Natural Hazards
and Earth System
Sciences

Discussions

Open Access

EGU

Interactive comment on "The use of genetic programming to develop a predictor of swash excursion on sandy beaches" *by* Marinella Passarella et al.

Anonymous Referee #3

This is a well written paper that proposes a new methodology to predict total and infra gravity swash elevation. As such it is of interest to NHESS and coastal scientists/practitioners. The methodology followed is correct and well explained. In particular, there is a very clear explanation of Genetic Programming and how this technique has been used for this work. This is very well written in a way which is suitable for non-experts approaching the methodology for the first time. The data used are of very good quality and there is a good explanation of the range of parameters covered by the dataset. The results are discussed in concise and detailed way and the accuracy improvement over existing relationships is demonstrated.

C1

My only minor suggestion is that the use of both MSE and RMSE is redundant and one of the two can be omitted.

Therefore I recommend publication after minor revision.

Minor corrections/suggestions.

Abstract Line 14: change the sentence: it contributes to the error, maybe it is contributes to the reduction of the error. However, beware of repetitions.

Also many repetitions of "wave runup" in the introduction, try to rephrase.

Line 123, concomitant is possibly better replaceable by "associated".

Line 143-147: specify the countries of the beaches named as not all authors might be familiar with these.

Line 170. You might want to specify which is your stopping criterion, and when do you consider the solutions stable.

Line 237: overfitting is mentioned, but it could be useful to explain what this is in the present context. Explanation in 240 occurs after the first use of the term and it is not clear.

Some sentences are written in present tense (e.g. we use at the beginning of Section 3.3, and "...finally selected" at the start of 4.2). Please make the tense consistent.

Also, in Line 314 it is mentioned that experiments in Ngarunui beach are carried out under mild dissipative conditions. Is the difficulty in predicting these results due to the particular combination of H and T (hence L)? It would be useful to be more detailed in explaining this.

In Line 356 it is claimed that the procedure followed is different from the use of a single data set. This needs clarification, as you always build one dataset that is divided in three for training validation and testing. The same was done in the development of

C2

ANN tools for overtopping in the CLASH project (van Gent et al. 2007), for example, when the dataset used was actually a composite one resulting from many datasets.

References

van Gent, M.R., van den Boogaard, H.F., Pozueta, B. and Medina, J.R., 2007. Neural network modelling of wave overtopping at coastal structures. Coastal Engineering, 54(8), pp.586-593.