Dear Editor and Reviewers,

Please find below, our response to the reviewers comments regarding the manuscript entitled, “The Influence of Infragravity Waves on the Safety of Coastal Defences: A Case Study of the Dutch Wadden Sea” for publication in Natural Hazards and Earth System Sciences.

We appreciate the time and effort exerted by the reviewers on our manuscript. We found their questions and suggestions very helpful, as they have highlighted areas where we can further clarify our writing and explanations. In the following paragraphs, we detail our position (in blue) regarding the modifications we have made to the manuscript in response to the reviewers’ comments (in black).

Comments from Reviewer #1:

1. Abstract, line 12: I don't think these foreshores have in many cases been incorporated in the design.

   The reviewer is correct that neither effects of foreshore vegetation nor the impact on IG waves are currently considered. However, the effects of the foreshore bathymetry on incoming wind-sea and swell waves is indeed taken into account. This is considered the standard practice in the Netherlands and is mentioned in the subsection "Foreshore Scenarios" in Section 2.2.5.

2. Abstract, line 12: overtopping volume, so not binary

   Yes, we have modified the text. Line 12 now reads: " …are expected to reduce the likelihood and volume of waves overtopping the dikes behind them ...".

3. Abstract, line 22: not position relative to the inlets and channels?

   That is correct that there is correlation with the inlets, since areas behind inlets are exposed to larger waves than areas behind the islands. We have modified the text. Lines 20-22 now read:

   "The spatial variation in this effect, observed for the case considered, highlights its dependence on local conditions—with IG waves showing greater influence at locations with larger offshore waves, such as those behind tidal inlets, and shallower water depths."

4. Line 30: I would say "the foreshore contributes to a reduced load and therefore reduced likelihood of failure"

   We have modified the text. Lines 31-32 now read: “The foreshore contributes to a reduced wave load at the structure and is therefore expected to reduce the likelihood of failure/”

5. Line 31: ref Dalrymple, mendez and Losada.

   We thank the reviewer for the suggested references. We have added the following:

Dalrymple, R. A., Kirby, J. T., and Hwang, P. A. (1984). "Wave Diffraction Due to Areas of Energy Dissipation." Journal of Waterway, Port, Coastal, and Ocean Engineering, 110(1), 67-79

Mendez, F. J., and Losada, I. J. (2004). "An empirical model to estimate the propagation of random breaking and nonbreaking waves over vegetation fields." Coastal Engineering, 51(2), 103-118.

6. Lines 44-45: maybe not explicitly, but implicitly they have been in empirical formulas which relate the offshore forcing parameters to overtopping, and thereby include the intermediate processes.

The reviewer is correct that empirical formulae based on offshore forcing parameters would implicitly account for the intermediate processes. Therefore, risk assessments that make use of such formulae would indeed account or the effects of infragravity (IG) waves.

The formulae used in the Netherlands (and the wider Europe, see EurOtop Manual) are based on parameters at the toe. For conditions with shallow foreshores, these formulae e.g. Van Gent (1999) would indeed already account for IG motions. However, it is standard practice to use a phase-averaged wave model, like SWAN, to obtain the boundary conditions at the dike toe. As these models tend to exclude IG wave dynamics, these waves are often not considered in the analysis.

We have added the following text to Lines 49-52: "An empirical approach developed for shallow foreshores using offshore forcing parameters would implicitly account for the intermediate processes, including the effects of IG waves (Mase et al., 2013;Goda, 2010). However, the more widely-used formulae are based on parameters at the structure toe (EurOtop, 2018)."

References:

Goda, Y.: Random Seas and Design of Maritime Structures. World Scientific Publishing, 2010.

Mase, H., Tamada, T., Yasuda, T., Hedges, T. S., and Reis, M. T.: Wave Runup and Overtopping at Seawalls Built on Land and in Very Shallow Water, Journal of Waterway, Port, Coastal, and Ocean Engineering, 139, 346-357, 10.1061/(asce)ww.1943-5460.0000199, 2013.

7. Line 46: but it is corrected for if empirical formulas are used.

The reviewer is correct. In this manuscript, we make use of such an empirical formula to correct or the infragravity waves.

8. Line 60: i am sure there is a classical reference.

We have replaced the reference with the following:

Van Gent, M. (1999). "Physical model investigations on coastal structures with shallow foreshores: 2D model tests with single and double-peaked wave energy spectra." Hydraulic Engineering Reports, Delft.. ”

Hofland, B., Chen, X., Altomare, C., and Oosterlo, P.: Prediction formula for the spectral wave period T m-1,0 on mildly sloping shallow foreshores, Coastal Engineering, 123, 21-28, 10.1016/j.coastaleng.2017.02.005, 2017.

9. Line 68: ref.

We have added the following reference:

Roelvink, D., Reniers, A., Van Dongeren, A., Van Thiel de Vries, J., McCall, R., and Lescinski, J. (2009). "Modelling storm impacts on beaches, dunes and barrier islands." Coastal Engineering, 56(11-12), 1133-1152.

10. Lines 70 -72: was the problem the "case" (ie. combination of profile and/or forcing conditions, or the "method" ie using xbeach.

Our findings suggest that the problem was "method" specific; not due to the use of XBeach Surfbeat (numerical model) but due to the use of inappropriate empirical formulae to estimate the actual overtopping discharge. This is discussed in detail in Section 4.2 along with the findings when different overtopping formulae are applied.

We have modified the text (lines 74-76) for clarity: “This rather striking finding requires further investigation; particularly, to determine if the large IG-wave influence reported by Oosterlo et al. (2018) is valid or if it was merely an artefact of the method used.”

11. Line 82: contrasting?

We chose the word "inconclusive" with reference to the work of Oosterlo et al. (2018) that suggested that further work was needed before the influence of the IG waves could be described with confidence. We have added this citation to the text for clarity:

Oosterlo, P., McCall, R., Vuik, V., Hofland, B., Van der Meer, J., and Jonkman, S. (2018). "Probabilistic Assessment of Overtopping of Sea Dikes with Foreshores including Infragravity Waves and Morphological Changes: Westkapelle Case Study." Journal of Marine Science and Engineering, 6(2), 48.

12. Figure 2: aren't SS parameters and vaklodingen input into the empirical formulas?

We thank the reviewer for highlighting this. We have adjusted the figure to be more linear so that inputs are better described.

13. Line 129: plus: storms occur in the winter season with little vegetation present

That is correct. We have modified the text so that lines 129-132 now read:

“Lastly, wave attenuation by vegetation is not included in the probabilistic analysis because: i) storms generally occur in the winter season, where there is little vegetation present; and ii) it is very likely that almost all vegetation present will flatten or break under extreme forcing (Möller et al., 2014;Vuik et al., 2018a)”

14. Figure 3: what is there to reverse shoal on etabar?

We thank the reviewer for highlighting this. We have modified the figure so that the reverse shoaling is clearly applied to the significant wave height alone.

15. Figure 3: typo

We thank the reviewer for spotting this! We have corrected the typo.

16. Figure 3: write out the caption, it should be understandable without having to read the text.

For clarity, we have expanded the caption to describe each of the parameters.

17. Line 141: ref

We have added the following reference:
Booij, N., Ris, R. C., and Holthuijsen, L. H. (1999). "A third-generation wave model for coastal regions - 1. Model description and validation." J Geophys Res-Oceans, 104(C4), 7649-7666:

18. Line 143: in this paper? why? you add an empirical relation which should cost no comp effort.

The reviewer is correct. The added empirical relation did not have a significant impact on the computational demand. However, the probabilistic model itself (FORM) does require computation effort--around 10 minutes per dike section. We have modified the text as follows (now lines 147-149):

“To reduce the overall computational time of the probabilistic calculations (around 10 minutes per dike section), the output locations were reduced to one every 1.5 km.”

19. Line 144: but if the waves at the location of the hydra output were already breaking, you underestimate the offshore wave height by simply reverse shoaling with linear theory.

The reviewer is correct. A check was made to see whether the waves at the hydra output location were breaking by comparing the actual ratio of wave height to water depth to the breaker index calculated using the Battjes and Stive (1985) formula (Equation 4 of this manuscript). The actual ratio of wave height to water depth for the Hydra output locations varied from 0.133 to 0.51 (mean of 0.37) while the estimated breaker index varied from 0.68 to 0.87 (mean of 0.79). This suggests that the assumption of non-breaking waves may be considered reasonable.

We have modified the text so that lines 155 to 157 now read: “The assumption of non-breaking waves is considered reasonable since the average ratio of the wave height to water depth at the Hydra-NL output location was 0.37, while the average ratio for breaking waves—referred to as the breaker index—was estimated as 0.79 (using Equation 4).”

20. Lines 148 – 149: so you are comparing output at two locations?

Yes, that is correct. We compare estimates at the dike toe to estimates approximately 1 km offshore (obtained by reverse shoaling). We have amended the text to improve clarity. Lines 153-154 now reads:

"Compared to the Hydra-NL estimates (a few hundred meters offshore), the reverse-shoaled significant wave heights (to 1km offshore) were…"

21. Table 1: m

We thank the reviewer for spotting this! We have added "m" to the table.

22. Lines 162-163: no? obliqueness is a function of the orientation of the contour lines, so for dike sections that are not perpendicular to NW-SE you still get refraction

The reviewer is correct. Our approach assumes a perpendicular dike orientation. This results in a milder slope than if the transect was considered from North to South. This approach is considered comparable (in terms of correctness) to assuming obliqueness and using a "correction" factor in the empirical overtopping model, which were developed for perpendicular wave attack.

We have added the following text to describe this (lines 171-173): "It should be noted that assuming a NW to SE orientation results in an artificially milder dike slope for dikes that are not perpendicular to the assumed transect. This is taken into account in the average dike slope calculation."

23. Line 164: if your transect is not perpendicular to the contour lines you will get artificially milder slopes

The reviewer is correct. See our response to comment #22 above. We have taken this into account in our average dike slope calculation..

24. Line 166: what is this based on?

This is based on the value used by Vuik et. al (2018) to represent the spatial and temporal variations in the bathymetric data of the Dutch Wadden Sea. We have added the following citation:

Vuik, V., Van Vuren, S., Borsje, B. W., van Wesenbeeck, B. K., and Jonkman, S. N. (2018). "Assessing safety of nature-based flood defenses: Dealing with extremes and uncertainties." Coastal Engineering, 139, 47-64.

25. Line 168: what this exceeded in the data? Ie where there milder slopes?

The calculated foreshore slopes ranged from -0.04% to 4% with an average of 0.14%; where a negative slope indicates a slight downward slope towards the dike.

Our approach is in-line with the standard assumption (e.g. CLASH database (Steendam et al., 2004) and Keimer et al., 2021) that slopes milder than or equal to 1/1000 (0.1%) are all considered (near) flat or horizontal. We have added the following text to clarify this (lines 179-182:

"Given the range of validity of the empirical formulae (Equations 12 and 18), a minimum foreshore slope of 1/1000 (or 0.1 %) is considered here. This is in-line with the common approach where slopes milder than or equal to 0.1% are all treated equally as (near) flat (Keimer et al., 2021;Steendam et al., 2004). Note that the calculated foreshore slopes ranged from -0.04% to 4% with an average of 0.14%, where a negative slope indicates a slight downward slope towards the dike toe."

References:

Keimer, K., Schürenkamp, D., Miescke, F., Kosmalla, V., Lojek, O., and Goseberg, N. (2021). "Ecohydraulics of Surrogate Salt Marshes for Coastal Protection: Wave-Vegetation Interaction and Related Hydrodynamics on Vegetated Foreshores at Sea Dikes." 147(6), 04021035

Steendam, G. J., Van der Meer, J., Verhaeghe, H., Besley, P., Franco, L., and Van Gent, M. (2004). "The International Database on Wave Overtopping." Coastal Engineering 2004, 4301-4313.

26. Figure 4 caption: typo

We thank the reviewer for spotting this! We have corrected the typo.

27. Figure 4 caption: how is the average slope determined?

The average slope is determined as a best-fit line (least squares method) of the foreshore elevation data points from the dike toe to 1km offshore. We have added text to clarify this (lines 176-177):

“The average slope was determined as best-fit line (least squares method) considering the foreshore elevation data points between the dike toe and 1km offshore.”

28. Lines 180-182: but that fact would lower the significance of including IG waves. So you should reflect on this in the discussion

Actually, increasing the dike crest actually increases the influence of the IG waves, since the higher load needed for failure is reached earlier when IG waves are included. This is indeed discussed in Section 3.2.2 in our analysis of the effect of different crest levels.

29. Line 185: ref Booij et al 1999

We have added the following reference:

Booij, N., Ris, R. C., and Holthuijsen, L. H. (1999). "A third-generation wave model for coastal regions - 1. Model description and validation." J Geophys Res-Oceans, 104(C4), 7649-7666.

30. Equation 2: I thought it upgraded to Baldock formulations?

The Battjes and Janssen formulations remain the default in SWAN version 41.31A (used in this study) according to the user manual obtained here:

http://swanmodel.sourceforge.net/download/zip/swanuse.pdf.

31. Line 197: sqrt missing

We thank the reviewer for spotting this; we have corrected the equation.

32. Line 198: and a mean of?

The mean value is determined using Equation 4 above. We have amended the text as follows for clarity (line 212):

"...standard deviation of 0.05 and mean value determined using Equation 4."

33. Equation 8: why these pre-determined frequency ranges? I would let that depend on the peakfrequency and then let the IG range go to 0.5*fp. see also your figure 16 which shows that the split frequency should be at f=0.1 Hz.

We agree with the reviewer that for conditions with a single visible high-frequency peak, that 0.5*fp is the more accurate separation frequency. However, in the Dutch Wadden Sea, the wave spectra typically show multiple peaks. In order to avoid contaminating the infragravity signal with that of swell (around 10 s or 0.1 Hz), a separation of 0.05 Hz is typically used.

We have included an example of measured wave spectra to highlight this in Appendix B. Lines 229-234 now read:

"It should be noted that for conditions with a single, clearly-defined peak frequency (f_p), the frequency separating IG and SS motions is typically taken as f_p⁄2. However, as wave spectra in the Dutch Wadden Sea typically show multiple peaks, a separation frequency of 0.05 Hz is typically used to avoid contaminating the IG signal with that of swell (with periods around 10 s or 0.1 Hz). This choice of split frequency is consistent with previous studies in the area (Engelstad et al., 2017;De Bakker et al., 2014) and coincides with the minimum in spectral density in the observed wave spectra at the field site (Figure B. 1, Appendix B)."

34. Line 215: true, but for design conditions, the peak period increases so the split freq will lower

As mentioned in our reply to the comment #33 above, we agree that the separation frequency should be dependent on the high-frequency peak for uni-modal spectra; however, for the Wadden Sea where spectra typically have multiple peaks, it is considered more accurate to have a fixed separation frequency. This is to avoid mixing swell and infragravity waves.

35. Equation 10: what is the effect of neglecting the variance above 1 Hz?

There is little to no variance above 1 Hz. Therefore, there is no impact on the calculation. We have included an example of measured wave spectra to highlight this in Appendix B.

36. Line 219: ref

We have added the following reference:

Smit, P., Stelling, G., Roelvink, D., van Thiel de Vries, J., McCall, R., van Dongeren, A., Zwinkels, C., and Jacobs, R. (2010). "XBeach: Non-hydrostatic model." Report, Delft University of Technology and Deltares, Delft, The Netherlands.

37. Equation 11: the factor H_IG should be dimensionless

The reviewer is correct. H_IG is indeed dimensionless. The factor (0.36) has the dimension m^-0.5, which was previously not stated clearly. We thank the reviewer for spotting this oversight and we have corrected the description in the text.

38. Equation 11: is this the incoming component only?

This could be either the incoming component alone or incoming+reflected. It depends on the gamma_d factor which accounts for reflection depending on the structure slope.

We have now added the expressions for each of the individual factors (Equations 13 to 17).

39. Equation 12: so \tilde{H_IG} has a dimension m^{1/2}?

It is dimensionless. See our response to comment 37 above.

40. Line 225: what are the values of these factors?

These factors can have values ranging from 0 to 2.5 depending on the various environmental conditions. We have now added text and formulae to describe how each factor is determined (Equations 13 to 17).

41. Line 240: how much more gentle?

We have modified the text as follows (line 262):

"..typically 1:500 to (near) flat...".

42. Equation 13: dimensionless?

Yes, it is dimensionless.

43. Line 299: what about bottom dissipation?

Bottom dissipation is indeed present; however, this is expected to be minor compared to wave breaking. For clarity, we have added the following text (lines 331-332):

"Note that the influence of bottom roughness is included in both scenarios and represented by default parameter values in the numerical model."

44. Figure 7: H_ig IS equation 12. what is the difference in input between the open circle and the solid?

    We thank the reviewer for highlighting this. We agree that the previous description was confusing and we have adjusted the manuscript to improve clarity.

    H_IG is a stand-in for H_m0,IG,toe/H_m0,SS,toe, which can also be derived from the measurements (Equation 11 for definition). It may also be estimated empirically using Equation 12. We have adjusted Equations 11 and 12 and their descriptions to better reflect this.

45. Figure 7: how can \tilde{Hig} be measured? Does it take the observed wave heights? What about the gamma factors? how are they observed?

    See our response to comment #44 above. We have adjusted the text to clearly describe how H_IG may either be derived from the measurements (Equation 11) or estimated using the empirical equation (Equation 12).

46. Figure 9: why a histogram and not a table?

    Since the failure probability is typically quoted in powers of 10 (e.g. 10^-3 or 10^-6), we felt it would be easier to visualize the differences if plotted on a logarithmic scale.

47. Figure 9: what about SS+veg?

    We have now added SS+veg to Figure 9.

48. Figure 11: this figure undermines your thesis that IG waves are important factors that are missing as the alpha values are small. This would be the takehome for me, so please correct the reader for drawing this conclusion

    We thank the reviewer for this comment. The sensitivity factors only describe how sensitive the calculation is the uncertainty in the parameters (described by their standard deviations). It does not, however, account for the sensitivity of the results to the mean value.

    Since the plot may distract the reader from the main results, we have moved this analysis to Appendix D.

49. Line 411: parentheses

    We thank the reviewer for spotting this typo. We have corrected it.

50. Line 418: see my earlier comment. repeat this statement above

    We have adjusted the text above (Abstract) to reflect the impact of the inlets.

51. Line 420: then why is the pf higher in the extreme west?

Points in the west are exposed to higher levels of wave forcing. This is made visible in Figure 13a where points with higher Pf have higher values of H_m0,deep^2 x T_m-1,0,deep. This is explained in lines 440-443:

"In **Error! Reference source not found.**a, an offshore forcing parameter ($H_{m0,deep}^2 T_{m-1,0,deep}$), which is proportional to the offshore energy flux, is used to represent the combined influence of $H_{m0,deep}$ and $T_{m-1,0,deep}$. In **Error! Reference source not found.**b, the influence of variations in $\bar{\eta}$ and the bed level at the toe ($z_{b,toe}$) are represented by $h_{toe} = \bar{\eta} - z_{b,toe}$. The calculated P_{f1} shows a strong positive relationship with $H_{m0,deep}^2 T_{m-1,0,deep}$ (R^2= 0.65), meaning that higher forcing results in higher failure probabilities."

52. Figure 12: why is it dark red here?

The influence of the IG waves depends on several factors, as shown in Equation 12. The dark red area shown to the West in Figure 12 corresponds to two areas exposed to higher offshore waves (>2.4 m, while most of the other areas typically have waves <2 m) and very shallow water depths at the dike toe (>1.4 m, while most of the other areas have water depths > 3 m). This water depth is a function of the still water level and the (high) bed level at this location.

We have added the following text (lines 471-473): "This is also seen by the areas with higher IG-wave influence to the West in Figure 11b, which correspond to points with higher offshore waves (> 2.4 m, at the design point) and shallower water depths (<1.4 m), compared to the other locations (with offshore wave heights typically < 2m and water depths at the toe > 3m)."

53. Line 432: why is this, as \bar{eta} was the largest contributor?

The sensitivity factors only describe how sensitive the probabilistic calculation is to the uncertainty in the parameters. It does not describe the correlation between the mean values and the resulting failure probabilities. We note that the discussion on the alpha values distracts from the main focus of the paper; so we have moved it to Appendix D.

54. Line 436: what are the parametric differences between saltmarshes and mudflats?

The characteristic differences are that: 1) saltmarshes are typically found in areas with less exposure (lower offshore wave heights and periods); and 2) saltmarshes typically have higher bed levels, which may translate to lower water depths at the dike toe, depending on the still water level.

55. Figure 13 caption: and what are the colors?

The colors represent the mudflats vs saltmarshes. This is listed in the legend of the figure.

56. Line 447: why not show this in a similarly styles plot as figure 14a?

As suggested, we have added this as subplot 'c' in now Figure 13.

57. Line 456: please go into details about some of the large and small values, and separate out batyy from forcing

We now describe the specific "hot spots" in more detail. See our response to the reviewer's earlier comment (#52).

58. Line 461: relative, not in absolute sense

The reviewer is correct. We have modified the text as follows (lines 470-471):

" Therefore, areas with low water depths at the toe relative to large offshore waves are expected to have a greater IG-wave influence on T_(m-1,0,toe). "

59. Line 497: And it takes the boundary conditions at the toe, so the transformation is not included

The laboratory experiments used to derive the equations for shallow foreshores (Van Gent 1999) did include the transformation. The problem comes when models that do not correctly account for the transformation are used to obtain the boundary conditions at the toe.

60. Line 501: I would leave this part off. It is not essential for your takehome and message

We acknowledge the reviewer's suggestion; however, the authors consider this discussion critical, since the choice of overtopping model significantly affects the calculated failure probability and explains the differences found here, compared to that of Oosterlo et. al (2018). We have added the following text to make this connection clearer and more relevant (lines 548-550):

"As Oosterlo et al. (2018) applied the EurOtop (2007) formulae to a dike with an average slope of 1:8, it can be concluded that the large influence of the IG waves—where including IG waves increased the failure probability by 10^3—reported by the authors was indeed due to the method used."

61. Line 541: I don't understand. H_IG should account for the bottom slope as through equation 12.

H_IG does indeed account for the bottom slope's influence on the height of IG waves compared to that of the SS waves. However, Tm-1,0, by definition, gives a disproportionate amount of weight to energy at lower frequencies and is very dependent on the spectral shape. Therefore, while the ratio of the wave heights may be the same for different foreshore slopes, the spectral shape varies (Figure 15) and so too does the value of Tm-1,0 (Table 4). This is also described in Appendix C. To make this clearer, Appendix C is now referenced on line 561.

62. Lines 568-570: this needs more ink. was it case (as in site/forcing) specific or method specific or both? I would suggest to spend an entire subsection on the comparison with Oosterlo.

We agree with the reviewer that comparison with Oosterlo is important. This comparison is actually done in Section 4.2. To make this connection clear, we have added text to lines 540-542 (see response to comment #60 above).

63. Line 578: I would argue it from the other side. Given a crest level of 8 m, in both scenarios the dike does not fail. But with what increase sea level does it start to fail under either scenario?

Remarking that HIG decreases relatively when the water depth increases. So don't look at a hypothetical crest height but the actual (higher) one.

Our work focused on the required crest level and not on different sea-level rise scenarios.

Though sea-level rise was not the focus of the current paper, a rise in sea-level can be interpreted as a reduction in crest level—without taking into account the influence of wave breaking. As the reviewer pointed out, HIG decreases with increasing water depth; however, the sensitivity of the probabilistic calculation to the an increase in water level (and reduction in crest level) would result in higher failure probabilities. This coupled with the fact that higher sea levels would result in less wave breaking and higher Hm0,SS at the toe, it is likely that the influence of HIG on the calculation would decrease.

64. Line 602: Make this is stronger statement. What has been found? A lot of readers only read the abstract and conclusions.

We have re-written the Conclusions so that it is stronger and more concise.

65. Line 611: this is the conclusion

We agree with the reviewer. See our response to comment #64 above.

66. Line 622: combined with this statement

We agree with the reviewer. See our response to comment #64 above.

67. Line 634: these are not conclusions but recommendations

That is correct. See our response to comment #64 above.

Comments from Reviewer #2:

1. General: In my opinion the manuscript fits well with NHESS and addresses relevant scientific questions for the journal audience and even wider readers. This manuscript clearly provides a relevant contribution to natural hazards and their study, namely by analysing the impact of infragravity waves in coastal structures.

We thank the reviewer for this comment.

2. General: The scientific quality is in my opinion excellent. I have followed the discussion between the fellow reviewer and the authors. I was very pleased to see that all the issues raised by my colleague were quickly address and that the authors recognized many of the limitations point out. Nevertheless, I must say I found the scientific approach correct. It is multidisciplinary and despite relying mostly on numerical modelling approaches, the field validation guarantees reliability of results presented. The fact that the manuscript is superb in English and structurally helps the reader a lot. Furthermore, I also think the figures are of good quality. In a sentence, the results are presented in a clear, concise, and well-structured way. Figure number and quality are suitable and very informative.

We thank the reviewer for this comment.

3. Discussion: Having said this, I believe the case for the relevance of infragravity waves is not sufficiently stressed and more conclusive and clear evidences are missing. This was also noted by the other reviewer and it is a crucial aspect of this work. It is stressed in the title, in the abstract, etc. but results do not seem to so clearly demonstrate the reasoning forward. I believe the authors must be less enthusiastic and more cautious when writing the discussion. It is crucial that they address the shortcomings and discuss reasons for the poor discrimination made (for example on Figure 10). I am also curious about the error associated with the models and would like to see that clearly mentioned in the methods. The approach used is a succession of different model data and I am wondering if the sum of errors is not above 20%... I am mentioning this because in the abstract you describe increases of 1.1.to 1.6…

We thank the reviewer for highlighting this. To account for the error in the empirical models, the estimates are multiplied by normally distributed factors with mean values and standard deviations to represent the bias and scatter (errors) associated with each model. This is already described in the Methods. Likewise, the error (uncertainty) in the SWAN numerical model is captured in the breaker parameter (Equation 4) which is treated as a stochastic parameter with a standard deviation of 0.05. The error of each model is also shown in in Figures 6 to 8 as error bars.

We have added the following text to lines 507-513:

"To account for the error in the empirical models, the estimates are multiplied by normally distributed factors with mean values and standard deviations to represent the bias and scatter (errors) associated with each model (Section **Error! Reference source not found.**). This uncertainty is also shown in **Error! Reference source not found.** to **Error! Reference source not found.** as error bars. As the overall approach is a succession of different numerical and empirical models, it is important to note the combined error. The combined error (or uncertainty) may be expressed using a coefficient of variation, which is equal to the combined standard deviation normalized by the combined mean. If we consider the means and standard deviations of Equations **Error! Reference source not found.**, **Error! Reference source not found.**, **Error! Reference source not found.** and **Error! Reference source not found.**, the combined coefficient of variation or uncertainty is 0.15 or 15%.

4. Discussion: Furthermore, roughness is never mentioned and I think it is a crucial physical aspect when we are discussing overtopping. On line 564 you state: "the influence of saltmarsh vegetation on coastal safety under extreme forcing remains an important issue for future research." I was somewhat disappointed that this theme was not further discussed as it deserves. So, my suggestion is to add a paragraph further discussing this topic after line 600, for example.

We thank the reviewer for highlighting this. Bottom roughness is indeed a wave dissipation process to be considered here. We have added text to describe this. See our response to comment #43 of Reviewer #1.

We agree with the reviewer that wave overtopping is sensitive to slope roughness; however, here we do not consider roughness elements, as the grass-covered dikes of the Wadden Sea are considered smooth. We have added the following text and literature to describe this (lines 286-288):

"The dikes of the Dutch Wadden Sea are typically grass-covered and therefore treated as smooth (without roughness elements). To simplify the calculation, the dikes are assumed to be uniformly sloping (without a berm). However, it should be noted that the presence of berms and roughness elements can significantly reduce the overtopping discharge (Bruce et al., 2009;Chen et al., 2020;Van der Meer, 2002)."

The influence of saltmarsh vegetation is discussed in Section 4.4. We agree with the reviewer that the theme is important and deserves further discussion and research. However, more measurements of wave-vegetation interaction under very severe storms are needed before further conclusions can be drawn. Here, we discuss our vegetation results based on measurements taken during two storms (with exceedance probabilities of 1/5 per year); but given the uncertainty surrounding vegetation and whether or not it will remain standing under more severe conditions, we are unable to expand on this further.

References:

Bruce, T., Van der Meer, J., Franco, L., and Pearson, J.: Overtopping performance of different armour units for rubble mound breakwaters, Coastal Engineering, 56, 166-179, 2009.

Chen, W., Van Gent, M., Warmink, J., and Hulscher, S.: The influence of a berm and roughness on the wave overtopping at dikes, Coastal engineering, 156, 103613, 2020.

Van der Meer, J.: Technical report wave run-up and wave overtopping at dikes, TAW report (incorporated in the EurOtop manual), 2002.

5. Materials and Methods: Finally, I am worried with the limited number of events analysed and with the narrow spatial distribution studied. To support some of the bolder statements regarding the relevance of infragravity waves the authors should have provided a more extensive database. Despite this, I feel this is a very good contribution to this scientific subject and deserves to be published on NHESS after some minor changes are made.

Here, we apply numerical and empirical modelling tools that were previously validated using physical model tests and now validated here against field data. Nonetheless, we agree with the reviewer that a larger database of field measurements would improve the work. As field measurements under actual storm conditions are difficult to obtain, we made use of the most appropriate dataset available for the Dutch Wadden Sea, to-date. In our manuscript (lines 51-653), we recommend that additional field campaigns focused on measuring infragravity waves be carried out in the Dutch Wadden Sea to provide the data necessary to further validate and support the conclusions drawn here.

6. Conclusions: Another aspect is the extention of the Conclusions. They must be more focused and a couple of paragraphs could be deleted as they are very generic.

We thank the reviewer for highlighting this. We have re-written the Conclusions so that it is more concise.

7. References: A final note on some self-citation and what I consider to be an average reference list. There are a few classic papers missing….

With guidance from Reviewer 1, we have added a number of classic references to the manuscript. This work brings together several tools and methods that were developed in previous works by the authors; hence, the apparent self-citation. However, we have minimized and replaced the references with more classic examples, where appropriate.

8. General: Case presented is not sufficient to prove major importance of infragravity waves.

Our findings indicate that neglecting infragravity waves results in an overestimation of safety. We also highlight that the magnitude of the impact will differ per case, as it is dependent on local conditions (forcing and bathymetry). For the case considered here, the change in failure probability was by a factor of 1.1 to 1.6 times with equivalent cost of approximately M€1/per km. This may (or may not) be considered major to some but highlights that IG waves do indeed play a role.

9. General: Error associated with models and poor discrimination.

See our response to Comment #3 above.

10. General: Roughness.

See our response to Comment #4 above.

11. Conclusion must be shortened.

See our response to Comment #6 above.