Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2019-81-RC1, 2019

Natural Hazards and Earth System Sciences
Discussions

Interactive comment on "Ensemble models from machine learning: an example of wave runup and coastal dune erosion" *by* Tomas Beuzen et al.

Anonymous Referee #1

Received and published: 30 April 2019

General Comments:

Beuzen et al., present a probabilistic, data-driven model for ensemble predictions of wave runup and dune erosion. They show their model has a lower RMSE and bias than other frequently used, deterministic runup predictions, and has the ability to produce confidence bounds on runup (and thus dune erosion) predictions. This technique could bring value to erosion forecasting as well as to the development of probabilistic hazard zones. As such, I feel that the research forms an important contribution. The manuscript is well written and describes the motivation, model, and results very clearly. However, I feel the authors need to discuss the limitations of their approach more outright as well as improve some of the structure of the results and discussion sections. Nevertheless, if these comments are adequately addressed, I find that this

Printer-friendly version

Discussion paper

paper would be a valuable addition to the literature.

Specific Comments:

As mentioned above, this manuscript develops a novel method for including uncertainties within estimates of wave runup, an important contributor to coastal erosion hazards. While there is certainly merit to the results presented, I feel that the manuscript would benefit from a discussion of the limitations and assumptions of the modeling technique as well as some changes to the overall structure.

My main concern is that the authors neglect to discuss the limitations of their methodology. For example, the authors' state that machine learning models perform poorly when forced to extrapolate a prediction, and that it is important to use training data sets that capture the full range of variability of the data. While the authors use training data from a year, this may not account for interannual variability of the wave climate and long-term trends or shifts in storm tracks and intensities. The authors should discuss how representative the year of training data is of past years, wave climate-wise, or other years of measurements, runup-wise, as well as the precautions in a limited set of training data. Furthermore, rare, extreme events may not have occurred during that particular year. This is often also one of the issues with deterministic/empirical parameterizations of the R2% from field studies – that extreme conditions haven't been captured. On that note, the authors seem to suggest in their discussion that this technique is more reliable than Monte Carlo techniques because the uncertainty is learned directly from data. Monte Carlo techniques however, seek to represent conditions that haven't necessarily occurred by modeling large sets of physically plausible conditions. Both the GP and MC techniques seem equally useful, with different motivations.

Following on, there is no adequate discussion of the model's poor prediction for large erosion events seen in Figure 8 between profiles 28 -34. The authors suggest that the under-prediction of large erosion events could be due to a non-optimized Cs parameter, stating that, "...increasing Cs would lead to better mean ensemble predictions of

the large dune erosion volumes but over-prediction of the smaller events." Does this mean that the optimum Cs within table 1 for each ensemble grouping captured erosion over this set of profiles better? Is it more important to weight the mean or extreme conditions?

Then, in Lines 534-535, "the results demonstrate the ensemble approach is less sensitive to the choice of Cs than a deterministic model so it can be useful for forecasting with non-optimized model parameters". So on one hand, they suggest that erosion can be better predicted with an optimized Cs, and on the other hand it will be useful for forecasting with non-optimized model parameters. Can the authors please expand on these seemingly contradictory statements?

In Figure 8, there are some points that fall well outside of the range of uncertainty in erosion measurements. The authors state that, "regardless of the value of Cs chosen, an advantage of the GP approach is that uncertainty in GP predictions can give an indication of dune erosion, even if the mean dune erosion prediction deviates from the observations. " While there is truth to this statement, there are locations where the uncertainty does not characterize the observations at all, and this should be discussed.

Furthermore, I understand that the main contribution of the paper is the ensemble technique for modeling wave runup rather than the erosion model calibration, however I found the sections focused on the transport coefficient, Cs, to break up the flow of the manuscript. Specifically, the section describing previous research on Cs (Lines 415 – 440) can be shortened and put into an appendix or supplemental information so the reader can focus instead on the results. I felt the most important information in this section was what Cs value was being used, how much it ranged, and specifically Lines 430 – 440. Additionally, the results that appear in the discussion section (section 5.2) could be moved to the end of the results section. While these results are important for weighing the importance of Cs, they felt out of place in the discussion section.

Finally, this manuscript left me wondering what type of applicability this has to other

locations. For example, is this method limited to locations with data?

Comments on Specific Lines:

Lines 294 – 295: The authors may want to consider citing the technical methods for extracting wave runup as this is not a trivial task.

Lines 392 – 393: What is the resolution of SWAN model? 10m?

Line 316: does “this value” refer to 5% or N = 416?

Figures:

Figure 3: Why is the wave runup histogram plotted in red?

Figure 7a: Does the blue display the runup prediction or the total water level prediction? It looks to me like there are tides within the blue values. Also, is it possible to plot some of the wave runup data on 7a? Authors show observations of erosion and dune toe, however, their model is predicting wave runup so that would be interesting to see in the time series sense too.

Technical Corrections:

Line 35: The authors use the term “significant wave period” to describe the wave period variable in runup formulations and then later mostly use “peak wave period.” As “significant wave period” is not used as typically, I’d recommend sticking with “peak wave period” or simply “wave period.”

Line 285: Please define SSE acronym, and if not used again, no acronym is necessary.

Lines 446 – 447, Lines 509-510, Lines 545-546, Lines 595 – 596, Lines 619-620: repetition of similar variations of the following text, “drawing 10,000 samples takes than one second on a standard desktop computer.” I’d recommend the authors say this a few times, then perhaps the term “computationally efficient” as this felt repetitive to read 5 times.

Line 603: remove “were used”

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2019-81, 2019.

Printer-friendly version

Discussion paper