Natural Hazards and Earth System Sciences Discussions

Nat. Hazards Earth Syst. Sci. Discuss., author comment AC1
https://doi.org/10.5194/nhess-2021-22-AC1, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

Jason Goetz et al.

Author 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-AC1, 2021

General Comments

Dear Editor, dear Authors,

this a well-written and interesting paper on the automatic calibration and validation of a framework for regional debris-flow modelling. Besides the modelling of debris-flow initiation sites with a GAM, the GPP model is used for debris-flow path and runout modelling in the upper Maipo river basin, Andes of central Chile. The authors develop and present a novel approach for model optimization and validation, including several aspects like uncertainty in parameter selection, spatial transferability, and the models's sensitivity to sample size. The results are well presented and discussed, including very nice and informative figures to illustrate the findings. Most parts of section 2 (material and methods) are also well written, but I think this is the section which could be improved most by adding some more detail on some of the aspects (see specific comments below). Apart from that I think the paper is well suited for publication in NHESS. It is also really nice to see that the tools developed for this paper (as well as the data) are also made available to the public.
With best regards.

We thank this reviewer for their highly constructive comments. As shown in our following response, we will make improvements to the methods section to help clarify our approach on the use of the AUROC as a metric for runout path modelling, sampling debris-flow and non-debris-flow source areas, and how the the GPP implementation of the random walk model limits runout distance. Additionally, we believe by addressing their comments on the potential multiple optimal PCM (runout distance) model solutions, we enhance our discussion and further demonstrate the suitability of our approach for regional debris flow runout modelling.

Specific Comments

Section 2.1.2
Please use a different color for debris flows and roads in Fig. 1, they are both grey and can't be distinguished very well.

We will update the colours in the figures.

Section 2.1.3
Regarding the sampling of presence and absence of source points: how do you exactly

determine the non-source points? Do you somehow guarantee that the samples are not "too close" to mapped source points? There are much more non-source than source points in your study area, how does this influence the results? This affects training as well as validation, please elaborate.

We will rephrase some sentences in this section to help clarify how non-source points were sampled and how many.

"The non-source (i.e. absence) points were determined by random sampling locations within the mapped sub-basins outside of the debris flow polygons. The resulting training and test data contained 541 source points and 541 non-source points."

We guaranteed that the samples were not too close to source points by sampling outside of the mapped polygons. As mentioned in L.102. We used the commonly applied 1:1 sampling ratio of source and non-source points.

After denoising, you apply a sink filling algorithm to the DEM, which one?

We used the sink filling algorithm from Planchon and Darboux (2001). This citation will be added.

Section 2.2.1
Regarding the rating of the random walk performance (line 160): performance was rated higher if observed debris-flow tracks were within the modelled paths. Please provide more details on how this was done exactly, e.g. did you also take the number of cells into consideration that were outside the mapped track? Otherwise you might get optimized parameters that overestimate the process area.

We accounted for the cells outside of the mapped track. The ROC curve plots the true positive rate (TPR) vs the false positive rate (FPR; Zweig and Campbell 1993). Therefore the AUROC does consider cells inside and outside of the mapped tracks. We will add a brief description of the AUROC to the paper to help clarify this, as well as the Zweig and Campbell 1993 citation.

Regarding the random walk parameter optimization before the runout optimization (two-stage approach): in order to optimize the random walk parameters, wouldn't you also require to use some kind of friction model to limit the runout distance? This overlaps with the previous question, please explain.

The Gamma (2000) random walk model implemented in the GPP model (Wichman 2017) does not have controls for runout distance. The flow paths will continue downslope until neighboring cells have a higher or equal elevation compared to the central cell being processed. We will add this detail to the paper.

Regarding runout distance optimization: here, you use a minimum area bounding box to measure length. What impact has the character of the derbis flow path on this concept? For example, take (1) a quite short, more or less straight debris-flow path versus a (2) very long path, which runs from a hillslope into a channel with a distinct change of direction, let's say 90°? Then you get (1) a bounding box matching the real length quite well and (2) a bounding box which is almost square, strongly underestimating the runout length.

This is an excellent question. It was also brought up by Reviewer #2.

It is possible that the runout length of the minimum area bounding box can be underestimated when a debris flow makes an abrupt 90 degree change in direction. This

may occur for some iterations of decreasing sliding friction coefficient (or increasing mass-to-drag ratio) past the actual optimal value. However, to mitigate this issue in optimal parameter selection, we use the AUROC to break any ties in performance. Longer runout paths should have a lower AUROC. We will add the following paragraph to the discussion to better explain this issue:

"Although we obtained a unique regional model solution, runout distance relative errors were only slightly higher than the best performer for pairs of sliding friction coefficients and mass-to-drag ratios across a band in grid search space of lower sliding friction coefficients. This insensitivity of performance to different combinations of PCM model parameters may be due to the uniqueness problem. Our approach using the minimum area-bounding box could also contribute to this observed parameter insensitivity. Abrupt changes in flow perpendicular to the initial flow direction, such as a flow meeting a channel, may only slightly increase the length of the bounding box for several iterations of decreasing sliding coefficient (or increasing mass-to-drag ratio). However, our approach of breaking parameter ties using the AUROC ensures that we select the parameter set that best fits the runout extent in addition to distance – slides with longer runout distances should also have a lower AUROC performance"

Regarding the optimization of the 2 parameters of the PCM model (sliding friction coefficent "my" and mass-to-drag ratio "M/D"): a general problem with the PCM model calibration is, that there is some mathematical redundancy between the parameters. I.e., you can achieve the same runout length with different parameter combinations of my and M/D. How does your calibration approach handle this? Please add some information on this, because this may also have some impact on other sections of the paper, e.g. section 3.2 ("low sensitivity for a large range of parameter combinations"), section 3.5 ("no clear spatial pattern in optimal my and M/D parameter combinations across the study area"), section 4.1.2 ("we observed high variability in optimal PCM parameters").

If there were ties in the PCM model, we select the parameter set that resulted in the highest AUROC of the runout path performance. We did explore this, and found that this was not a major issue for regional optimization. Only 5% of the repeated spatial cross-validation iterations (n = 5000) had multiple optimal solutions. For individual events this occurrence rate was much higher (56%). However, after using the AUROC to break ties, the vast majority of individual events (97%) had a unique solution - we will add these results to sections 3.2 and 3.5. For the remaining cases, which still had ties, we simply select the first record. Additionally, in the case where relative error may seem insensitive to perpendicular changes in flow directions, the AUROC enables us to select the optimized flow distance that best matches the flow extent of the mapped debris flow tracks.

We will add the following to the discussion to better highlight this issue with optimizing the PCM model.

"The two-parameter PCM model has a uniqueness problem (Perla et al 1980). There are possibly infinitely many pairs of the sliding friction coefficient and mass-to-drag ratio that result in the same runout distances. When optimizing individual events, we did observe this phenomenon. The majority of individual events had more than one optimal combination of parameters. Obtaining a unique solution was not an issue for the regional optimization in this case study for the given grid search space. Likely this is due to having to satisfy the runout distances for a variety of hillslope conditions and lengths across the study area. Through our investigation of sample size, we observed a reduced variability in PCM model optimal solutions for larger sample sizes (Figure 15)."

Section 2.2.2
You assessed the transferability of optimized model parameters by 5-fold spatial cross-validation. In section 2.2.1 you state that you are using a random sample of 100 debris-

flow tracks for optimization. Is this the sample size you use here too? Or how is this related?

It is the same sample. We will add, "Based on our random sample of 100 debris-flow tracks, ..." to help clarify this.

Section 2.2.4
To calculate the AUROC, you used 1000 samples of both debris-flow and non-debris-flow locations. How did you sample the non-debris-flow locations? Thematically similar to my question on the non-source point sampling.

We randomly sampled locations outside of debris flow polygons. We will rephrase this to, "The AUROC was calculated using a sample of 1,000 debris-flow runout locations and 1,000 non-debris flow locations outside of the debris-flow polygons".

Section 3.1
You write that areas with slightly concave profile curvature were modelled as more likely being source areas. So far plan (not profile) curvature was used, and it is also plan curvature that is shown in Fig. 5.

Thanks, this was a typo. We mean plan curvature.

Section 3.2
I think it would improve the reading of Table 2 if you would name the "third" model component "Runout distance (spatially varying friction)" instead of only "Runout distance" (like the "second" model component).

Good point! We will make this change.

Section 3.4
In line 294 you write "... the modelled runout paths failed to follow the flow direction ...": is this due to a general problem of the flow path model or is this caused by errors in the DEM?

This is likely a problem of the errors in the DEM than the flow path model. We previously mentioned this in the discussion - however, we will add references to works that cover these issues in more detail,

"Poorly individually optimized events could be attributed to locally poor DEM quality (Horton et al., 2013) and mapping uncertainties (Ardizzone et al., 2002)".

In line 299 you write that "these cases were related to missclassifying stream erosion ...": was the runout lemgth over- or underestimated in these cases?

Runout was underestimated for these cases (Figure 11c), likely due to the relatively gentle slope of these stream channels.

Section 4.1.2
This section (mostly) discusses the runout distance model, please also add a few sentences on the runout path model.

Thanks, we will add the following interpretation of the runout path model results to the discussion,

"The best-performing regional random-walk parameters allowed for maximum lateral spreading of the runout path given the range of parameters for optimization. Individual

events tended to also optimize for high lateral spreading, but not as strongly as the regional model. We believe this high lateral spreading may be due to the location of the observed debris flows relative to simulated paths and the quality of the DEM. A large proportion of the observed debris-flow tracks were located at the fringe of the most frequent simulated paths. Thus, a higher slope threshold and exponent of divergence are required to capture these fringe debris flows. Additionally, the surface of DEMs with resolutions greater than 20 m can be too general to capture minor gullies that may have high flow accumulation (Blahut et al 2010b). The 12.5 m resolution ALOS DEM used in this study is derived from downsampled SRTM data, and would likely contain some of the topographic generalizations of the original DEM (~ 30 m spatial resolution). Despite potential issues with DEM quality, similarly to Horton et al. (2013), we illustrated valuable results can still be achieved."

Technical corrections

p1, l5: fix typo in "Germany"

Corrected.

p1, l29: remove additional blank after "learning"; "source" instead of "sources"

Corrected.

p2, l30: remove additional blank after "and"

Corrected.

p2, l46/47: not sure if a comma should be used instead of a semicolon in the enumeration

Corrected.

p2, l55/56: add missing periods after "al" in three citations

Corrected.

p2, l56: remove additional blanks after "be"

Corrected.

p3, l73: "our" instead of "out"

Corrected.

p3, l77: add missing periods after "al" in three citations; Moreiras et al. 2012 and Serey et al. 2019 are missing in the references, please add

Corrected.

p3, l80: add missing period after "al" in the citation; Sepulveda et al. 2006 is missing in the references, please add

Corrected.

p3, l81: add missing period after "al" in the citation

Corrected.

p3, l83: add missing period after "al" in the citation

Corrected.

p4, l95: add period at the end of the table description

Corrected.

p5, l115: add missing "the" in "with ___ remaining set"

Corrected.

p6, l136: the PCM model was developed by three authors, so it isn't "Perla's" model, please rephrase

We will make this change.

p9, l208: throughout the text you use a hypen in "debris-flow", here you write "non-debris flow"; should this be changed?

Will rephrase to, "locations outside of the debris-flow polygons"

p9, l214: add the missing "a", the package is called "Rsagacmd"

Corrected.

p11, l243: "mass-to-drag ratio" not "mass-to-drag-ratio"

Corrected.

p14, l278: "towards a threshold of 0.5" instead of "thresholds"

Corrected.

p14, l279: There's quite a break between the two sentences, I had to read it twice to realize that "The resulting runout prediction map ..." was meant to be that with a threshold of 0.7. Maybe it would help to start a new paragraph here or to reformulate the sentence to something like "The runout prediction map resulting from the best threshold ..."

Thanks for the recommendation. We will rephrase it to, "The runout prediction map resulting from the best threshold".

p18, l314, Figure 11: "... runout path (a), ..." "... relative error (b), actual runout length error (c), and ..."

Corrected.

p22, l264: "source conditions to spatially" instead of "source conditions spatially"

Corrected.

p22, l373: "parameters of the PCM model" instead of "parameters the PCM model"

Corrected.

p24, References: please add the missing references and also have a look at the formatting - there are many references in which the author's first names are not shortened to the initials

Corrected.

References

Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., & Reichenbach, P. (2002). Impact of mapping errors on the reliability of landslide hazard maps. Natural hazards and earth system sciences, 2(1/2), 3-14.

Gamma, P.: Dfwalk – Murgang-Simulationsmodell zur Gefahrenzonierung, Geographica Bernensia, G66, 2000.

Horton, P., Jaboyedoff, M., Rudaz, B. E. A., & Zimmermann, M. (2013). Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale. Natural hazards and earth system sciences, 13(4), 869-885.

Planchon, O., & Darboux, F. (2001). A fast, simple and versatile algorithm to fill the depressions of digital elevation models. Catena, 46(2-3), 159-176.

Wichmann, V. (2017). The Gravitational Process Path (GPP) model (v1. 0)–a GIS-based simulation framework for gravitational processes. Geoscientific Model Development, 10(9), 3309-3327.

Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical chemistry, 39(4), 561-577.