We sincerely thank the reviewers for their careful reading of our manuscript and for their valuable comments and suggestions. Their feedback was crucial to improve the structure and focus of the manuscript. In the revised version, we have clarified terminology, focused validation on high-energy events, summarized the results and provided additional explanations regarding the EWS's computational chain, site-specific conditions, and forecasting methodology. Detailed technical analysis such as extended validation and assessment of the EWS's forecasts, have been moved to Appendix. In the pages that follow, we provide a point-by-point response to all reviewer comments, indicating the corresponding changes made in the revised manuscript (as shown in the tracked changes file).

On behalf of all authors,

Antonis Chatzipavlis

The paper focuses on a highly relevant topic: Early Warning Systems for storm-induced overwash in a coastal region. While the paper is well written and structured, I believe three key topics require further attention from the authors before publication.

Key comments

(1) The first and most important of these is the correct definition of Overwash and Flood. In fact, the authors only evaluate "smaller" berm overtopping (traces) and "larger" berm overtopping (which they call overwash and flood) and never mention dune overwash or the flood/inundation of assets. I have made several comments on this issue below, and I think the authors should consider Sallenger's regimes for this study. They could even introduce a "berm overtopping" regime as a complement.

Response: We agree that the terminology required changes and additional clarification. In the revised manuscript, we explicitly specify that in the case of Poetto beach, episodic flooding occurs within the zone between the berm crest and the dune toe, as a consequence of berm overtopping. According to Sallenger (2000) storm impact scale, the study site falls within the *swash regime*, since no events exceeding the dune system and causing true overwash or inundation of back-beach assets have been recorded by the coastal videocameras. Consequently, our analysis focuses on the assessment of flood risk limited to the berm–dune zone. This is now clarified in Lines 166-172) and 182-187 of the revised manuscript, whereas consistent changes in terminology (distinguishing berm overtopping from dune overwash) have been applied throughout the text and the relevant figures (Fig. 8, 9 and 10).

(2) The authors should also clearly indicate the indicators to be used in the EWS regarding assets. While they recognise the necessity of indicators for the EWS in the Introduction, they never mention what thresholds are used or what warning levels have been defined (and why). This is never clearly stated throughout the paper. What warning levels exist, and how will they be sent to the first responders, managers, and the population?

Response: We agree that the description of the thresholds, indicators, and warning levels within the EWS required further clarification. In the revised manuscript, we have expanded the relevant section (lines 189–206) to explicitly describe the criteria used to define flood risk levels based on TWL exceedances relative to the local morphological threshold. Four warning levels (*Low, Moderate, High, and Very High*) are defined according to the degree of TWL exceedances, accounting for a 20% calibrated error margin derived from comparison with coastal video-monitoring data. Additionally, the revised text now explains how the NEPTUNE-EWS automatically issues e-mail alerts to first responders, whereas Figure 2 illustrates the operational workflow of the system and the process of issuing and disseminating warnings.

(3) The entire validation uses all data, including a large majority of low-energy conditions, whereas the EWS is designed for storms. Thus, the authors should clearly define the thresholds above which it is considered a storm for the purposes of this study, and then validate only values above these thresholds. Validation against mild or low wave conditions is not very meaningful when dealing with this type of work.

Response: We appreciate this important observation from the reviewer. The manuscript has been substantially revised to clarify the definition of high-energy conditions and to focus the validation analysis accordingly. Specifically, for validation of wind and sea level, high-energy were isolated by applying the storm criterion of Sanchez-Arcilla et al. (2008), selecting wind events from directions to which the beach is exposed (i.e. between 117° - 202°) and capable of generating increased hydrodynamic conditions (> 4 Bf). These events represent 11% of the 2.5-year available records (69 effective storm events; depicted on Figures A1, B1 and C1). For validation of wave parameters, only recorded significant wave heights exceeding 0.7 m, corresponding to the highest 16% of the 45-day records of the AWAC logger were considered. These revisions are reflected in Section 3.2 of the revised document with new text (lines 390-512 and 534- 584) and figures (Figure 5 and 7) specifically focused on the validation of these high-energy events. Extended validation for the full datasets is provided in Appendices A, B, and C, following also the structural suggestions made by Reviewer 2.

Detailed Comments

Detailed comments are provided below, including reinforcement of the points already mentioned.

Methodology

* Figure 1: Please include the name Poetto Beach in Figure 1b); Please include the acronym CAM after videocameras in the caption of Figure 1.
Response: The figure has been updated as suggested. The name "Poetto Beach" has been added to Figure 1b, and the abbreviations CAM1 and CAM2 have been included in the figure caption to refer to the coastal videocameras.

* Lines 149-151: Not easy to understand and split between these two concepts. Is overwash simply when the berm is surpassed by the run-up, and is flood simply when there is extensive water over the berm? Is that correct? Please be clearer. Why don't you use Sallenger's terminology for overwash and inundation? Why is the berm crest used instead of the dune crest? This is berm overtopping, not overwash.

Response: In the revised manuscript, we explicitly specify that in the case of Poetto beach, episodic flooding occurs within the zone between the berm crest and the dune toe, as a consequence of berm overtopping. According to Sallenger (2000) scale, the study site falls within the *swash regime*, since no events exceeding the dune system and causing true overwash or inundation of back-beach assets have been recorded by the coastal videocameras. Consequently, our analysis focuses on the assessment of flood risk limited to the berm–dune zone. This is now clarified in lines 150–155 (166-172) and 165-170 (182-187) of the revised manuscript, whereas consistent changes in terminology (distinguishing berm overtopping from dune overwash) have been applied throughout the text and the relevant figures (Fig. 8, 9 and 10).

What do you mean by "limited water coverage"? Is it the maximum limit inland reached by the water? The meaning of limited and extended water coverage is unclear, which is paramount to understanding the rest of the paper.

Response: To avoid confusion, the original line and its associated paragraph have been removed from the manuscript. In the revised text, the terminology has been updated from "limited or extended water coverage" to "limited or extensive berm overtopping". Specifically, limited overtopping refers to detected traces of partial water coverage or/and *Posidonia oceanica* washes, whereas extensive overtopping refers to flooding of the berm-dune zone, as observed on the frames of the coastal videocameras. This distinction is now reflected in (lines 182–187) of the revised manuscript.

above you have beta in italic for the foreshore slope, but not here. Please check for consistency. Please check all acronyms.

Response: The foreshore slope symbol has been standardized throughout the manuscript. All occurrences of $\tan\beta$ are now presented in italic and all acronyms have been checked for consistency.

Why is berm height used and not dune height? If there are dunes inland (even low ones), flooding should only occur after the dunes (or an obstacle where there are no dunes) have been overpassed, and not the berm. It is possible to have berm overtopping with no flooding for infrastructures. Please justify this choice.

Response: In Poetto beach the area between the berm and the dune toe is periodically flooded at specific locations (P1 and P2) during elevated surge events due to berm overtopping. These events cause damage and economic losses to the coastal facilities (mainly beach bars/restaurants) that are located within this berm-dune zone. Importantly, flooding rarely exceeds beyond the dune crest and no such events have been recorded by the coastal video monitoring system. This is now clarified in lines 166-172 of the revised manuscript.

Integrating measured bathymetry with the lower-resolution data from EMODNET is not simple, especially since EMODNet data often has relatively high vertical errors. How did you merge the data into the grid, and how were these potential vertical errors considered?

Response: The bathymetric grid used in SWAN simulations was generated by merging the retrieved high-resolution bathymetry and the EMODnet bathymetric database. To date, EMODnet represents the best available harmonised Digital Terrain Model (DTM) for European seas, showing however relatively high vertical errors in shallow depths. To overcome this and retrieve the best possible accuracy, the retrieved high-resolution nearshore bathymetry replaced the EMODnet values covering the same area. Moreover, a triangular interpolation where samples are organized into a Delaunay network allowed computation of the nearshore bathymetry (up to ~40 m depth) and generation of the final grid. This description has been added to lines 238-244 of the revised manuscript.

* Lines 223-226: Is this enough? Waves forecasted with these global models are often underpredicted during storms. It is most likely that the EWS will also underestimate the flood hazard. What has been done to minimise this problem and ensure that there is no underestimation?

Response: For Poetto beach, the downscaled wave height forecasts were found to be slightly overestimated during high-energy events (> 0.7 m) over the 45-day validation period, with a relatively low bias of - 0.06 m. This is now reflected in the revised wave validation section (lines 533-546) and in Figures 6 and 7. We acknowledge the reviewer's point that global models can sometimes underestimate wave heights during storms, and that wave period and direction may also exhibit significant discrepancies, as suggested also by Reviewer 2. This is now explicitly discussed in lines 525-527 and lines 943-948 of the revised manuscript, supported by relevant studies (e.g. Ruju et al., 2019; Alday et al. 2022; Oikonomou et al., 2023; Korres et al., 2021,2023). Furthermore, we suggest the inclusion of a future correction module to enhance the reliability of CMEMS and ECMWF forecasts for operational use (Lines 1200-1201).

* Lines 234-235: The tide gauge should only measure the sea level. How did you compare wind and wave data with data from the tide gauge?

Response: We have clarified the validation procedure. Wind forecasts were validated against a meteorological station at the port of Cagliari (elevation 3 m, 10-minute recording interval), using the closest ECMWF grid point located 7.2 km offshore (lines 250–270). Sea level forecasts (CMEMS Med-PHY) were validated against measurements from the tide gauge at Cagliari port, co-located with the meteorological station (lines 300–302). Wave forecasts (CMEMS Med-WAV) were validated against an AWAC logger deployed for 45 days at 18 m depth, located 2.1 km from the nearest CMEMS virtual buoy at a similar depth (lines 290–293).

* Line 246: Replace Ruju and Viola 2024 by Ruju and Viola (2024). Check the references throughout the text. There are some minor typos.

Response: The reference has been corrected accordingly and similar typographical errors throughout the manuscript have been addressed (line 314 of the revised document).

* Table 2: Table 2 is first included in the text and only mentioned afterwards. It should be the other way around. Please revise it for the final version. The explanation of the variables should not be inside the table. Some of the variables have already been explained in the text, so there is no need to repeat them.

Response: The text now precedes Table 2 to improve the flow of the manuscript. The table has been revised to avoid repetition by removing the last three lines (Ho, Lo, and tanβ – foreshore slope).

Results

The retreat is true, but only for the berm. The dune does not appear to have major changes in position (only in height). Therefore, it does not affect the entire profile.

Response: We agree with the reviewer. The manuscript explicitly distinguishes between shoreline and berm changes in both position and height, while noting that the dune exhibits only changes in height, without significant displacement (Lines 352–360 of the revised document).

Is this the reason for the flat and lowered berm profiles at P2 in the last two surveys?

Response: Indeed the flat and lowered berm profiles at P2 in last two surveys are linked with man-made works (bulldozing by beach operators) that took place during this period. This is now clarified in lines 356-357 of the revised document).

Why has the dune crest lowered? Wave overtopping? Wind removal without compensation at the front? Human trampling? This has never been explained or analysed, even though the paper is about overwash (runup exceeding dune height).

Response: The manuscript has been updated to clarify that the lowering of the dune crest is attributed to both anthropogenic and natural factors, including human trampling, limited vegetation cover, and insufficient sediment input from the beach fronting the dunes (Lines 360-361, revised document).

I disagree that this threshold is commonly used. In the vast majority of the existing works, the dune height is used as the threshold for coastal flooding. As can easily be seen from the two profiles, the berm can be overwashed, and there will still be enough berm for water percolation and infiltration before reaching the dune elevation. Flooding of human settlements will only occur after overwashing the dune crest and not the berm crest, unless there is significant occupation on top of the berm during winter, in which case, the risk to that occupation needs to be assessed. Is that the case?

Response: We thank the reviewer for highlighting this important point. We acknowledge the reviewer's point that dune height is commonly used as a reference threshold for coastal flooding (e.g. Armaroli et al. 2012; Harley et al., 2012;2016), which is now described in Lines 166-167. Nevertheless, in the case of Poetto beach there are several coastal facilities operating between the berm-dune zone, and which are periodically flooded due to berm overtopping, causing significant damages and disruptions. Importantly, flooding rarely exceeds beyond the dune crest and no such events have been recorded by the coastal video monitoring system. These clarifications are now given in Lines 167-172 and Lines 182-183 of the revised manuscript.

"flooding threshold values in NEPTUNE-EWS". This must be clearly explained. Why is he berm crest used instead of the dune crest? In Sallenger's regimes and all other derived works, it is always the dune elevation and never the berm elevation. If there is a dune between the road, the houses, and the waves, it should be the dune crest to be used, not the berm crest. If there is no dune, then you should use the maximum elevation before any human structure. If there are buildings between the dune and the berm crest, then two thresholds are needed: one for the flood of the infrastructure between the berm and the dune (which can be the berm crest) and another for the flooding of the assets behind the dune (the dune crest elevation). However, this simple approach (using berm crest) is not capable of distinguishing between different levels of berm overtopping, dune overwash or inundation, and therefore it does not work properly as an EWS.

Response: The use of the berm instead of the dune crest as flooding threshold reflects the specific site-conditions of Poetto beach, as explained in our response to the previous comment. The records of the coastal videocameras allowed the distinction between i) limited berm overtopping; and ii) extensive overtopping, led to flooding of the berm-dune zone at the identified vulnerable beach locations P1 and P2 (as referred in Lines 565-568). A descriptive example of this distinction evident for event S1 is provided in the text (Lines 670-703) and shown in Figure 8. the complete optical timeseries for all six studied surge events is archived on Zenodo (10.5281/zenodo.15426775), accompanying this work.

"the" and not them

Response: Typo has been corrected.

Would it be possible to correct the values or apply a correction that would suit and improve the data, particularly for storms?

Response: The reviewer refers to the mismatch observed during validation of the full wind dataset, which is now reported in Appendix A of the revised version. The newly added validation of the wind data during high energetic conditions showed a lower bias (0.7 ms^{-1}) and improved correlation. This is now sufficiently described in Lines 763-474 and shown on the new Figure 5 added. Although the ECMWF model is found to provide slight underestimations of wind velocity, these minor discrepancies have limited effect on the system's predictions. Moreover, it has to be considered that the meteorological station used for validation is located at the port of Cagliari, whereas ECMWF point of forecasts is located 11.2 km offshore the meteorological station. These considerations are further discussed in the discussion section (934-942 of the revised manuscript)

* Line 369 and data analysis for the 45 days period: The analysis was performed on the full 45-day dataset. However, for an EWS focused on storms, the analysis should cover a longer period and focus only on deviations during storms. The authors should have selected a threshold for storms (at least a simple Hs threshold) and analysed the data above that threshold only. From the records, I can see only one or two potential storms for the analysed period. The fit is very good for the highest storm, and better than what I would expect from global models at coastal locations.

Response: We acknowledge the reviewer's point for the limited time-window used for validation of the wave data. However, the available *in-situ* wave records in Poetto cover this 45-day period. Nevertheless, a threshold of 0.7 m was applied representing the higher 16% of all values. (Lines 533 – 536, revised document). Overall, wave heights were found to be slightly overestimated during high-energy events (> 0.7 m) over the 45-day validation period, with a relatively low bias of - 0.06 m. This is now reflected in the revised wave validation section (lines 533-546) and in Figures 6 and 7.

* Line 382: "low wave heights". Yes, but for a storm model, the concern with small wave heights is probably minimal. This is the reason why the analysis should focus on wave conditions with the potential to cause Overwash, rather than on all conditions.

Response: We consider this point to be addressed in our previous response, as validation is focused on higher-energy wave conditions (> 0.7 m), representing the higher 16% of the dataset. This is now reported in the newly added section "3.2.2 Wave validation".

* Line 397: Are these waves relevant to the analysis? Do they produce overwash? If not, why analyse them?

Response: We consider this point to be addressed in our previous response, as low-energy events have been excluded from the analysis. This is now reported in the newly added section "3.2.2 Wave validation".

* Lines 426-427: It seems that there is a 10 cm vertical difference between most of the shown records. However, the BIAS shows almost no difference (1 cm). So, why is there an underestimation during storms when the opposite should occur? The model is most probably not correctly predicting storm surge.

Response: We investigated this discrepancy by examining the sea level data during the 69 high-energy events, as identified by using the storm criterion of Sanchez-Arcilla et al. (2008). Our findings suggest slight underestimations of the CMEMS forecasts with a mean residual of 0.04 m for these events. This is now analytically reported in added Lines 506-515 and Figure 5 of the revised document.

<u>* Line 438:</u> What happens when you average the entire set of storm data? Is there a vertical difference? According to the BIAS, the difference for the entire series is almost 0. Is it the same for storms, or is the model unable to correctly reproduce the storm surge? For a model that considers storms, the analysis should focus mostly on storm data and not on all data.

<u>Response:</u> <u>Response:</u> We consider this point to be addressed in our previous response, as we have now included the validation during energetic conditions. Extended validation of the whole dataset is now reported in Appendix B.

<u>* Lines 439-440:</u> "Overall, the forecasted sea level values used by NEPTUNE-EWS are found to be slightly underestimated, but in a minor way that has limited effect on the early warning system". I believe so, but wouldn't it be better to apply a correction?

<u>Response:</u> This part of the text is now revised showing the results during energetic events, whereas validation of the whole dataset is now reported in Appendix B. Considerations are further discussed in the discussion section (934-942, revised manuscript), whereas a correction module for improvement is discussed in Lines 1113 and 1200-1201 of the revised document.

<u>* Table 4:</u> You have some values of mean Hs of 0.6 m to 0.8 m and periods of around 6-7 s. Is that really a storm capable of generating Overwash, or is it because you defined the berm crest as the limit for Overwash, meaning that lower waves can produce overwash events?

<u>Response:</u> The low values of mean Hs between (0.6 and 0.8 m) correspond to the days immediately before and after the storm peaks. This approach was intentionally adopted to evaluate the EWS performance more comprehensively by capturing the full evolution of each event, including onset and decay phases.

<u>* Figures 14, 15 and 16 (caption):</u> I think this is the first time you have referred to a Flood State, mentioning classes No trace, Overwash and Flood. What do these terms mean? You defined these terms in a very simple (and not entirely accurate) way at the end of 2.1 study Area, which is strange. This definition is the most important in the paper and should not be at the end of the Study Area. You should clearly state what a Flood State (or Regime, since not all conditions lead to flooding) means. You should use Sallenger's regimes. Does Flood correspond to Inundation in the Sallenger regime? I don't think so. How do you differentiate between Overwash and Flood? I think you are only analysing berm overtopping and not true dune Overwash, as it should be.

<u>Response:</u> The figures and their captions have been revised accordingly, and new text has been added to clarify the definitions and terminology (lines 671-673), revised manuscript.The records of the coastal videocameras allowed the distinction between i) limited berm overtopping; and ii) extensive overtopping, led to flooding of the berm-dune zone at the identified vulnerable beach locations P1 and P2 (as referred in Lines 565-568). A descriptive example of this distinction evident for event S1 is provided in the text (Lines 670-703) and shown in Figure 8. the complete optical timeseries for all six studied surge events is archived on Zenodo (10.5281/zenodo.15426775), accompanying this work.

<u>Lines 491-581:</u> The event-by-event description is tedious and should be simplified, probably using a table to summarise these results.

<u>Response:</u> The event-by-event description has been condensed and simplified in the main text (Lines 698–857, revised manuscript), while the full analytical descriptions are now provided in the newly added Appendix D.

<u>* Line 499 and figures 14 and 15</u>: In figures 14 and 15 you distinguish between overwash and flooding. Here, however, you combine both. Why? Why are you not adopting Sallenger's terminology and concepts?

<u>Response:</u> The reviewer refers to figures 8 and 9 of the revised document. Terminology has now been modified accordingly thorough the manuscript distinguishing between events of limited and extended overtopping. Sallenger's scale is not adopted in the case of our study area as we refer to berm overtopping.

Discussion

<u>* Lines 638-640:</u> This is true, but in general, I would say that for beaches that are exposed to wave action, the runup must be included. In sheltered areas such as estuaries and backbarriers, however, this is not necessary because small waves with short periods do not contribute to effective flooding. Therefore, in this case, it is important to include the wave-induced runup. This does not necessarily contradict other studies, but rather represents the correct definition of what is relevant for each environment.

<u>Response:</u> We agree with the reviewer's observation. This point, referring to Lines 1718–1720 of the revised manuscript.

<u>* Line 641:</u> "flooding predictions". This could refer to berm flooding or berm overtopping. You never mention flooding of roads or structures, even though the entire introduction points to that. I would say that the introduction of this paper misleads the reader, and the definitions are not clear enough to effectively state that you are only analysing berm overtopping and not dune overwash or infrastructure overwash/flood. This must be clear throughout the paper.

<u>Response:</u> Additional clarifications have been incorporated into the Study Area section to explicitly state that, for this specific case study, the analysis focuses on episodic flooding occurring within the berm–dune toe zone as a result of berm overtopping (Lines 166-172 of the revised document).

<u>Lines 643-657:</u> This description from line 643 until 657 partially repeat the results.

<u>Response:</u> We agree with the reviewer. To avoid repetition and confusion for the reader, this part of the text has been removed.

Conclusions

* Line 689: This paper presents the validation of a potential warning system, rather than the warning system itself. All the data and discussion is about the validation, not the EWS and its implementation.

Response: In the revised manuscript, significant structural changes have been implemented. The validation section has been condensed, while a more detailed description of the EWS including its computational chain and forecast implementation, has been added to better emphasize the system itself rather focusing on the validation.

* Line 700: You are forecasting berm overtopping, not coastal flooding. These are quite different.

Response: We believe this distinction has now been sufficiently clarified after justifying the use of the berm height for the specific site-conditions of Poetto beach. The terminology throughout the manuscript has been revised accordingly.

* Line 701: The meaning of a storm event for the region was never defined, and this should be done before any data analysis. The data analysis should focus primarily on the storm events and exclude calm conditions.

Response: We believe that this aspect has been addressed by excluding the low-energy conditions from validation of the wind, wave and sea level data. As addressed in detail in responses of previous comments.

This is definitely a good quality work that merits publication. However, while I am positive, I believe that the structure and presentation can be improved substantially. The paper presents efforts towards a potential EWS and not the EWS itself. There is a lot of information on validation and the paper is rather long. For that reason, I see a lot of potential to produce two papers and that would be my first suggestion if the authors decide to take that way:

• One paper presenting the modelling chain and the validation results. I think this is the most solid part and I would be happy to review a revision, which focusses only on that aspect of the work

• Second paper discussing about how the model output is converted to early warnings and their accuracy. I believe that here there still needs to be done some additional work in order to present something more robust and closer to the current state of the art.

Obviously, the authors can choose to try to include everything in one paper, but they need to try to prepare something easier for the reader to follow. I believe that to do that, they need to try to focus only on the important information, and even move some technical parts in the appendix (like extended model validation, different run up formulas, etc).

Response: We value the reviewer's constructive suggestions for improving the structure and presentation of the manuscript. In the revised version, the manuscript has been substantially restructured: Validation results have been summarized, focusing on high-energy events, as also suggested by Reviewer 1. Description of the computational chain of the EWS and the specific site conditions of the study area are clarified and sufficiently detailed in Methodology (Section 2), following appropriate text modifications. The event-by-event assessment of the EWS's performance during the surge events has been condensed to make it easier for the reader to follow and interpret. Analytical technical parts are now presented in Appendix. Specifically, extended validation of the wind, sea level and wave climate is reported in Appendix A, B, and C, respectively, while a comprehensive assessment of the EWS's performance during the identified surge events is reported in Appendix D.

The system is based on forcing from other models to run SWAN. The model performs rather well in the validation for Hs, but is less accurate for the period, direction and other secondary parameters. This is a known issue of such wave models and there are several previous studies reporting similar performance that could be cited to support to present effort.

Response: We have addressed the known issue of the discrepancies in validation of the wave climate. Clarifying text has now been added (Lines 520 – 523) and citing relevant studies that report similar findings for wave model performance(Ruju et al., 2019; Alday et al., 2022; Oikonomou et al., 2023; Korres et al., 2021, 2023).

The authors use empirical run up formulas to assess potential overtopping events and in this part there are a lot of aspects that merit clarification. First of all, since there is a video system on-site that could at least provide run up maxima, the authors could validate the different fomulas, or even fit a site-specific one. Second, any discrepancy between observed and estimated R2 can be the result of a combination of model/formula error, beach face slope inaccuracy and the variable capacity of the changing bathymetry to attenuate incoming wave energy. For all of the above there is a substantial amount of literature, that's why I am in favor of splitting the manuscript in two separate papers.

Assuming that the use of the video imagery information can solve the run up estimation accuracy and that the beach face slope doesn't change a lot, as stated in the manuscript, the authors need to address the effect of the nearshore bathymetry. I add below some examples from relevant literature that could help (the order is chronological and the list is not exhaustive):

https://onlinelibrary.wiley.com/doi/abs/10.1002/esp.2264

https://core.ac.uk/download/pdf/36733581.pdf

https://www.sciencedirect.com/science/article/pii/S1674237019300833

https://www.sciencedirect.com/science/article/abs/pii/S0141118722003601

There is a lot of literature about how to estimate accurately the wave run up and the authors discuss some. It is not clear that using a model like XBeach could improve the performance if the bathymetry is not updated (which is practically not possible for known reasons), but that could be investigated. In case there are data, the authors could have estimates of potential seasonal profiles that could be used. Also if the wave run-up cannot be accurately estimated then the authors could consider issuing warnings based on some percentile threshold of the TWL and not the berm/dune height itself.

Response: We acknowledge the reviewer's point about validation of the run-up formulas based on video imagery, Precise positions of the run-up maxima could not be determined due to the lack of appropriate optical metadata. However, we consider that the use of the raw frames is sufficient for validation of berm overtopping in the context of this study. Regarding the control of nearshore bathymetry on wave run-up, and the inclusion of Xbeach. New text has been inserted in the discussion section addressing this important aspect. The empirical formulations, while calibrated against local observations, assume a static beach profile that is periodically updated by field surveys in the EWS. To account for seasonal variations, representative seasonal profiles (e.g. winter/summer profiles) can be extracted from topographic data of adequate temporal coverage, using statistical and clustering techniques (e.g. Zhao et al., 2020; Athanassiou et al., 2021; Otto et al., 2021). Such techniques can identify seasonal changes and trends of major nearshore topo-bathymetric features (i.e. dune, berm and bars), driving the run-up. We highlight the need of considering the site-specific variability of the nearshore bathymetry and its features in Lines 1027-1032 of the revised manuscript. The latter approach may prove to improve forecast realism, without the need for frequent updates of the beach topo-bathymetry. We also agree that the inclusion of Xbeach may not improve the performance of the system as despite the fact that run-up could be simulates more dynamically, the benefits are constrained if nearshore bathymetry is not frequently updated, which is

practically challenging. Nevertheless, this is an effort suggested for experimenting. This is now analytically discussed in lines 1092-1100 of the revised manuscript.

Finally, if they combine all sources of incertainty (e.g. range of slope, range of R2, range of dune height) they could issue forecasts in a probabilistic framework (i.e. probability of overtopping, rather than a binary outcome).

Of course all the above are suggestions that I hope could be helpful.

<u>Response:</u> Regarding issuing forecasts in a probabilistic framework. We suggest the incorporation probabilistic methods or/and deterministic ANNs to enhance forecast reliability under uncertain wave and surge conditions, similar to the approach of Zampato et al., (2016). This is discussed in lines 1110-113 of the revised document.