Natural Hazards and Earth System Sciences Discussions

Nat. Hazards Earth Syst. Sci. Discuss., referee comment RC2
https://doi.org/10.5194/nhess-2022-177-RC2, 2022

Comment on nhess-2022-177

Anonymous Referee #2

Referee comment on "Development and evaluation of a method to identify potential release areas of snow avalanches based on watershed delineation" by Cécile Duvillier et al., Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2022-177-RC2, 2022

OVERVIEW

This paper proposes a new method for identifying potential release areas (PRA) for snow avalanches based on terrain characteristics and validates the approach using a long-term avalanche cadaster dataset. The research is situated in France, and the study area includes three mountain massifs in the French Alps. Overall, the topic is interesting and relevant for the avalanche research community and the NHESS readership.

While I appreciate the authors' desire to create a practical, transparent, and computationally efficient algorithm for PRA identification that uses easily accessible datasets, there are several substantial weaknesses in the present study that, in my opinion, prevent this manuscript from being a meaningful contribution to the literature in its current form. Properly addressing several of my concerns would require a substantial redesign and/or expansion of the study, and I am unsure whether that can be accomplished within the current peer-review process. I hope that the following comments can help the authors to further develop their research.

PRIMARY ISSUES

Selection of terrain characteristics and threshold for PRA identification

While the selection of terrain characteristics included in the PRA identification algorithm is based on existing literature, the reasons for their selection (or the refusal of other characteristics) are only discussed superficially. Furthermore, the selections of the parameter thresholds (e.g., 1400 m elevation threshold, incline range) do not seem to be

well grounded in evidence. I recommend that the authors conduct a proper grid search to determine the ideal parameter settings for their PRA identification algorithm. This is particularly important because they use a low-resolution DEM (25 m), which results in incline values that are biased towards lower values. This means that the thresholds described in the literature are not necessarily applicable. While the current sensitivity analysis might intend to do this, it is not done in a very rigorous and scientifically valid way. See additional comment on sensitivity analysis below.

Selection of datasets

Several datasets used in this study seem to be of lower data quality than established best practices in the field of PRA identification suggest. For example, the forest data set seems to have considerable limitations and the DEM is of much lower resolution than suggested in the literature. While I do not have a problem with a let's-do-the-best-we-can-with-what-we-have approach (not everybody has Swiss quality datasets available!), these choices need to be clearly explained and potential shortcomings evaluated and discussed.

Watershed delineation

While I appreciate the simplicity of the watershed delineation approach, delineating PRAs is not new. The OBIA approach described in Bühler et al. (2018) does the same thing in a more sophisticated way. In my opinion, Section 3.1.3 and Fig. 4 explain the calculation of the flow direction, but do not actually show how the watersheds are delineated. Since the authors' method uses standard tools available in open-source GIS software, it might be more useful for the reader to get a detailed description of how these calculations are done in freely available GIS software.

Validation of PRA identification

I see several fundamental challenges in the current validation approach that, in my opinion, provides a very biased perspective on the performance of the PRA model.

1) The authors' choice to only evaluate the performance of the model within areas of documented avalanches means that they only test whether the PRA algorithm can identify start zones in known avalanche path (true positives). It does not provide any insight about the algorithms ability to ignore terrain where avalanche do not start outside of the known avalanche paths (true negatives). While the authors explain their approach when they introduce their modified confusion matrix (L290+), this does not seem to be very meaningful to me. As explained by the authors, avalanche cadaster datasets are not widely available and have limitations in many areas. The purpose of PRA models is to identify PRA in areas where direct observations are not available. In my opinion, a more meaningful approach would be to validate the model in areas with high confidence in the

avalanche mapping record and include both avalanche terrain and non-avalanche terrain so that the complete confusion matrix can be properly evaluated. As can be seen in Fig. 3 and 7, there are considerable areas outside of the avalanche path areas that the algorithm incorrectly identifies as PRAs.

2) Applying the PRA model steps (e.g., > 1400 m, slope incline between 28 and 60°, watershed delineation, etc.) to the CLPA dataset before conducting the validation completely defeats the purpose of a validation. Obviously, the model will perform well if the validation only includes terrain with the same characteristics. In the end, the authors only evaluate the steps in the PRA algorithm that are not included in the CLPA preprocessing (slope curvature?). This is a fundamental weakness of the paper.

3) The simplified confusion matrix and calculation of the accuracy and error rates derive directly from the authors' choice of only examining true positives and false positives. As pointed out above, I do not think this is meaningful. Simply assuming that the true negative is 100% is not meaningful and leads to inflated accuracy rates. It is also unclear to me why the authors use percentages in their confusion matrix calculations. Confusion matrices are general populated with counts, which is possible for evaluating both the identification of individual PRAs and the total area of PRAs.

4) Nowhere in the manuscript is explained how the author identify a match between a PRA identified by the algorithm and the validation dataset. Is 100% overlap required or do the authors use a different rule to distinguish true positives from false positives?

5) Only focusing on the accuracy rate is a very simple evaluation of performance. Furthermore, since the error rate is simply the complement to the accuracy rate, having the error rate in all the tables does not add any value. The use of this simple validation measure is very much at odds with the content of the paragraph on model evaluation in the introduction (L87+), where the authors seem to highlight the value of more advanced evaluation approaches. This seems a missed opportunity for contributing to the literature.

6) The repeated statement that the validation in this study is done over large areas of terrain (i.e., entire massifs) is incorrect. The validation was conducted within documented avalanche paths within these massifs, which, as highlighted in Fig. 7, are generally very small areas.

Challenges in sensitivity analysis

As mentioned above, the sensitivity analysis does not seem sufficiently rigorous to provide meaningful insight. For example, the authors only compared the benefit of the 1400 m elevation threshold to not having an elevation threshold at all. Why not test other threshold values (1300 m, 1500 m, etc.)? The sensitivity analysis also only examines

certain parameters and leaves out others without explanation. In my opinion, properly deriving the parameter selection and thresholds from the available data is a critical piece that needs to be included in this paper.

Given that the authors use a low-resolution DEM that is far below the quality recommended in the literature, it seems critical that this choice is justified with a proper sensitivity analysis. While the performance of the lower-resolution DEM will likely be lower, its use can still be justified based on a cost-benefit argument, but it will be important to have the comparison to better understand the consequences of this choice.

Some of the main results of the sensitivity analysis do not seem to match common sense. The fact that including slope incline increases the accuracy rate by less than 2% seems wrong as incline is one of the primary determining characteristics of avalanche terrain. In my understanding, this odd result is the direct consequence of only looking at terrain within documented avalanche paths and processing the validation dataset with PRA algorithm rules, which clearly highlights the limitations of the validation approach taken in this study.

Comparison with other algorithms

In my opinion, comparing the algorithm introduced in this paper with some of the established methods would be very important for highlighting the value of the new approach. I think this should be included in this manuscript.

Limited discussion

In its current form, the discussion primarily repeats information from earlier sections of the manuscript (i.e., introduction, methods and results) without adding much value. This is partially due to the fundamental limitations mentioned above. The tone of the discussion is also quite casual (e.g., L529: "All in all, the validation data we use is certainly not perfect and our validation approach may potentially favour the comparison with our detected PRAs, …"), which does not seem appropriate for a scientific publication. See the technical comments section for additional comments. The outlook section does not seem to offer any novel ideas as it primarily discusses already existing application cases for PRA maps (e.g., large scale mapping of avalanche hazard and risk) and existing research extensions (e.g., PRAs conditional on snow and weather conditions, probabilistic detection rules).

SECONDARY ISSUES

Set up of the research objective and expectations

To motivate their study, the authors provide a fairly comprehensive, even though not completely up-to-date, summary of the existing literature on PRA identification in the introduction. In this overview, they identify several limitations of the existing approaches (e.g., disagreement about relevant terrain factors, delineation of individual PRAs, validation of PRA algorithms) to set the stage for their research objective on L103. This setup creates the expectation (explicitly or implicitly) that the algorithm introduced in this manuscript will address these issues. There are multiple issues with this. First, I do not agree with all the claims that are made in the introduction. The terrain parameters included in the various PRA algorithms do not differ from each other that much, Bühler et al. (2018) have presented an approach for meaningfully delineating PRAs, and the most cited algorithms have been validated with mapped avalanche datsets. Second, the paper does not deliver on these expectations due to the methodological issues mentioned above. This results in disappointment and sets the paper up for failure.

I think the manuscript would benefit from a more focused introduction that describes the research objective more honestly and positions the study within the existing literature more accurately. I have no problem with a study that aims to create a simple approach for PRA identification based on easily accessible datasets, but this objective should be clearly stated at the beginning of the paper to create meaningful expectations.

Language and tone

Overall, the quality of the English in this manuscript is not very high, and the text includes many terms that are not meaningful in this context (see partial list in technical comments). While the use of these terms might the result of a poor translation from French, it is important to check the manuscript for proper use of terminology before publication.

Furthermore, the tone of the writing is rather ambitious and glowing. Examples include "In this paper, a method that well identifies …" (L16 and L103), "… the CLPA is a very valuable source of information, … and, arguably, among the rare existing ones …" (L197), and "… the CLPA is almost surely a true avalanche extent." (L188). The further exasperates the reader's sense of unfulfilled expectations mentioned above.

Figures and tables

The figures included in this paper are not of high quality. Several are hard to read (e.g., Fig. 2), and the figure layout and formatting of the legends seem different in every figure. In my opinion, Fig. 5 and 6 do not add any value beyond what is already explained in the text. It is also unclear to me why the validation maps for the Mont-Blanc and Maurienne

massifs are currently in the supplementary material and not included in the main manuscript.

Captions of tables are typically presented above tables. In all confusion matrix tables, the different components of the confusion matrix should be properly labelled.

TECHNICAL COMMENTS

Abstract

L10: 'lacunar' is not a meaningful word in this context.

L18: 'Confrontation' should be 'Comparison'.

Introduction

L30: Extremely convoluted sentence.

L64: 'Retained' should probably be 'included'.

L83: 'on the field' should probably be 'in reality'.

L98: I am not sure whether the existing PRA algorithms are 'competing'.

L104: Should be '… where avalanches can occur'.

L108-120: This preview of the methods and results is not necessary in the introduction. It is best to finish the introduction with the statement of the research objective.

Data

L134: ‘reputed’ should be ‘well known’.

L176: All abbreviation need to be properly introduced the first time they are used in the manuscript. There are additional abbreviations that have not been introduced.

L176: It should be ‘It consists of…’.

L179: I don’t not understand what is meant with ‘… is mainly produced at the destination of …’.

L192: ‘… near human stakes …’ should be ‘… near human assets or settlements.’

L225: ‘In order to …’ can be simplified to ‘To …’. There are several cases of this in the manuscript.

L241: The sentence that describes how the thresholds and parameters are chosen (reference to Sect 4.2) does not seem to belong here.

L245+: The description of the algorithm provided in this section does not seem consistent with the information presented in Fig. 5.

L261: The statement that PRA identification target primarily large avalanches needs to be stated much earlier in the manuscript as it is a fundamental assumption of the study.

L271+: The description of the CLPA seems repetitive as it discusses information that was mentioned previously already.

L280: Typo: PRAS should be PRAs. There are several instances of this typo in the manuscript.

Results

L327+: The description of the confusion matrix provides exactly the same information that is shown in the table. Hence, it does not add any additional value.

L349: I do not know what you mean with ‘probative’. There are several uses of this work in the manuscript.

L350+: The explanation provided here seems rather speculative and not well grounded.

L371: Not sure what you mean with ‘parametric study’.

L380: In academic writing, the term ‘significant’ should only be used in the context of statistical significance. Use ‘considerable’ or substantial’ instead.

L382: ‘Use ‘more substantially’ instead of ‘more largely’.

L406: The last sentence in this paragraph is too hand-wavy and not grounded in evidence.

L410: Table 6 does NOT show that not including the forest layer results in the worst performance. The accuracy rates without forest are higher than with the theia dataset.

L421: Delete ‘eventually’.

L425: Use ‘tower’ instead of ‘pylon’.

L435: Delete ‘eventually’.

Discussion

L447+: There is no need to repeat the information from the intro at the beginning of the discussion section.

L455+: I do not think these objectives have been achieved by this study.

L469: I am not sure what is meant with ‘probative’.

L475+: This discussion primarily repeats information from the methods and results section without adding much value.

L499: Reword this sentence. It should be ‘… by comparing X to the processes CLPA dataset.’.

L503+: This sentence seems like an excuse and is not very convincing. The readers’ expectations should be managed by properly describing the research objective in the introduction.

L507: ‘Envisaged’ should be ‘envisioned’.

L510: ‘Confronted’ is the wrong word here.

L529: Similar to the sentence on L503, this sounds like an excuse. Limitations should be discussed more seriously.