NHESS-2024-158

Forecasting avalanche danger: human-made forecasts vs. fully automated model-driven predictions

By Frank Techel, Stephanie Mayer, Ross S. Purves, Günter Schmudlach, and Kurt Winkler

April 5, 2025

Dear Frank and co-authors:

Thank you very much for your patience since our processing of this manuscript has taken much longer than it should. I have now had the time to reread your manuscript in detail and study the reviewer comments as well as your responses to them. Based on my analysis, I recommend major revisions for your manuscript. While your study provides valuable and cutting-edge insights on the value fully automated, model-driven avalanche hazard predictions, I believe considerable changes are needed to make the manuscript clearer, more accessible, and more impactful for the NHESS readership.

In addition to the comments provided by the reviewers, I would like you to consider the following comments and suggestions.

1) Simplification of analysis

I encourage you to simplify the manuscript as much as possible and focus it on what is really needed to address the two main objectives of the study. Right now, it seems to me that you are presenting everything that you explored for this study, but not everything might be necessary for supporting your main conclusions. For example, it is really necessary to conduct the entire analysis on nowcasts and forecasts, which seems to make Fig. 5 much more complicated than it has to be? You establish very early on that nowcast and forecast predictions are highly correlated (this could be done in the methods section since it is not related to your primary research question). Once this is established, it does not seem necessary to continuously include both in the analysis. The trends shown in Fig. 5 are almost identical and presenting them both does not seem to add much value.

Similarly, it is unclear to me what the true value is of calculating the event ratios for human triggered avalanches in two different ways: a) with the reference distribution, and b) the sum of the observed events and non-events derived from the GPS points. Both approaches produce essentially the same results (aside a few differences which you attribute to limitations of the GPS dataset), but including them both requires considerable additional explanations, assumptions, and processing steps. To me, it seems that this additional complexity only makes the study harder to understand but does not add much additional value.

These two simplifications would have several benefits: a) require fewer datasets to be introduced, b) reduces complexity of Fig. 2, c) simplify the description of the analysis (Section 4.4), d) reduce the number of panels and data series in Fig. 4, e) simplify Table 1, and f) reduce redundancies in the results section.

2) Description of research objectives and structure of manuscript

The research objectives stated at the end of the introduction are "(1) Is the expected increase in the number of natural avalanches or in locations susceptible to human-triggering of avalanches predicted by spatially interpolated model predictions? and (2) Do fully data- and model-driven predictions achieve performances comparable to human-made avalanche forecasts?".

If I understand your manuscript correctly, you address Objective 1 by comparing the trends in the three model predictions (instability, danger level, and natural avalanches) to trends in event ratios of natural and human triggered avalanches. You then move on to address Objective 2 by comparing the model trends again the operational avalanche danger ratings in the avalanche forecast.

First, I find the wording of the second research objective as well as the title of the manuscript rather grandiose and too general. I think it would be better for both the title and the description of Objective 2 to be very explicit about your focus on the danger rating, which is only one aspect of human-made avalanche forecasts.

Second, as pointed out by one of the reviewers, the methods and result sections are difficult to follow. I wonder whether restructuring these two sections more clearly along the two research objectives would help make the information easier to understand and more accessible to the reader. Right now, the introduction of the different dataset and description of the analysis approach seems partially out of order and a bit detached from the research objectives. Hopefully, this would allow you to introduce the different aspects of the analysis in a more logical order and not having to refer to sections across the manuscript (e.g., L137, L209). Furthermore, I hope that my earlier suggestion can help to further simplify the text.

2) Simplification of danger rating model output

I understand that your simplification of the danger rating model output produces a single binary variable with probabilities from 0 to 1 similar to the output of the other models, which makes your subsequent analysis easier since you can use the same approach for all models. I have not thought this all the way through, but you could do this for multiple danger rating thresholds. I am not sure whether this would provide any additional insights, but regardless, I think the methods section would benefit from a more in-depth explanation of your choice.

3) Generalization

Drawing universal conclusions about the ability of avalanche forecasting model chains based on a dataset from two seasons that includes considerable assumptions and simplifications without providing any measure of uncertainty is a bit challenging. I understand that the additional analysis by Florian Herla might be out of scope for this paper, but I encourage you to either a) include something along these lines to provide the reader with some sense of uncertainty in your analyses, or b) adjust the wording in your discussion and conclusion to make it clear that the results are just a description of the available dataset and generalization should only done with caution.

4) Discussion

The discussion seems a bit meandering and go beyond the main objectives of the study. Overall, I would appreciate a stronger focus on the discussion of the two research questions at the beginning

of the discussion before relating the results to other studies and avalanche forecasting in general. I encourage you to ground your discussion of the state of avalanche forecasting more strongly in your results and make those links more obvious throughout the text.

While I appreciate you being up front about the limitations of the study right at the beginning of the discussion section, the section primarily repeats information that was presented in the methods section already.

I would also like to encourage you not to include new analyses in discussion section (Section 6.4). In my opinion, it would be better to discuss this earlier in your manuscript since it relates to design choices you make for your study.

Other minor comments

p. 3+ - Section 2: The subheadings in Section 2 seem unnecessary since they each just consist of a single paragraph.

p. 4 – Fig. 1: Is it necessary to label all AWS with their elevation? It completely clutters the map, and many labels overlap and are therefore hard to read.

p. 5 – Fig. 2: I really appreciate these types of flow charts that explain the design of a study. However, I find this one difficult to understand. Part of the reason might be the complexity of the study (see earlier comment) and the fact that the figure is introduced at the beginning of Section 3 and never referred again later in the text. It might be helpful to talk the reader through the figure more explicitly.

p. 6 – Section 3.2: Please use a consistent spelling for snowline.

p. 6+ - Section 3.3.1: The summary statistics presented for the reference distribution, events and non-events project a level of precision that is beyond the accuracy of the data, These number could probably be rounded to the closest number of 10 m.

p. 8+ - Footnotes: It seems like your footnotes could be replaced with regular citations.

p. 10 – Calculation of relative event ratio: Please explain why calculating the relative event ratios is beneficial for the presentation.

p. 11 – L274: Could you please describe how you applied the chi-squared test in more detail. It is not clear to me what this adds to your analysis.

p. 12 – Fig. 4: Could the panels in this figure be arranged that they are more in line with the analysis approach. I believe this is the case for the first two columns (reference distribution -> events -> event ratio), but it is different for the final column where the top chart refers to non-events whereas the other two charts relate to events (human-triggered avalanches). See my earlier comment about simplifying the analysis in general. Maybe Figures 4 & 5 could even be combined.

p. 12 – Table 1: It is also difficult for me to completely understand the link between Fig. 4 and the numbers presented in Table 1.