Response to review by Karl Birkeland

Dear Karl,

Thank you for reviewing our manuscript, and providing detailed, constructive feedback. Please find below our response to the points raised and the suggestions made. As most of the comments are technical corrections, we will provide a point-by-point response to these points when submitting the revised manuscript. We intend to address all these recommendations when revising the manuscript.

L2: I'm not sure I would say that an avalanche forecast "terminates" with assigning the avalanche danger. This is because after assigning the avalanche danger you still have the issue of how to communicate all the nuances to the public in the written avalanche forecast.

We agree. The forecast process is followed by compiling the product for communication. But the assessment process terminates with assigning a danger level. – We will rephrase accordingly.

L4: I think this is true for some avalanche forecasts where things are simplified quite a bit. But the nuances of the conditions can be captured in other ways than the danger rating (forecast discussion, videos, etc.).

We fully agree that there are ways to communicate nuances in the conditions. However, at the level of this final avalanche danger assessment, the danger level is a strong simplification of the continuous nature of avalanche conditions.

L13: I always treat "data" as plural. So, this would be "data are" rather than "data is". I just looked it up and I believe that used to be pretty standard, but it sounds like some people are now using "data" as singular as well. I'm not sure what NHESS prefers, but if they prefer it to be plural you may need to change this throughout the paper.

Thank you for pointing this out. - We will revise throughout the manuscript.

L31: “for each identified avalanche problem type”. Note: I think it is important to clarify that this assessment is done for all avalanche problem types, as per the CMAH.

We agree.

L48: Here I would suggest adding a sentence saying that forecasters attempt to bridge this gap between a continuous phenomena and a discrete number using the text of their avalanche forecasts. Then, instead of starting the next sentence with "In addition", I'd substitute in "Regardless"

Thank you for this suggestion. We will take that up when revising the manuscript.

L 76: “are simply asking whether or not there is a”

L81: “the correlation between the range of”

L119: This is interesting. Will you discuss why you think this is the case sometime later in the paper?

As Rune Engeset (Reviewer #1) also wondered why this was the case, here some more details: After the first season of using the sub-levels, the forecasters discussed issues and challenges relating to their use. In general, forecasters were comfortably assigning a sub-level to a danger level for dry-snow conditions. This was different for wet-snow conditions: forecasters had essentially always assigned sub-level + for wet-snow conditions. We interpreted this as forecasters being unable to make the relative comparison within the danger level for wet-snow conditions. At the time, we (the forecasters) did not analyze in detail as to the why this was the case but we suspect that forecasters at SLF have more difficulty assigning a danger level in a wet-snow situation compared to dry-snow conditions. As the absolute judgement precedes the sub-level assignment, this means that reliably assigning a sub-level is essentially impossible.

L129-130: You said earlier you did not consider wet snow days. So, at this step did you exclude days that had only wet snow avalanche problem types?

This is correct. As forecasters assigned sub-levels only to danger levels describing dry-snow conditions, we excluded all forecasts, which had only a wet-snow or gliding avalanche problem as main avalanche problem (L119-121, L130). – We will clarify that we excluded these forecasts.

L197: This is a unique and interesting dataset, and it's awesome you are using it. I wish such data were available for backcountry ski tourers in the US.

Section 4: The methods look solid to me.

L343-344: This is really interesting, and it shows that the Swiss forecasts are doing a good job of identifying days when avalanches are triggered by humans. I've always wanted to see numbers like this, so it's exciting to see this result.

L355-364: More interesting results.

L 390-403: These are further interesting results. Some of the small bumps in the ECT curve are a bit puzzling. The RB curve seems to do a little better in terms of having fewer "bumps" and a more consistent increase.

The results shown in this section and Figure 8 confirm previous findings that RB test results have a stronger correlation with instability compared to ECT results (e.g. Techel et al., 2020).

L 405-422: Nice work utilizing these models to further strengthen your results.

Table 4: I like this summary graphic.

L 435: “For the first research question,”

L 443: “second»

L 425-457: Nice section!

L 464: «Forecasters»

L 458-483: In this "Implications for forecasters" section, I think it would be useful to discuss how the forecasters viewed this additional task when putting together the forecast. Was this difficult for the forecasters, or did it not really take much additional time?

When describing dry-snow conditions, forecasters felt comfortable assigning a sub-level to a danger level. Based on our own experience (authors/forecasters Frank Techel, Kurt Winkler) and talking to our forecaster colleagues, we can say that it takes very little additional time, once a decision concerning the danger level has been made. However, we believe that the step of assigning the absolute danger rating requires more cognitive effort than the relative assessment made afterwards. During the forecaster briefing, when a consensus on the forecast is sought, the sub-levels are discussed as well. From our experience, this does not take a lot longer than discussing any of the other elements, which are communicated in the forecast. Please note that this is our subjective take on this. – When revising the manuscript, we will mention that the effort and time required to assess the sub-level is small.

L485: Based on your assessments that you presented in this paper, I agree with you.
However, in terms of the "Practical applications", do you see any potential drawbacks of providing this information to the public? You touch on this a little bit below when you discuss the "comprehensibility" of the forecast products.
One thing that comes to mind for me is that already some members of the public - and even some professionals - utilize the avalanche forecasts as a "go/no-go" tool. Of course, this is completely the wrong way to apply the forecast since it is, by definition, a regional product and ultimately the user has to make a slope-scale decision. But I worry that some members of the public as well as some professionals might use the "+" and "-" system to justify pushing a little bit closer to the edge. In other words, maybe before they wouldn't ski things when the danger rating was considerable. Now they might ski it when the danger is "Considerable -" or even "Considerable =", but not when it is "Considerable +". I don't know if this is really a worry for you in this scientific paper, though you may want to consider discussing it since I believe this paper will attract wide readership among avalanche forecasters and so your words here might carry some weight.

We agree that it is challenging to present this additional information to forecast users in an appropriate manner. From our perspective, it is clear that the bulletin structure has to remain in line with the information pyramid. We consider the avalanche danger level, aspects and elevations where the danger prevails, and the avalanche problems as more important than the sub-level. Concerning the communication of the sub-levels, it is intended to display these only on the fourth level of the information pyramid, together with the danger description. From the danger description, skilled forecast users can extract nuances describing the avalanche conditions. With the sub-levels, these nuances will be summarized in a more systematic and explicit way. As a side-note: for our own information, we did also analyze the danger descriptions as a function of the danger level – sub-level combination, using the data and methodology described in Hutter et al. (2021). These results are not published, but they showed that the danger descriptions (likely trigger, number of potential triggering locations/avalanches, avalanche size) also correlate with the sub-levels. – As we have discussed potential challenges arising with the communication of the sub-levels already (L488-495), and as the goal of the study is to explore the consistency in the assignment of the sub-levels, we will not go into more detail in our manuscript.

Frank Techel, on behalf of all co-authors

References

Hutter, V., Techel, F., and Purves, R. S.: How is avalanche danger described in textual descriptions in avalanche forecasts in Switzerland? Consistency between forecasters and avalanche danger, Nat. Hazards Earth Syst. Sci., 21, 3879–3897, https://doi.org/10.5194/nhess-21-3879-2021, 2021.

Techel, F.; Winkler, K.; Walcher, M.; van Herwijnen, A.; Schweizer, J. (2020). On snow stability interpretation of extended column test results. Natural Hazards Earth System Sciences, 20, 1941-1953, 10.5194/nhess-2020-50