves a clarified description notably concerning the inputs and outputs of the hydrodynamic model. A more detailed description of the modeling components will be added in chapter 2.3 (see answer to comment on page 7 line 9-10 on the modeling method).

Line 9: Describe SPM abbreviation before using it (skew surge peak method). Please define also its calculation (is it the difference between max. sea level and max. tidal level?)

=> The definition and calculation of the skew surge will be detailed as “the difference between the maximum simulated water level and the predicted high tide level”

Line 12: Rephrase from “peak period Tp, wave directions Dp and winds Du” to “peak period - Tp, wave direction - Dp and wind direction - Du” => OK

Lines 19-20: Why is SPM (surge level) extracted at the location of the buoy (50 m) and not at the location of the tide gauge? As I see it the proper methodology would be to extract the tide level (T) at 50 m depth and use this value in conjunction with the sea level measured close to the shore (tide gauge) for estimating surge skewness.

⇨ Please refer to the answer for comment Page 6 – line 5.

Line 28: Correct “then” to “the” => OK

Line 30: Be consistent for the term “pseudo-historical”. Add the quotations “” also in line 21.

=> the quotations will be removed to be consistent with the rest of the manuscript

Lines 29-30: Here you state that “storms (Fabien, Klaus and Domingos) occurred at low tide coefficients”. Previously, In line 26 you state that “Table 1 shows the characteristics of the monitored 8 storms during high tide”. By checking your table it seems that indeed Klaus, Fabien and Domingos occurred during low tidal level. Please correct this paragraph accordingly.

=> The text seems correct. High tide is the time when the water level reaches its highest level during a tidal cycle. This term is used to designate this precise moment whatever the tide coefficients.

Lines 30-31, and Lines 1-2, p. 6: It is not clear how these pseudo-historical events have been created. You mention that you used the high tide of the synthetic events to construct 80 pseudo-events, but what is the range of the forcing parameters you used for this treatment and which concept was used ? Also, the pseudo_StormName_TideLevel(NGF) is not shown or mentioned elsewhere in the manuscript.

⇨ To construct the pseudo-historical events, we use the 8 historical storms (i.e. their forcing conditions characterized by parameters Hs, Tp, Dp, U, Du, SPM) and we associate them with 10 different tidal coefficients (described by 10 high tide levels T) instead of the historical tidal coefficient. The purpose is to create new events with the same historical storms but with lower or higher tidal coefficients. The nomenclature “*Pseudo_StormName_…”* is reused in chapter 4.2 to describe the performance of the metamodeling approaches on these pseudo-historical events. To facilitate the understanding of the readers, we will detail in the manuscript the name of the parameters of pseudo-historical events extracted from historical storms or the 10 high tide levels.

Page 7

Line 2: Add full name in italic “Pseudo_StormName_TideLevel(NGF)”. => OK

Lines 9-10: Need to be more descriptive on the modeling strategy you used here. Did you use the output of WW3 from Copernicus Marine Service, coupled with the UHAINA model to simulate hydrodynamics, or did you use the WW3 model itself. This should be clearly stated within the text. Also, why UHAINA model is selected amongst others found in the relevant literature. Are there any advantages compared to the usage of other well known hydrodynamic models for which some of your storms have been successfully validated in the past. Specifically, Xynthia is a well-known event that has been previously validated by several models (e.g. Xbeach).

Line 11: Be more specific to your grid resolution. Suggest to refer to the specific range of your grid resolution in meters (offshore, nearshore and land).

- ⇨ This chapter describes the modeling chain set up within the project (including a local WW3 model and the UHAINA model) to downscale conditions from 50 m depth into the lagoon and on land. These models have been chosen because they enable to resolve the physical equations needed. Xbeach was not chosen because the wave runup and the wave overtopping are not taken into account in the study that focuses on marine flooding by overflowing only.
- ⇨ The text will be completed as follows for the model description: “The marine flooding model set up for the project includes a chaining of WW3 spectral wave model (Tolman, 2014) and UHAINA hydrodynamic models (Filippini et al., 2024) that resolves Saint-Venant equations. The grid is an unstructured mesh with a resolution from kilometric offshore to decametric inside the lagoon (from 20 m to 50m) and on land (~10 m) (Fig. 3). This chain of models is forced by tidal harmonics, atmospheric surge and waves at the western boundary (that passes through the Cap-Ferret buoy) and by the wind on the entire domain. It can simulate (1) the propagation of tide, atmospheric surge and waves inside the lagoon; (2) the generation and propagation of the additional surge induced by local winds and the breaking of waves and (3) the marine flooding by overflowing.”
- ⇨ The text will be completed as follows concerning the forcing conditions: “For historical and pseudo-historical events, forcing conditions come from time-varying observations at the Cap-Ferret buoy (for waves and wind) and reanalysed data of Météo-France Arpège model (for storm surge). For synthetic events it corresponds to the set of parameters at high tide determined with the statistical analysis (U, Du, Hs, Tp, Dp, SPM) that are applied as stationary conditions during the tidal cycle.”

Lines 14-16: English check of this sentence.

- ⇨ The sentence will be replaced by “Despite the resolution of the mesh remains limited on land, the model represents the main obstacles for water flows thanks to the line constraints that represent the protection structures (dikes), the little walls, the embankments and the roads.”

Page 8

Figure 3: Suggest using a different color than the already used one for showing the box (red) to indicate the land limits of your grid.

=> the red color for the box will be changed (red=> black)

Lines 8-9: Suggest to rephrase from “consideration of temporality “ to “temporal resolution”.

=> this is not what we meant (see answer to next comment)

Lines 9-12: Be clear of the temporal resolution of the synthetic (is it hourly) and the historical, pseudo-events (what's the resolution you used here and how did you reconstruct the bulk parameters).

⇨ The forcing conditions for synthetic events are stationary on the tidal cycle. Some details will be added in the text following the modifications described for comment on page 7 lines 9-11.

Line 18: Replace "Concerning" with "Regarding". => OK

Line 19: Remove "also". => OK

Lines 22-24: Check the use of english here and rephrase this sentence.

=> the sentence will be clarified and split into two sentences : "The interested reader will find further information about simulations sensitivity tests in Lecacheux et al. (2023). They corroborate that (1) the modeling chain reproduces correctly TWLmax and Hmax on the ten municipalities (2) TWLmax and Hmax are mainly controlled by conditions at high tide, which justifies the use of storm parameters at high tide (U, Du, Hs, Tp, Dp, SPM) in the simulations."

Page 9

Line 5: Suggest being careful here when referring to "more than millennial" events, concerning the span of data coverage.

=> this comment will be deleted to avoid any misunderstanding (see answer to comment on page 11 line 1)

Lines 13-14: This sentence is not clear, please rephrase "It induces that strong storm events are simulated with more tide levels than moderate ones ...". Guess you mean "higher levels".

=> the sentence will be clarified as follows : "It induces that strong storm events are simulated with more tide levels than moderate storm events (because moderate events generate marine flooding only for the highest tidal coefficients)."

Lines 15-18: How did you conclude to 220 and 32 simulation parameters for learning and testing/validation? This should be clear in the text. What are the combinations used ranges). Believe that these modeling scenarios should be submitted as a supplementary material accompanying this manuscript.

=>The 220 scenarios for the learning database correspond to the subset of the initial 500 scenarios (50 storm X 10 tide levels) that generated marine flooding.

=>The number of 32 pseudo-historical events for the validation corresponds to the subset of the 80 pseudo-historical events (8 storms X 10 tide levels) that generated marine flooding.

It will be clarified in the text as follows: “To sum-up, the learning dataset is composed of 220 simulated scenario and the validation dataset of 32 simulated scenarios (corresponding to the subset of the initial scenarios that generated marine flooding)”

Page 10

Lines 2-5: This should be moved to Lines 15-18, p.8. Also, check the use of english of the first sentence.

⇨ We would prefer keeping these lines in chapter 2.4 that describes the database with water level and flood maps ranges

Line 6: Remove “although”. => OK

Lines 20-22: Check the use of english in this sentence (rephrase also problematics).

⇨ The sentence will be replaced by : “But being exposed does not necessarily mean being vulnerable. The impact of a flooding event depends not only on its severity, but also on the stakes involved and the challenges faced by crisis managers”

Page 11

Line 1: Again, I strongly disagree the inclusion of storm scenarios with a 1000-year return period. Do you feel confident to include such a scenario on the basis of your hydrodynamic forcing, based on the studied storm events? How do you justify it.

⇨ The mention of 100-year return period will be removed. The study does not intend to associate return periods to the scenarios because (1) the return period is not a key element in the operational use of flood forecast; (2) we do not have a sufficiently long dataset to precisely estimate a 1000-year return period event. The objective of this remark was to underline that we consider extreme scenarios, but we acknowledge that this can be confusing for the readers.

⇨ Now, the most extreme events of the dataset correspond to extreme storms configurations (obtained with the multi-variate extreme value analysis presented on Fig. 2 (A)) associated with high tide coefficients. Thus, even if we cannot precisely estimate the return period of such events, we are confident that they remain realistic and they must be included in the database whose purpose is to cover the range of possibilities.

Lines 12-13: 30 cm and 50 cm range of SSHmax doesn’t sound logical. Think you refer to the range of differences amongst SSHmax. Please clarify and rephrase appropriately.

⇨ The sentence will be clarified as follows : “These thresholds generally correspond to increments of about 30 cm in most cases (but the increments can reach 50 cm in some cases)”

Lines 13-15: Check the use of english.

⇨ This sentence will be clarified as follows : “Figure 6 shows the thresholds of the categories for each municipality and the distribution of the scenarios of the learning database (box plots) and validation database (points) inside the categories. “

Lines 16-21: Check the use of english.

⇨ This sentence will be clarified as follows : “Even if the municipalities comprising wetlands (notably Biganos, Audenge, Ares) are more exposed to marine flooding, their urban areas and critical stakes are impacted for higher TWLmax than the other municipalities”

Page 12

Lines 1-6: Check the use of english.

⇨ The text will be replaced by: “This preliminary work highlighted the need for forecasting water levels (TWLmax) and the associated categories of impact (CAT) for each municipality first, so that crisis managers can rely on predefined action plan typologies. Forecasting marine flood maps (Hmax) should then be undertaken as a second step, in order to complement and refine the characterization of the event within each impact category.”

Lines 9-10: Max. water height is defined as Hmax. This could create confusion with sea level height. Suggest to rename this parameter to “flood depth”. Also, as Hmax is commonly used to describe the max. wave height, this can create confusion for the reader.

⇨ We would rather keep this terminology to remain consistent with other publications and avoid the introduction of mistakes in the text and figures with numerous updates.

Line 11: Rephrase from “Let us define” to “This is defined as” or something relevant. => OK

Line 21-23: Check the use of english. A better description is needed here. Suggest to revise the entire 3.1.1 section, considering also the next comment.

=> The sentence will be clarified: “For given metocean forcing conditions, it consists in querying the learning database to identify the scenarios whose conditions are the most similar, and retrieving the corresponding TWLmax and Hmax maps”

Page 13

Lines 3-6: This is an interesting finding, how is this justified, is this be supported by relevant works? I’d suggest to enhance this paragraph.

⇨ In this paragraph we will develop the methodology (notably through a sensitivity analysis) and justify the choice of a single analog.

Line 23: Exclude “etc.” => OK

Lines 10-13: Suggest to revise this sentence to be more understandable to the reader.

=> The sentence will be clarified: “On this basis, we identify, for each municipality, the analog flood map in the learning database whose TWLmax values are closest to those predicted by the GP-based model”

Page 14

Line 21: Provide the full name for RF. => OK

Page 15

Lines 2-6: Be consistent with the equations format, according to the journal specs. Numbering needs to be revised for all equations. => OK

Line 15: Capitalize “hmax”. => OK

Line 20: “(in M2 and thus M3)” – This is a point where you can state that M3 is used to improve M2.

- ⇨ We choose to insist on this point rather on paragraph 3.1.3 by adding the following concluding remark : “Thus, M3 can be considered a refinement of M2, as it uses the same methods to compute TWLmax and the flood extent, while improving the reconstruction of flood depth (Hmax) map through the use of an autoencoder.”
- ⇨ Fig. 7 will also be modified, as suggested, to better describe the link between M2 and M3.

Page 16

Figure 8: Need to revise the legends shown on (A) and (B) for consistency. Also, suggest to revise the figure’s caption to be more understandable to the reader. Indicate also the relevant areas shown on the map of the lagoon on panel B. The use of an arrow could be more indicative.

=> OK

* Entire Section 3.1: M3 is actually used for the improvement of M2 approach. This can be confusing for the reader. Think that this could be to represent visually through a diagram, showing also the main processes involved in M1, M2 and how M3 is nested in M2.

=> See answer to comment on page 15 line 20.

Page 17

Figure 9: As Figure 8. => OK

Lines 6-7: Check the use of english. => OK

Lines 7-9: Use either (1), (2), (3) or i); ii; iii); as previously done in the text.

⇨ The manuscript will be harmonized with (1), (2), ...

Page 18

Figure 10: From the results, it is evident that Pseudo_Klaus_2.0 (Audence) error is reduced by 25%. However, rest of the "Preudo" cases (and the points) are not significantly affected. This should be highlighted in the text (before line 5). => OK

Page 21

Figure 13: CAT panel (upper). As M2 and M3 share the same predictions, the results duplicated for M3 can be excluded, and instead rename the panel for "M2" as "M2-M3". => OK

Page 22

Line 2: Here you make use of the entIre word "Section" instead of "Sect." Suggest to keep it as it is and revise rest of the manuscript accordingly, where needed. => OK

Line 3: Make use of "Table" instead of "Tab.". => OK

Page 23

Lines 10-18. Considering the time passed since submission, if there is any update of the outcome of the works stated here, this could be mentioned or/and used for a comparative analysis