Natural Hazards and Earth System Sciences Discussions

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC1
https://doi.org/10.5194/nhess-2021-253-RC1, 2021

Comment on nhess-2021-253

Caterina Samela (Referee)

Referee comment on "Machine-learning blends of geomorphic descriptors: value and limitations for flood hazard assessment across large floodplains" by Andrea Magnini et al., Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2021-253-RC1, 2021

Premise

In the present study univariate geomorphological assessment of flood hazard is investigated and thoroughly compared with a multivariate assessment, in which several DEM-based geomorphic descriptors are blended together by means of decision trees. The main goal is to understand if an improvement can be obtained by applying a machine learning aided multivariate flood hazard assessment relative to a univariate approach, and to quantitatively evaluate this improvement.

These techniques are designed in particular, but not exclusively, for large-scale analyses and in data-poor environments

This branch of research has experienced in the last two decades an increasing interest, where DEM-based methods are becoming more and more frequently associated with more or less complicated machine learning algorithms.

Among the rich literature on this topic, the authors clearly list in the introduction five key elements that are addressed simultaneously in their study, making it different from all previous works. Also research questions are clearly formulated, the analysis of results is complete and leads to convincing and satisfactory answers.

The article is clearly written and organized, with a good quality in terms of presentation. Authors could consider to implement improvements in the presentation of the work according to the reported comments.

The manuscript is excellent as regards its scientific significance and scientific quality. Applied methods are valid from a scientific and technical point of view, and the results are discussed with clarity of concepts.

After this premise, I consider this work a substantial contribution in this field and an article of great interest to NHESS readers.

Major comments

- Terminology: "flood hazard" maps is a terminology more appropriate to maps derived by hydrologic/hydraulic simulations. Topography-based (hydrogeomorphic) maps are generally termed in literature as flood-susceptibility maps, or flood-prone areas map or floodplain maps (see e.g. indersson, S., Brandimarte, L., Mård, J., and Di Baldassarre, G.: Global riverine flood risk – how do hydrogeomorphic floodplain maps compare to flood hazard maps?, Nat. Hazards Earth Syst. Sci., 21, 2921–2948, https://doi.org/10.5194/nhess-21-2921-2021, 2021.)
- One of the most important issues addressed in this work is the estimate of the water depth, a parameter of fundamental importance especially for estimating expected flood damage. Compared to the large number of published studies on the delineation of the areal extent of flood hazard areas, in the literature there are fewer studies concerning the estimation of water inundation depth with simplified methods. This is an added value of this work. However, since DEM-based methods find their primary purpose in applications in data-scarce environments (although not exclusively), it is perhaps worth considering the fact that while reference data to calibrate the classification problem are often available also in this contexts, on the opposite flood hazard map providing water depth values (to use for calibrating the regression problem) are more difficult to find. In addition, this data should be characterized by good accuracy in order to train a simpler but reliable model based on it. I think a consideration on this aspect can find a place in the manuscript.
- I wonder about the choice of identifying the calibration area by setting a constant-radius buffer. In this study, testing the performance outside of calibration areas is part of the application, so it was possible for authors to perform a sensitivity analysis on the accuracy obtained with different buffers. However, readers who want to apply the methodology with no possibility to validate the results (e.g. in poor data environments) are left without guidance on how to set this constant buffer. Here, in the same work for the same study area, two different buffer values are considered the best for the two reference maps (2 km for the 500-years PGRA flood hazard map, and 5 km for the JRC 100-years flood map). Instead, a topographical-hydrological criterion (e.g. the one used by Degiorgis et al., 2012) offers the possibility of being adopted and re-applied in any context, responding at the same time to the characteristics of the study area and of the available reference map. This consideration does not influence the relevance of the investigation and the interesting results obtained, but is made thinking about how to replicate the study in different case studies.
- The analyses are made up of a series of steps and sub-steps (and further sub-steps), not always easy to follow along, that are listed in the first lines of Section 4 "Framework of the analysis". Then, subsections 4.x do not follow any of the previous subdivision. Did you consider that the methodology would be easier to read, follow and reproduce if a subsection is dedicated to each of the major 4 steps?
- In tables 1, 2, 3 can be unclear the difference among the results of the first two rows

and the other rows. Section 4.2 reports that the models have been applied a first time using the entire domain of the calibration areas, and then the models were applied again four other times after selecting four subdomains of the calibration area (to test extrapolation performances). I believe this should be better clarified and the section 4.2, in general, could be reorganized. For example, it first describes what happens in phase (3), then in phase (4), and toward the end of the section is nominated phase (2). Is there a possibility to simplify and re-order this description?

Minor comments

- Section "1 DEM-processing to obtain geomorphic descriptors": line 98 "composite indices, if they are derived based on a combination of different other indices." You probably mean "based on a combination of different other features".
- Section "1 DEM-processing to obtain geomorphic descriptors": in the list of composite indices, references for the GFI and LGFI are lacking.
- Section "1 DEM-processing to obtain geomorphic descriptors": line 127, the occurrence of inunvdation. Correct the word inundation.
- I believe Figure 1 could be improved and enriched to help readers understand the DT algorithm adopted in this study.
- Figure 2 appears before it is mentioned in the text. Consider moving it to Section 3 Study area.
- 4 Framework of the analysis, lines 187,188: "The first step (a) is the definition of the conceptual scheme of the processing operations to set up the models, which is based on the following four points"… but actually five points are listed (i, ii, iii, iv, v).
- Figure 5, green panel: "Training DTs and calibrating GFI with the same training set….". Why only the GFI is mentioned in the panel? Is this performed only for the GFI or also for the other selected geomorphic descriptors (GDs)?
- Figure 5, red panel: correct typo in "perofrmance".
- Line 210, typo in "calbiration".
- Results, line 278: "In order to find out the relevance of each input GD in the DTs' structure, the Gini importance for each model is reported in Table 1". But Table 1 reports the performance metrics of the Classification problem, instead.
- Line 293, a Table 3 is mentioned, while a Table 2 did not appear in the text yet.
- Line 334, typo in "ligth".

Recommendation

Overall, I do not find any major deficiency in the article and I recommend publishing the manuscript after minor revisions.