General comments

This manuscript investigates the mechanisms leading to a series of very large run-up events that took place within a couple of hours at different locations along the US Pacific Northwest coast. The authors show that the series of extreme runup events coincides with the arrival of very long swells at the coast and hypothesize that these long swells were responsible for the generation of unusually large (initially bound) infragravity waves that ultimately led to the large, 5-6 min period, run-up events that were filmed by beachgoers.

The manuscript addresses a topic that is relevant to the scientific community and presents some interesting results. The analysis is however limited by the type of data that is available (e.g., no nearshore wave measurements except the 1min-resolution tide gage water level data) that prevents strong conclusions to drawn. I think that it is acceptable as long as the limitations of the analysis are clearly stated (in the core of the manuscript but also in the conclusions). This is a point that needs improvement in my opinion (see comments on below the interpretation of the DART data for instance). Furthermore, several aspects of the methodology, particularly in Section 6.2, need to be clarified and possibly contain errors (see detailed comments in the next section).

I therefore recommend major revisions.

Specific comments
The paper refers a few times to the runup of bound infragravity waves. I find the formulation confusing as the infragravity waves that run up the beach are not bound anymore (released in the surf zone). This is coming back at different locations within the paper, in slightly different forms. For instance it is stated line 265 that “th[e] bound wave travels at the speed of the wave groups towards the shore and as a free wave [...] after reflection” which is not strictly correct as IG waves are not bound until the shore. Or in line 312 the authors discuss the “dissipation of the bound infragravity waves” and refer to van Dongeren et al. 2007 to do so, while this paper examines dissipation close to the shoreline, i.e. at a location where the IG waves are not bound anymore.

Line 32: I don’t think the reference to Roeber and Bricker is appropriate as an example in which trapped waves lead to extreme runup (they actually explicitly state that “Resonant amplification over the reef flat [...] did not contribute significantly...”).

Line 62: “the amplification of low frequency motion was found to decrease dramatically during storms”. What does that mean? How does it relate to the previous statements?

Line 106: I do not understand the sentence “Data is [...] averaged over 1 minute to produce 1 min timeseries”? Is 1 minute the sampling interval or the duration? (probably the first one?)

Lines 180-182: I find that the time delay between observations of “larger than usual” oscillations at the DART sensors and at the tide gages difficult to visualize and as a result the numbers given in the text seem quite arbitrary (I could argue based on visual inspection of figure 5 that the largest waves at Garibaldi occur several hours after the point identified as the approximate onset of the anomalous water level fluctuations... which would in turn change all the time lags). As this time delay seems important to the current narrative (it is used later to support the fact that the observed large oscillations at the DART sensors could be due to IG wave reflection), it would be useful to define periods with “higher than usual magnitudes” in a more systematic way (e.g., it starts when the elevation is larger than X*std(eta)?).

Lines 195-207. There has been a number of recent publications on coastal hazards induced by meteotsunamis such as Shi et al., 2020 (Nature communications) and Anarde et al. 2020 (JGR-Oceans) that could be relevant to this discussion. For instance Shi et al., 2020 showed that both single-peak meteotsunami waves and long lasting meteotsunami wave trains could be generated (thus not only soliton like waves as mentioned in the manuscript). Anarde et al., 2020 showed that pressure disturbances much lower than the 5HPa mentioned in the manuscript could trigger significant meteotsunamis.

Lines 211-212: The fact that the long waves that are responsible for the extreme run-up events were generated as bound infragravity waves is a key result of this paper which is in my opinion not sufficiently supported. I understand that data availability limits the analyses that can be conducted, but still feel that stating that 5-min is a “plausible period for wave groups when carrier waves are 25 s” (see also line 260) is in
itself not a very convincing statement (for such a key result). Have the authors considered using the measured spectra at the wave buoys to reconstruct a timeseries using a random phase (which is probably not a too bad assumption in these depths) and examine the group structure? They could for instance look at the spectrum of the envelope of such a reconstructed timeseries and check that it indeed peaks around 5 min? That would lend some additional support to the fact that such a wave field could indeed force bound waves in the proper frequency range.

- Line 222: What is the typical period of the waves measured by the DART sensors during the period where they recorded at high temporal resolution? Is it indeed close to 5min and therefore consistent with the 5-min peaks in the tidal gage spectra?

- Lines 276-279 (and following): van Dongeren et al. 2007 (and I expect that Battjes et al. as well, although I haven’t re-checked) used shoaling zone data in off-resonant conditions to examine the relation between the growth rate (alpha) and the normalized bed slope (beta) and demonstrate that the infragravity wave amplitude was bound between $h^{(-1/4)}$ and $h^{(-5/2)}$. This means that the relation between beta and alpha (that is used in this paper) was derived for conditions in which the carrier waves were not in shallow water, contrary to what is stated at the start of this paragraph (and a few times later on, e.g., line 295).

- Lines 291-305: I find it difficult to understand what the model is exactly doing, and even more difficult to follow the description of the model outputs that follow in the second paragraph. Some additional information is needed (see questions/remarks in bullet points below). A figure showing the cross-shore evolution of short wave and infragravity wave heights could also help.

- Line 294: I would expect that the infragravity wave amplitude is calculated until the moment the short waves break, not when the infragravity wave break as stated here (as van Dongeren et al. derived eq. (3) based on shoaling zone data)?

- Lines 296-298: I guess the short-wave height is shoaled according to linear wave theory until $H=\gamma h$? It would be good to say it explicitly.

- Line 303: I find the formulation confusing. I do not expect that the model described in the paragraph above is needed to determine that 25 s waves are in shallow water for $h<15m$ (which depends only on the dispersion relation). Also, how is this info used to determine that the infragravity waves have a 0.87 m amplitude at this depth? Does it mean that eq. (2) is used until the point where the waves enter shallow water (let’s say $h/L=0.05$)? In that case, does it mean that van Dongeren et al. relation is used only shoreward of that point? (while to the best of my knowledge this empirical relation was not derived using shallow water data - see also comment #11)

- Lines 305 and 306: Different alpha’s are used for the 10s and 25s short wave cases. I understand that these depend on the beta-values, but how are these beta’s calculated? Do they differ only because of differences in breaking depth for the 10 and 25s short-wave periods? Or is omega, the IG wave frequency, changing as well when the carrier frequency is changing? If that’s the case, how is it changing/what is the assumption to calculate this frequency?

- Line 306: an infragravity wave of amplitude 2.3 m (so 4.6 m high) in 5.8 m depth seems very large. Is it a typo or is there a problem with the model?

- Line 308: It seems unlikely that waves that are 5m high offshore break at a depth of 1.8 m (I would expect the breaking depth to be closer to 6-7m). Is this a typo?

14. I have similar comments/questions on the paragraph discussing the dissipation patterns based on the beta_H-value
- How is beta_H calculated in both cases? Again, beta_H depends on the infragravity wave period (or frequency), not on the short-wave period. So some explanations are missing.
- According to the definition of beta_H, a larger IG wave height H means a smaller value of beta_H, and thus a smaller R (meaning more dissipation). So unless other parameters are varying (such as the long wave period), I do not understand how it can be concluded that the case with larger infragravity waves is dissipating the least (lines 322-324).

15. Finally, the link between dissipation and run-up does not seem that straightforward to me as we are comparing the runup of waves that have different incoming heights (and maybe period?). A wave that dissipates more can still lead to a higher runup than a smaller that would have dissipated less...

16. Lines 388-389: "This conclusion is supported by far offshore sensors which did not detect the incoming waves but did detect the returning ones...": The formulation suggests that the authors were able to discriminate between incoming and reflected waves -> reformulate?

17. It should be clear in the conclusion that the predictor developed in this study is (likely to be) highly site-specific (depends on a dimensional parameter, involves a proportionality factor that already varies strongly along the considered stretch of coast...).

Technical corrections

- “very low frequency” is usually used to describe motions happening at a much lower frequency than the long swells described in this paper, which I find a bit confusing.
- Line 228: shouldn’t it be figures 8 and 9 instead of figures 9 and 10?
- Line 346: I guess that sqrt(m_{-3}f_{-3}) should be sqrt(m_{-3}) here?