Natural Hazards and Earth System Sciences Discussions

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/nhess-2021-22-RC2, 2021

Comment on nhess-2021-22

Anonymous Referee #2

Referee comment on "Optimizing and validating the Gravitational Process Path model for regional debris-flow runout modelling" by Jason Goetz et al., Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2021-22-RC2, 2021

General comments

The paper presents an approach to optimize the parameters of the Gravitational Process Path model for regional debris-flow runout modelling. It addresses the evaluation of the source areas as well as of the runout path and its length. The approach is illustrated with a case study in the upper Maipo river basin in the Andes of Santiago, Chile. The method and the sensitivity analyses are interesting and add value to the field of regional debris flow modelling. The paper is well written, and the figures are of high quality. I recommend publishing it after consideration of two main concerns I have about performance metrics that may require additional work.

Main concerns

I have two main concerns about the performance metrics of the runout distance and runout path:

AUROC for the path: you process the AUROC as defined by: "Model performance was rated higher if the random walk model contained observed debris-flow tracks within its simulated paths" (2.2.1). The problem here is that there is no "false positive" in your approach, and thus the model is not penalized for over-predicting. The approach is correct for the source areas but not for the runout path. As we can see in Fig. 2, the extent of the modelled debris flow is much larger than the observed one, but the AUROC is almost = 1. It means that your model needs to spread as widely as possible to have a good score. I get the difficulty of comparing potential events to a single observed event, but you might then use another contingency table score that does not have false positives. Using a ROC-type score is misleading here if there is no false positive.

Relative error for the runout distance: your approach of using a bounding box on the median frequency (2.2.1) to quantify the runout distance is interesting, but I have an issue with it. Most observed debris flows will likely propagate to the valley-bottom, where they might meet the main river. My problem is that when you model the debris flow propagation with small friction values, it is likely to reach the main river and continue perpendicularly, thus not increasing the bounding box for some iterations of the parameter values. We can see such behaviour in your Fig. 2. There is, therefore, a discontinuity as too long propagations are less penalized than too short ones. I believe this might play a role in the results of Fig. 6, where the runout length error remains low for a large range of sliding friction coefficients. It might provide a misleading impression of insensitivity. Or is it the case that most propagations reach a flatter area where they quickly stop anyway? Although an approach based on actual length (for example, defined by a D8) might better represent the difference in runout distance, it might not be trivial to use the median frequency criteria. What about using the median length of all random walk runs for one setting, provided it's a piece of information you can get? This problem should be at least discussed and considered in the interpretation of the sensitivity analyses. Then, interpretation such as in l. 380-381 ("This may indicate that the combination of random walk and the process based PCM model dictates a general runout pattern that is insensitive to values within a broad and nearly optimal range of physically reasonable parameters") might not be stated this way. Same for l. 407 ("we observed a general insensitivity in runout distance performance of the PCM model to a range of parameters").

Specific comments

Figure 1: The caption should be a bit more comprehensive, explaining, for example, the random sample.

Section 2.1.3: It might be useful to describe the fundamental principles of the AUROC in 1 sentence.

Section 2.2: Please provide more details about the models and their parameters. For example, mention the random component in the iterative simulations of the random walk and give more information about the persistence factor and the exponent of divergence. As they are key parameters for the rest of the paper, adding a few sentences to describe them and 1-2 equations would be beneficial for the readers.

Section 2.2.1: Please mention that you do an exhaustive grid search.

Section 2.2.4: You have chosen 1000 "non-debris flow locations". However, could these be excluded to be potential source areas for future events? Could they become source areas under certain triggering conditions?

Figure 5: It would deserve some more interpretation. For example, what can explain the role of elevation in debris flow conditioning? Why is the slope angle contribution decreasing after a certain threshold? What about the plan curvature?

Section 3.3 & Figure 9: Is the runout frequency relative to a single source? How are they combined when different propagations overlap? Please add some clarifications.

Section 3.4, l. 299: "these cases were related to misclassifying stream erosion": can you identify such information from satellite imagery?

Section 3.4, l. 309-311: not so clear; please clarify.

Figure 12a: You do not mention plot 12a in the text, i.e., the slope threshold values in the grid of other parameters.

Section 3.6 & Figure 14: You might mention again here that these scores are processed on the test data.

Section 4.2: The ability to optimize the runout path and the runout distance separately is related to the fact that the random walk mainly controls the path/spreading, and the PCM controls the runout distance. The influences of these algorithms are quite distinct.

Conclusion: Should contain some more results of your study.

Technical corrections

l. 5: “y” is missing in Germany

l. 73: “our” instead of “out”

l.186: “We explored *for* such spatial…” ?

l. 378: what do you mean by “ambiguous events”?

l. 386: “very *a* specific problem”