Response to reviewer comments

Sykes, J., Haegeli, P., and Bühler, Y.: Automated snow avalanche release area delineation in data sparse, remote, and forested regions, Nat. Hazards Earth Syst. Sci. Discuss. [preprint], https://doi.org/10.5194/nhess-2021-330, in review, 2021.

Note to Editor

Based on your suggestion in the ‘comments to the author’ to add sources of maps in figure captions we reviewed all our maps and added appropriate sources. The majority of the map data was generated by the authors from raw data sets and according to the NHESS submission guidelines for maps and aerials it does not require sources. Any maps that used data provided by an educational license with a commercial company or using open access datasets were noted in the captions.

Reviewer 1 - Markus Eckerstorfer

Overall

Accepted subject to minor revisions.

First off I would like to commend you on an interesting and generally well-presented study. I took the opportunity to review your study based on my interest in using EO data for avalanche applications as well as based on the opportunity to learn a couple of new things, especially about deriving PRAs.
I am suggesting minor review as I do not see any major shortcomings in your study or how you present it. I would like you to consider some suggestions of restructuring paragraphs as well as providing a bit more information on certain technicalities. I commented on these issues directly in a pdf version of your manuscript that is attached.

Author response:
Thank you for your generous evaluation of our research. The comments included in the PDF supplement were very useful, but largely focused on specific edits to the manuscript regarding figures, writing structure, and technical details so we have not addressed them specifically here. We have addressed those comments and implement the changes you suggested in our revised manuscript. Please see the comments marked with ‘ME supplement’ in our track changes manuscript to see the revisions.

General Comments

1.1 PRA Validation and Updating

How often should one update PRAs based on changes in vegetation and snow climate, especially in the light of rapid climate change? And how good are avalanche observations then anyway for validation purposes?

Author response:
Considering impacts of climate change on PRA modeling is an interesting suggestion that we neither considered nor plan to include as an added topic in our discussion. Overall, we believe that changes in vegetation and snow climate across different regions has a large impact on the optimal input parameters for PRA models. Therefore, it follows that gradual change in vegetation and snow climate within a single region could lead to changes in the characteristics of potential release areas. Likely changes can be split into two categories: 1. Landscape changes that affect the extent of PRA due to moving treeline and growth of new vegetation. 2. Snow and avalanche climate changes due to weather pattern shifts and increased regional temperatures.

Gradual landscape changes due to vegetation cover would not require the PRA input parameters to be updated because the nature of the terrain is remaining constant. Simply updating the PRA model using a more recent forest layer could help make the extent of the PRA more accurate if there have been significant changes in forest distribution and density. Evaluating the rate of forest change could be accomplished by producing a new forest classification roughly every 10 years and comparing it to the original forest data. If significant changes have occurred the PRA model could be updated, otherwise there is no need to allocate computer resources to reprocess PRA models.

The impacts of climate change on snow climate are likely to be complex and variable across different mountain regions. Existing research on this topic has found that the largest changes are at lower elevations where the boundary between snow and rain for precipitation events has a large impact on snowpack development. Another suggested impact of climate change is more frequent high intensity storms which could increase the frequency and magnitude of avalanches at upper elevations. The impact of these changes on PRA modeling are not obvious, therefore a reevaluation of the optimal PRA model input parameters could be warranted if notable differences in snow and avalanche climate are observed in a given area.

Avalanche observation data is the best method we have to validate PRA models but they are still limited. Two primary reasons for the challenge of validating PRA models are: 1. A general lack of long term spatially accurate avalanche observation datasets worldwide. 2. Even the most complete datasets available are limited to a roughly 50 year time scale which does not necessarily capture the full spectrum of snowpack and avalanche conditions that are possible. Current efforts to map avalanches by satellites might be a good option to complement existing databases but have to be extended into further regions (e.g. Eckerstorfer et al. 2017, Bühler et al. 2019).

PRA models aim to capture all terrain that is capable of producing avalanches, which sometimes means that the output of a PRA model does not align with observed avalanches simply because the snowpack conditions necessary to cause an avalanche on a specific terrain feature has not been observed. This is a challenge of relying on human observations for validation and can create conflicts between the results of a PRA model and the local knowledge from a given area. This is an important caveat when producing PRA models and one that we will be sure to reinforce in the updated manuscript.

Despite these limitations, avalanche observations are the best data available to validate PRA models in our study area so our approach aimed to create a validation dataset based on the experience of local experts as a substitute for long term records. We believe this is a meaningful substitute that could be reproduced in many regions worldwide to help define optimal input parameters for PRA models in different ecological and snow climates.

Manuscript revisions: While we appreciate the comment, we are not planning to make any changes to our manuscript in response to it because it is outside the scope of this mansuscript.

Refs:

Bellaire, S., Jamieson, B., & Statham, G.: The avalanche climate of Glacier National Park, BC, Canada during 1965-2011. In Proceedings of 2013 international snow science workshop, Grenoble, 2013.

Bühler, Y., Hafner, E. D., Zweifel, B., Zesiger, M., and Heisig, H.: Where are the avalanches? Rapid SPOT6 satellite data acquisition to map an extreme avalanche period over the Swiss Alps, The Cryosphere, 13, 3225-3238, 10.5194/tc-13-3225-2019, 2019.

Eckerstorfer, M., Malnes, E., and Müller, K.: A complete snow avalanche activity record from a Norwegian forecasting region using Sentinel-1 satellite-radar data, Cold Regions Science and Technology, 144, 39-51, https://doi.org/10.1016/j.coldregions.2017.08.004, 2017.

Laternser, M. and Schneebeli, M.: Long-term snow climate trends of the Swiss Alps (1931-99), Int. J. Climatol., 23(7), 733–750, doi:10.1002/joc.912, 2003.

Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., Marchetti, M. Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. Forest ecology and management, 259(4), 698-709. 2010.

Wilbur, C., & Kraus, S. Looking to the future: predictions of climate change effects on avalanches by North American practitioners. In International Snow Science Workshop Proceedings (pp. 7-12), Innsbruck, 2018.

1.2 Cost effectiveness

You talk about the cost effectiveness of Spot 6/7 data. In my experience such data is expensive and for many avalanche operations and forecasting centers not financeable. Do you see other EO-based data sources (not only optical but also radar) that could be used?

Author response:
This comment touches on a common theme from all three reviewers, which is that high resolution satellite data is expensive and not feasible for most avalanche operations. This is true, and we will revise the manuscript to more accurately represent that our method is cost effective compared to other options for generating a 5m DEM. Unfortunately, a high resolution DEM is the fundamental dataset necessary for producing a high quality PRA model and there is no easy way around this. We evaluated as many options as we could to create a DEM in our study area and found that our workflow was anywhere from 2-10x less expensive than alternatives, although it required more technical knowhow to produce a DEM from raw imagery.

To our knowledge there are no freely available stereo imagery datasets that could be used to produce DEMs with 5m resolution. Producing DEM data from radar imagery is beyond our expertise, but to our knowledge the resolution of freely available data is insufficient to create the requisite 5m DEM. As satellite imagery continues to decrease in price over time and more sensors get available, the generation of high resolution DEM data will be more affordable in the future. Furthermore, national scale LiDAR mapping is becoming more and more common and could provide high resolution input data for those countries in the future.

Manuscript revisions: We added detailed information about data acquisition costs to our methods section (lines 184-187, 194-196) and conclusions (lines 740-742). We also suggested an alternative EO dataset for forest characterization in lines 660-663.

1.3 Representativeness of validation polygons

How representative are the validation polygons for the entire tenure? Are these runs characteristic for most of the other runs in the area in terms of topography, vegetation, and snow climate? From Figure 1 it seems that the validation runs are nestled in the northern part of the tenure only.

Author response:
We relied largely on the expertise of our local guides to select validation runs that are representative of the terrain across the study area. We targeted runs that had a combination of alpine and tree line terrain and were used frequently by the guiding operation so that the guides are very familiar with the terrain in a variety of snowpack conditions. The fact that our 5 runs were clustered in one corner of the tenure was not intentional and is a reflection of the operational realities of the guiding operation.

Manuscript revisions: We expanded the discussion of how we selected validation runs slightly in lines 324-328, and tried to highlight that we relied on the local expertise of our collaborating guides to select the most representative validation runs.

1.4 Services for DEM creation

It seems to me that there is quite some technical knowledge needed for DEM creation based on EO data. Are you aware of any services that could provide such data to overcome this technical threshold in order to make your method easier and wider applicable?

Author response:
There are several commercial software programs that can produce DEM based on stereo imagery, but they require costly subscriptions and technical knowhow. We opted to use open source software because of the well documented methods, prior research demonstrating high accuracy of DEMs using similar stereo imagery, and the ability to minimize future costs to develop DEMs using our documented workflow. The technical details of our DEM processing workflow are captured in our supplementary material, which includes a general description of our approach and copies of our code for generating the DEM.

An alternative solution is to purchase a commercial satellite stereo DEM from imagery providers such as Maxxar, Airbus, or ALOS, but these come with added costs and less control over the DEM characteristics. The time and energy saved by purchasing an off the shelf DEM product is significant, so this is not a bad option depending on resources available.

Manuscript revisions: We added a discussion of alternative methods for DSM production in lines 183-187 and 456-462. We also aimed to change the tone of the manuscript overall to acknowledge the technical skill required to generate DSM data from raw imagery while still justifying it as the most cost effective approach available during our data acquisition.

1.5 Flow chart of complete workflow

I am missing an overall flow chart of all the technical steps you carried out to derive PRAs. Would it be possible to provide such a figure since you talk a lot about how technically trivial all the work is?

Author response:
An overall workflow graphic was suggested by multiple reviewers and we plan to create one to help conceptualize our research workflow. We will also focus on limiting the discussion of how 'technically trivial' the work is, because that is not an entirely accurate representation of the process of generating the input data necessary to create a PRA model. Creating the input datasets requires technical knowhow, however if these data are available applying the PRA model is a relatively simple process described in Bühler et al. 2018.

Manuscript revisions: We added figure 1 (line 160) as an overall workflow graphic based on comments from all reviewers. We also tried to reference this overall workflow throughout the methods section to help clarify and simplify the methods.

1.6 Improvement of accuracy

After reading the methods section I expected a higher improvement of accuracy from Bühler 2018. You claim that an improvement of over 11 % is highly satisfying and I must believe you. What I am wondering, however, is if the additional costs and technical steps warrant the improvement, especially if you would calculate PRAs for the entire tenure of a BC forecasting region? I am also wondering what more / other things could be done to improve the accuracy of these PRAs even further, given that you err on the side of overestimation (I agree with you on that!).

Author response:
Your point about whether the improvement of our model is significant compared to the original Bühler 2018 version is noted and was suggested by other reviewers. We plan to add some context from prior research around these results to help reinforce them in the updated manuscript.

In addition, there are actually no added steps necessary to include our forest density methods when running the PRA model. The prior version used an optional binary forest mask to eliminate forested areas from the PRA model calculations. Our version uses that same forest mask to automatically calculate forest density and requires minimal additional processing time. The only added work for the user is to determine the input parameters for the forest density and forest slope scalar variables. In the updated manuscript we will clarify this point. The added workflow diagram we plan to create should also help clarify that our research has minimal impact on running the PRA model.

In regards to what other things could be done to improve accuracy, we have previously commented on statistical modeling of forest characteristic variables as a way to improve performance. Other main sources of error are the fact that we are using a digital surface model, therefore the terrain characteristics are not based on the bare ground surface topography. Using LiDAR to collect a digital terrain model would improve accuracy, but at a great financial cost.

The fact that the overall accuracy statistics in our research are fairly low (57.5%, 45.8%) is also partially due to our validation dataset and methods. There are unavoidable uncertainties when relying on human experience and recollection to create a spatial validation dataset. We attempted to minimize these and be transparent about them in the manuscript.

Manuscript revisions: We added a more complete background section on PRA validation to help develop the context of our results (lines 99-106). We also added references to this prior validation work in the results (line 569) and included specific ideas for further improvement to modelling forest character using satellite data in lines 658-663.

1.7 Extension to other snow climates

Finally, it seems to me that your study area and the area around Davos are not so different in terms of topography, vegetation, and snow climate. Could you maybe discuss the performance of your method in a more maritime setting? Finally, what would you suggest doing if there would be no local knowledge to determine avalanche terrain?

Author response:
Extending the model to additional snow climates is one of the main points of our discussion and conclusion section which we can expand upon further. This is a future research need that we can provide some insight on, but additional validation data is required to adapt the PRA model to a maritime snow climate. In a maritime climate the slope angle threshold would likely need to be increased based on generally deeper snowpack and less common persistent weak layers.

Further research using additional validation data in a maritime snow climate would enable the definition of unique default input parameters for different snow climates would help future researchers apply the PRA model in areas without validation data available. Our research highlights that only a few parameters changed between Galena and Davos, which supports that the defaults defined in the original Bühler et al 2018 paper are valid for other mountainous areas. The biggest changes are in slope angle distributions, which we believe to be driven by common weak layer types. Experiences with applying the model to different areas in Alaska, the Indian Himalayas, Afghanistan or New Zealand indicate that the model is well applicable to different snow climates and topographies. Additional research is necessary to refine the forest density layer generation for different ecological climates.

Manuscript revisions: While we appreciate the comment and opportunity to hypothesize about the application of the PRA model in other snow climates, we are not planning to make any adjustments to the manuscript as this topic seems beyond the scope of the current study. We suggest this as an area for further research in lines 749-750. Further, we added a section about generalizing our validation approach to new areas without existing validation data in lines 694-708.

Specific comments

Author Note:
The comments within our manuscript were largely specific to writing structure and appearance of figures which we will address in an updated manuscript but will not respond to individually in this document.
Manuscript revisions: The comments from the reviewer 1 supplement were all addressed in the manuscript and are labelled in the track changes document with 'ME Supplement'.

Reviewer 2 - Ross Purves

Overall

Reconsidered after major revisions

I enjoyed reading this paper which aims to apply a method developed in Switzerland to a Canadian test site. The underlying aim of the method is to identify potential release areas (PRAs), using a combination of terrain and vegetation data. PRAs are an important and useful tool in hazard mapping, and the paper describes a complete workflow for their derivation. In practice this means that the authors not only calculate PRAs, but they also generate a terrain model and a forest mask (which is non-binary). Overall, the paper seems to me to make a useful contribution and I think it could be published in NHESS.
However, I have some reservations about the paper in its present state. At the moment the approach taken is to my mind a very linear one, and although some experiments are made, it is difficult to see how the approach could be generalized in practice. Furthermore, in places the manuscript is overly wordy, which I think obscures the message, while in others the methods are described in insufficient detail to allow their replication. The authors have chosen to incorporate results and discussion in one chapter which makes teasing out the implications for the wider field unnecessarily hard, and leaves the impression of a discussion of the specifics of this case study rather than the purported much more general aim claimed by the authors. I have several specific suggestions for the authors.

Author response:
Thank you for your careful review and thoughtful comments. We have addressed many of the specific concerns mentioned here in more detailed comments below. Overall, we agree with the assessment that the manuscript could be improved considerably by simplifying and clarifying the writing and providing a better overview of how our methods fit in the broader research field. We plan to address all the concerns mentioned here in the updated manuscript and reevaluate the overall structure of the manuscript.

General Comments

2.1 Grid search on all parameter

The results reported compare, if I understand correctly, an optimized version of the PRA model including forest cover with the original Bühler et al. model. However, the original model uses the parameters derived in Davos. Given that the improvement is in any case rather small, wouldn't it be informative to do the grid search on the data from the Canadian test site and compare performance. In general, when we optimise models and compare, we should (in my view), always optimise our baseline too. You argue throughout the paper that the higher resolution data are needed. But you could easily show this by also running your model with the 18m DEM and 30m landcover dataset. It would be interesting to see how different the results really are.

Author response:

In regards to whether our model represents a significant improvement from the original Bühler 2018 version, the broader context of PRA modeling accuracy is important to consider. The earliest version of a PRA model simply considered all slope angles between 28-60° potential release areas. This method is highly effective at capturing all possible avalanche terrain, but has a tendency to overestimate the extent of PRA. More sophisticated modern PRA models use additional inputs such as curvature and ruggedness to make marginal gains in accuracy over this simple slope only method.

In prior validation research (Bühler et al 2018) modern PRA models demonstrate lower probability of detection (POD) rates compared to the slope only method but are able to improve upon probability of false detection (POFD) rates, leading to an overall more skilled model. To provide some context for the scale of improvement, the original Bühler 2018 model had roughly 13 percentage point improvement in POFD compared to the slope only method with a roughly 3.5 percentage point loss in POD. In that same validation research another modern PRA model (Veitinger et al. 2016) showed an improvement of roughly 9 percentage points in POFD rate with a roughly 2 percentage point loss in POD compared to the slope only model.

These results highlight the fact that PRA modeling is a field of marginal gains. Unfortunately, due to the nature of our validation data we cannot directly compare the accuracy metrics from our model to the prior validation research. However, in the context of this research field our 12 percentage point increase in accurate validation polygons and 15 percentage point decrease in underestimated validation polygons represents a significant improvement over the original Bühler 2018 model.

In regards to optimizing the input parameters for all the PRA model terrain characteristics, we did perform the grid search across all parameters and found that for slope angle maximum, ruggedness window, ruggedness maximum, and curvature maximum our validation data did not produce an input parameter with clear advantages over the Davos defaults (Figure 8). Therefore, we retained the default values due to the larger validation dataset available in Davos. Our validation effort serves as an example that changes in these input parameters may not impact the accuracy of the PRA model in different regions and the default settings may be universally applicable (also discussed in response 1.7 from Reviewer 1).

Finally, the analysis of how DEM resolution impacts PRA models has been carried out previously in Bühler et al 2013 and 2018, which indicated an optimal resolution of 5m. That preexisting research led us to adopt the optimal DEM resolution instead of performing this test ourselves and repeating the analysis. Instead our focus is to build upon the existing model of Bühler et al 2018 validate it in a new mountain region and forest ecosystem.

Manuscript revisions: We have expanded our discussion of prior PRA validation to highlight that our improvement over the original PRA model represents a significant increase in accuracy for forested terrain (lines 99-106). We also expanded our

explanation of the grid search methods (lines 156-158, 390-394) to show that we tested many iterations of baseline parameters before selecting our final values.

2.2 Cost

You argue regularly throughout the manuscript that your approach is transferable because you use open source software. But you bought data, and in fact you apply the method to a very small region in the end. I think more transparency about data costs is important, since this is likely a show stopper in many regions.

Author response:
This comment is similar to comment 1.2 from Reviewer 1. As stated in that response, we plan to place more emphasis on the relative costs of our method compared to alternatives for generating high resolution DEM data in the updated manuscript.

Manuscript revisions: We added detailed information about data acquisition costs to our methods section (lines 184-187, 194-196) and conclusions (lines 740-742).

2.3 Wordiness

The methods are in places very wordy, and though I like the idea of introductory sections describing the overall approach, which is then broken down into more detail, you are often very repetitive. Furthermore, lots of tables contain information which is then repeated verbatim in the text. I would suggest:

- Thinking about a graphical summary of your methodological approach as a whole and linking this to sections
- Removing or reducing obvious repetition. For example, I think it would be more useful to give parameter values in tables and cite these (rather than duplicate) in the text.
- The grid search (for which you simply cite Bühler (N.B. There are multiple 2018 Bühler papers in your references, you need to tidy up the citation)) is not described in sufficient detail to allow replication. I also couldn't find a clear description of this in the Bühler NHESS paper, but I am not sure if I looked in the right place. Since this is central here, I suggest describing the method more completely. By cutting other unnecessary text this shouldn't make the paper longer, but it will be much more useful to the reader as a standalone piece.

Author response:
Creating a workflow graphic to help illustrate our methods was also suggested by Reviewer 1 and we plan to include this as a new figure in the updated manuscript. Specific comments in the manuscript from Reviewer 1 also suggested changes in the ordering of writing sections and removing repetitive text to help clarify our descriptions of the research methods. These comments are very useful to help improve the communication of our research, and we plan to focus heavily on streamlining and simplifying the writing in the updated manuscript.

Your comments about clarifying the description of the grid search are very helpful, and we plan to update the manuscript to better highlight that we used independent methods from Bühler 2018 due to fundamental differences in the structure of our respective validation datasets. In addition, we have provided open access to our grid search processing scripts in an Open Science Framework repository that are available for readers that are interested in understanding the technical details of our methods.

Manuscript revisions: We reviewed the entire manuscript for wordiness and deleted repetitive sections – see comments with 'removed for conciseness' in track changes document. We also created Figure 1 to address the need of an overall workflow graphic (line 160). We added detail to our description of the grid search methods used to determine optimal parameters (lines 156-158, 390-394) and updated our references to make sure the appropriate paper was listed. Throughout the manuscript we included references to our Open Science Framework repository where our code and data are available for interested readers to understand the technical details of our methods.

2.4 Generalizability

The validation approach taken is interesting, but not really scalable in my view, which also limits the generalisability. Canada is rather special in having guides who repeatedly descend the same routes winter after winter, and discuss these in comprehensive guides' meetings, often on a daily basis. How about (at least in the discussion) discussing how other data could be used, for example in a Canadian context those derived from InfoEx?

Author response:
Expanding our discussion to include alternative methods of generating validation data is an excellent suggestion and adds value to our research by providing some guidance on how others could apply our method in new regions.

In our view the high level approach of working closely with local avalanche experts to generate a dataset of known avalanche release areas while accounting for the innate uncertainty in relying on human recollection to map avalanche boundaries is scalable to many different locations and types of operations. Many mountain regions across North America and South America have mechanized guiding operations that repeatedly descend the same routes over multiple seasons which could directly implement our methods.

In addition, other types of local experts such as backcountry avalanche forecast centers, roadway, railroad, or powerline avalanche forecasters, or ski patrols based in ski resorts have similar local expertise that could be leveraged to create a validation dataset in a similar manner to the methods we applied. In that sense our methods are generalizable because they rely on human expertise, which exists worldwide but has limitations that we have attempted to account for.

Manuscript revisions: We added a discussion of the basic principles we used to develop our validation dataset in the interest of making the methods more generalizable for other areas (lines 694 – 708)

2.5 Splitting of Results and Discussion

I would strongly suggest splitting results and discussion. I also note that the papers' objectives are rather implicit. If you more explicitly stated these in the introduction, then you could write a more standalone discussion. I'd be particularly interested in more comment on how realistic it is to extend your method to areas in other countries and some discussion of the likely costs.

Author response:
Thank you for highlighting that we should be more explicit about stating our research objectives to help provide a clearer structure for the manuscript. We initially elected to combine the results and discussion because we felt that directly comparing our results to the output of the original Bühler 2018 PRA model would simplify the structure of the manuscript. After reviewing the manuscript again in light of the reviewer comments we feel that it would be possible to separate the results and discussion but that the decision comes down to a personal preference. We plan to reevaluate our initial decision after updating the manuscript to address the comments of the three reviewers. The added discussion topics suggested by the reviewers could provide more content for the discussion section and justify switching the format of the paper, but we reserve the right to make the final decision after updating the manuscript.

Manuscript revisions: We feel that our research objectives are stated explicitly in lines 67-74. We added lines 694-708 and 740-742 to address the costs and considerations for applying our methods different areas. After reviewing the manuscript in its entirety, we elected to keep the original format of combing the results and discussion section. We made this choice because discussing the context of the original PRA model is central to contextualizing our results. We also feel that the additional discussion topics that we included in the manuscript based on reviewer feedback fit within the existing structure. We organized each section of the results/discussion section with results in the first paragraph(s) followed by discussion so that readers can easily find information within the manuscript. We also noted this structure in the introduction to the section to inform readers of our intended format.

Specific comments

2.6 Number of references

In many places you use multiple references for rather simple points. Often 1-2 references would suffice.

Author response:

We plan to review the manuscript with an eye for excessive citations and remove any that are redundant or unnecessary.

Manuscript revisions: We removed several references from the paper and reviewed the entire manuscript for areas that we could decrease the number of in text citations (lines 53, 80-81, 117, 141, 144, 148, 660, 760-761, 763).

2.7 Representativeness of validation polygons

The validation polygons are all concentrated in a rather small area. What are the implications of this?

Author response:
This is similar to comment 1.3 from reviewer 1. Our justification for selecting the validation runs is explained in that response.

Manuscript revisions: We expanded the discussion of how we selected validation runs slightly in lines 324-328, and tried to highlight that we relied on the local expertise of our collaborating guides to select the most representative validation runs.

2.8 Influence of surface hoar weak layers

On line 164 you say surface hoar is amongst the most common weak layers associated with avalanche problems. Later (L321) you say areas dominated by such problems were minimized in the validation dataset. Why?

Author response:
This point was also suggested by reviewer 1 in their supplemental comments on the manuscript.

Surface hoar is a unique weak layer because it can cause avalanches on much lower slope angles than any other. Creating a validation dataset that overemphasizes release areas that can only produce avalanche with surface hoar weak layer would have biased our dataset to being overly conservative, therefore we asked the guides not to focus on release areas with this characteristic. Our goal of creating a PRA model for a frequent avalanche scenario drove that decision.

Manuscript revisions: We expanded the discussion of excluding very low angle start zones from our validation dataset in the methods section (lines 342-345) and we discussed the implication of this choice in the results/discussion section (lines 583-590).

2.9 Avalanche problems in Davos

You don't comment on the avalanche problem in Davos, though you claim the parameters are rather universal (L452). This depends to some extent on avalanche problems, so wouldn't it make sense to discuss these.

Author response:
This is an excellent suggestion and we plan to add a discussion of common avalanche problems in the Davos region to the manuscript. However, the avalanche climate research that was carried out for western Canada which revealed the common avalanche problems we discussed is not necessarily available in other areas. Based on the literature we are able to find we will expand the discussion as much as possible. Based on local knowledge, the four avalanche problems in Davos are new snow, windblown snow, old snow (depth hoar), wet snow and glide snow. It is worth noting that the Columbia Mountains are a known surface hoar hotspot, with a higher occurrence of this type of weak layer than most regions where the avalanche climate has been analyzed.

Manuscript revisions: After reviewing available literature we were unable to find adequate information about the avalanche problem characteristics from the Davos area. In the absence of that information, we highlighted that our study area is highly prone to surface hoar avalanche problems compared to other areas that have been studied previously and therefore the decreased slope angle minimum threshold makes sense due to the relatively frequent presence of this type of weak layer (lines 579-590).

Refs:

Techel, F., Müller, K., and Schweizer, J.: On the importance of snowpack stability, the frequency distribution of snowpack stability, and avalanche size in assessing the avalanche danger level, The Cryosphere, 14, 3503–3521, https://doi.org/10.5194/tc-14-3503-2020, 2020.

Laternser, M. and Schneebeli, M.: Long-term snow climate trends of the Swiss Alps (1931-99), 23, 733–750, https://doi.org/10.1002/joc.912, 2003.

2.10 Table 3

I don't understand how the statement on L519 where you state that your method is correct for 57.5% of the polygons and then say this means it is within +/- 12.5% of the area specified by the guides. Table 3 doesn't help me here…

Author response:
Thank you for bringing this to our attention. We will clarify our description of the grid search method to explain why we use +/- 12.5% as a threshold for defining accurately captured polygons. This is part of a broader effort to review, shorten, and clarify our methods section which was highlighted by all reviewers as an area for improvement.

Manuscript revisions: We rewrote our explanation of how we determine accurate, underestimated, and overestimated polygons to attempt to clarify this point for the reader (lines 372-373, 545-546).

2.11 Figure 13

Figure 13 is really nice. But the very specific discussion of individual runs doesn't speak to a very generalisable method.

Author response:
We agree with this statement and plan to reword our manuscript to highlight that our research is a case study for applying an existing PRA method to a new area and that our research needs to be expanded to demonstrate that the method is generalizable. In addition, we will attempt to place the discussion in section 4.4.1 in a broader context that could be more generalizable to forest characteristics in other regions.

Manuscript revisions: We kept the discussion in section 4.4.1 largely intact because it highlights the biggest potential source of error for our remote sensing based method. While this discussion is specific to the forest characteristics of two specific runs, we feel that those runs are examples of common forest characteristics in many mountainous regions. The specific discussion of the sources of error are intended to highlight areas for future improvement of PRA modelling in forested terrain and are therefore valuable directions for future research. We believe our approach is generalizable but have only tested it in one location, so being transparent and specific about our sources of error is in the interest of helping future researchers test and develop our methods in other areas.

Reviewer 3 - Erich Peitzsch

Overall

Accepted subject to minor revisions

In this study, the authors use an existing potential avalanche release area (PRA) model and, with important modifications, apply it to forested avalanche terrain for a remote portion of the Columbia Mountains, British Columbia, Canada. They develop a relatively low-cost workflow using satellite imagery and a land cover dataset to generate high resolution DSMs that are used as the foundation of the PRA model along with a validation dataset derived from collaboration with experienced guides in the study area.

The manuscript is generally well written and organized. However, there are a few areas where clarification is necessary (see Specific Comments). The workflow and methods are sufficiently technical and sound and build off other well-established methods of DEM/DSM generation and PRA modeling. The interpretation is well supported by the data and results.

Overall, this work is very interesting and novel, shows promise for potential widespread application in other regions, and is a worthy contribution to avalanche mapping and risk management in remote or data sparse regions. The authors obviously understand the implications of avalanche terrain mapping and chose to bias an overestimation of PRAs, which is appropriate for this attempt at mapping PRAs in forested areas.

I recommend publication with minor revisions and ask the authors to address my questions/comments and consider the suggestions below.

General Comments

3.1 Discussion of sources of error

As presented, the two major sources of error in the model are a function of uncertainty in forest characteristics and validation polygons. The authors provide a nice assessment of forest characteristics and uncertainty of validation polygons, but these are treated independently and separate. Is it possible to quantify the proportion of error due to each? I'm not necessarily suggesting additional analysis of this but think some comment based on available data might be useful and could help inform future work by suggesting potential directions to improve the model in forested terrain (e.g. should we focus on better validation data or input forest layer data?).

Author response:
This is an excellent suggestion.

Manuscript revisions: We added lines 607-608 to provide our insight into the relative proportion of the error sources in the interest or directing future research.

3.2 Extension to other areas

The validated polygons are a critical and interesting component of this study. The data are from a remote part of the Columbia Mountains and the authors are fortunate to have tenured guides provide input for this dataset. The authors suggest these tools could be used in other regions, but how can these tools be applied in other remote regions where operations are just getting started or expanding into new terrain but without the validation of long-term local knowledge or observations?

Author response:
This is a common theme from comments 1.1 and 2.7. We will expand the discussion of how our methods could be adapted to other areas in the updated manuscript. While our exact methods are specific to our collaboration with mechanized ski guides, we believe that the overall concept of using local expertise to create a validation dataset which accounts for limitations of human recollection is promising for other areas. For regions without local expertise available we suggest the development of default parameters for different snow and avalanche climates that could be developed in areas like Davos and Galena where the expertise does exist. The PRA

model can be generated in remote areas using these default input parameters and then fine tuned based on local input as expertise in that area develops.

Manuscript revisions: We added lines 694-708 to address this comment.

Specific comments

Author Note:
Many of the comments in this section were specific edits within the manuscript. We removed those comments and focused on those which had broader questions that warranted discussion.

Manuscript revisions: We addressed all the specific comments from reviewer 3 in the manuscript (lines 11, 76, 176, 251-258, 281-284, 313, 335-338, 359-360, 362-364, 372-373, 376-382, 393-394, 397-398, 595, 648).

3.3 Methods - Classification of small trees

Methods
Sec. 3.2.2: How did you account for forest that may be classified as forest (e.g. “trees with rigid trunks” (line 420)) based on spectral signatures, but where the canopy height is small and subsequently buried by snowpack, or where the forest density (from Sec. 3.3.1) may change depending on snowpack height throughout the season or inter annually? Would this (mis)classification contribute to the uncertainty in forest characteristics?

Author response:
This is an excellent point, and certainly a limitation of image classification based on spectral signatures. When developing our classification model we tested different sets of training polygons to create the desired result. Overall, smaller trees with rigid trunks that were not in close proximity to more mature forested areas were not included in the model because we intentionally selected a classifier that was biased towards only identifying forested pixels where there were stands of mature trees. Our classification is more likely to omit sparse and immature trees that stand apart from forested areas and are likely to be buried by the winter snowpack.

Manuscript revisions: We moved a section from the results to the methods and expanded the discussion in that section to address how we approached this issue (lines 251-258).

3.4 Method - Representativeness of validation polygons

Line 309-311: These runs were familiar to the guides, but were they representative of the variable nature of the terrain/PRAs throughout the study area?

Author response:
This is a similar comment to 1.3 and 2.7 and was answered in those responses.

Manuscript revisions: We expanded the discussion of how we selected validation runs slightly in lines 324-328, and tried to highlight that we relied on the local expertise of our collaborating guides to select the most representative validation runs.

3.5 Method - Nature of validation polygons

Line 314-316: It's not clear to me what you mean by "consistent character?" Did they simply draw polygons around a sub-sample of similar terrain from the five selected runs? If so, how does this help with validation of variable terrain/PRAs throughout the entire study area if they are of "consistent character"? Please explain.

Author response:
Our process of generating validation polygons involved each guide breaking the validation runs into polygons based on their knowledge of the terrain. The extent of each validation polygon was defined by the area the guides thought had similar terrain characteristics, such as slope angle, aspect, forest cover, and wind exposure. Then each polygon was assigned a percentage of the total area that could be a potential release area. This process of breaking the terrain into manageable segments in order to describe the PRA extent was necessary due to the uncertainty in outlining exact release areas based on digital maps and human recollection.

Manuscript revisions: We expanded our explanation of how the validation dataset was created to better explain what we mean with "consistent character" and how that relates to the results (lines 335-338).

3.6 Results/Discussion - Surface hoar

Line 562-563: Overestimation of PRA extent in all terrain (e.g. wind/sun affected alpine terrain) not conducive to SH development or do you mean forested terrain not conducive to SH development?

Author response:
Thank you for this clarification. The errors would be more likely in alpine areas with wind/sun exposure that are not conducive to surface hoar growth. In forested terrain SH is not likely to be a cause of overestimation because we increase the slope angle minimum using the forest slope scalar.

Manuscript revisions: We added a comment to the manuscript to clarify this point (lines 592-593).

3.7 Results/Discussion - Forest description

Lines 586-590: Are the forest descriptions here (e.g. "The forest is very dense") characterizations by the guides or based on the forest layer you derived? I assume it is guides' characterizations because you address canopy height which isn't a part of the forest layer. Additionally, how well does the mapped forest layer align with guides' characterization of forested terrain? Do you think forest density, independent of the forest slope scalar, has a greater influence than your results suggest?

Author response:
These forest descriptions are based on conversations with the guides, and we will make sure that is indicated in the updated manuscript. The guide's impression of the forest map is that it is very spatially accurate but does not always capture forest gaps in mature forests that can be large enough to produce avalanches. This is partially due to the horizontal extension of the forest canopy that can obscure the bare ground in areas with mature trees and tall canopy height. The combination of horizontal canopy extension and shading caused by large trees can cause misclassifications in these mature forested areas.

Our results actually indicate that forest density has a large impact on PRA accuracy and degree of underestimated terrain. Other than slope angle and maximum curvature, forest density is responsible for the greatest increase in accurate and underestimated polygons of all the terrain characteristics included in the PRA model (Figure 8).

Manuscript revisions: Lines 620 – 625 clarify that these comments on forest character came from the local guides. The impact of forest density on the overall accuracy of the PRA model can be seen in figure 9 f, where the improvement in underestimated polygons is apparent. We did not add any additional discussion to the manuscript about forest density based on this comment as we feel the existing results support that forest density added to the accuracy of the model.

3.8 Results/Discussion - Identification of PRA for wide range of conditions

Lines 657-660: This is a very important point that highlights the importance of human use/mitigation and snowpack structure. While you mention the influence of snowpack structure on PRAs (vs. just more frequent release areas) in several places throughout the manuscript, consider emphasizing in the abstract and the conclusions that the model identifies PRAs for all potential release areas under a broad range of snowpack conditions rather than potentially frequent release areas.

Author response:
This is an excellent point and gets at the fundamental difference of PRA modeling versus avalanche forecasting. PRA models identify all terrain that is capable of producing avalanches regardless of the current snowpack conditions and may not align perfectly with common snowpack structure for a given area.

Manuscript revisions: We appreciate this comment but did not make any changes to the manuscript. Our validation data collection methods focused on ‘frequent release areas’ in order to identify avalanches that humans are commonly involved in. We intentionally asked our collaborating guides to omit very rare release areas or areas that require special and uncommon conditions in order to produce avalanches. In that sense our validation dataset is somewhat dependent on snowpack conditions. We feel that the

existing manuscript carefully describes how we generated this validation dataset and then used it to fine tune the PRA model. Over emphasizing that PRA models are independent of snowpack conditions would contradict this part of our methods and therefore we chose to leave the manuscript as is.