Reply to Reviewer #1

We thank the reviewer, Dr. Engeset, for his insightful comments. Below we provide answers to the major points and indicate how we will improve the manuscript.

General comments

The paper is well written, has a clear scope and analysis, and presents new results of scientific and practical value. It is an important contribution towards improving avalanche danger assessments and forecasting. It helps scientists and practitioners to assess and define avalanche danger, with practical implications ranging from developing machine learning tools to crowd-sourcing and interpreting signs and observations in the field. It also sheds light on the timing and combination of avalanche problems (new snow and persistent weak layer) based on observations, and provides directions for further research. The language, figures and tables are of high quality and the flow of the manuscript is sound and logical.

I recommend the editors to publish the paper after minor revisions.

Specific comments

Here are specific comments, which could be used to improve the manuscript:

- *Not only stability, but also distribution/frequency of instability and the avalanche size affect the LN data and thus the analysis. The effect of not accounting for this on the results could be described further in the discussion and mentioned in the conclusions, if possible.*

We agree and will include these points into the Discussion section of the revised manuscript.

- *The data comes from a limited geographical area, and the study would ideally include data from other snow climates and regions of the world. I guess such data are difficult to come by, but it could be a recommendation to include more data in future studies (e.g. explain to other scientist how to collect and share data in a common data set for an analysis of more a regional or even continental/global perspective). It would also be interesting to know which other observations theauthors would recommend for future studies, to improve the observational basis for quantifying avalanche danger (properties of the weak layer or the slab?)*

We agree that our data comes from a limited geographical area. As described in the Discussion section there is considerable agreement with regard to snowpack properties with previous studies, for instance, by Schweizer and Jamieson (2003). Their study included profiles from the Columbia Mountains of western Canada and Switzerland. This agreement suggests that characteristics of instability in a snow profile may not depend so much on snow climate – on the other hand, of course, their frequency of occurrence will depend on climate.

With regard to data collection, we think the issue is not so much about the type of data to be collected, but consistency. Hence most important are proper observations standards and observer training. Of course, the EAWS may play role here in establishing standards such as the SWAG published by the American Avalanche Association.

- *The introduction could explain how this study adds to – and differs from – the* *studies by Techelet al. (2020) and Schweizer et al. (2020)*

The present study includes the profile characteristics as well as the relation of the danger level to signs of instability. It even aims at describing the danger level based on observations, which is not possible based on the study by Techel et al. (2020) where the frequency distribution of instabilities needs to be known. The study by Schweizer et al. (2020) solely focussed on avalanche size.

- *In chapter 3.1.2, could you include statics for quality of fracture plane (smooth, rough, irregular)?*

The quality of fracture plane is known to be less indicative of stability than, for instance, the RB release type (e.g., Schweizer et al., 2008). Overall, 64 % of the tests showed a smooth fracture plane, 21 % a rough one, and 15 % of the failures were irregular. For rutschblock score 1 to 3, the proportion of smooth fracture planes was slightly higher than average, and vice versa for RB scores 4 to 6 (Figure R1). On the other hand, the proportion of irregular fracture planes clearly increased with increasing RB score. In conclusion, the observation of an irregular fracture plane indicates rather favourable conditions, whereas any other type of fracture plane does not allow any firm conclusions on stability.
We will add this information in the revised manuscript.

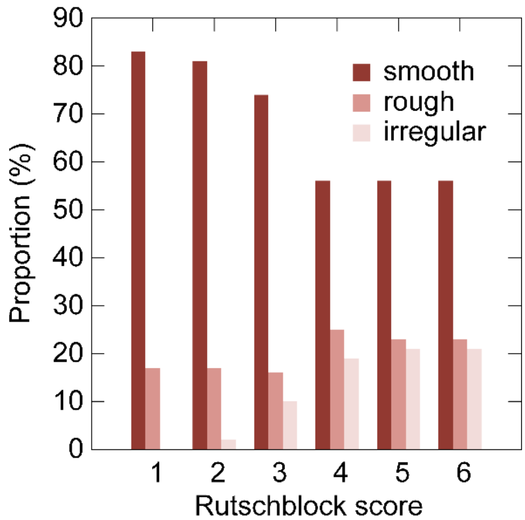

Figure R1: Proportions of the three fracture plane types per rutschblock score (*N* = 581).

- *I would suggest reversing the x-axis in the figure in the appendix (have very poor to the left and good to the right) to align with the other figures in the manuscript, and to insert this figure (as well as its results) in the results chapter. It could also be beneficial to add a paragraph discussing,comparing or summarising the results of the three different scales used for stability (categorising stability into 3, 4 and 5 classes)*

We will change the x-axis as suggested and consider its inclusion in the main text. Our original idea was to provide the figure in the Appendix primarily for illustration and to facilitate the comparison to the previous study by Techel et al. (2020). Including it in the main text has the potential for confusion and would require a more thorough introduction and more background information on this type of stability classification.

Technical corrections

Now follow specific comments, with reference to line numbers in the manuscript:

- *#30-37 This is a description of a regional forecast, I suggest you add the word "regional" somewhere, and briefly describes how regional warnings are different from other forecasts (object-based / slope-scale) and, if possible, if this study may be valuable also to slope-scale assessments/forecasts.*

We will explicitly refer to the regional scale in the revised manuscript. We implicitly do so in line 35. Stability evaluation is valuable for slope-scale assessments, but as pointed out in line 36 the slope scale or object-based assessments are different from regional forecasts.

- *#32 I could not easily find the EAWS 2019 in the reference list (it is on #570)*

Thanks for pointing this out. We will remove this inconsistency in the revised manuscript.

- *#39 Could you improve the English of the sentence starting with "Even"?*

We will revise the sentence.

- *#59-60 Consider rewrite and simplify the first sentence, e.g. "but it was not possible todistinguish 2–Moderate from 1–Low"*

We will revise this sentence.

- *#84 Explain why you selected dry snow conditions only (dry slab selection could also beexplained in #113)*

Stability evaluation for dry-snow conditions is of particular relevance to mitigate the persistent weak layer avalanche problem type. Profile data are primarily indicative of dry-snow conditions. Also, signs of instability such as whumpfs and shooting cracks relate to dry-snow slab avalanches. We will provide somewhat more rationale in the revised manuscript.

- *#104 Could you improve the definition of the "adjacent layer" and explain when it is above andbelow the failure interface?*

We follow the approach that was introduced by Schweizer and Jamieson (2003). In all profiles, the failure interface was reported. While in most cases the weak layer is obvious, we considered the softer of the two layers as the failure layer (FL) and the layer across the failure interface as the adjacent layer. If there was no difference in hardness, we selected the lower layer as the failure layer, and the layer above the failure interface as the adjacent layer.

- *#108 Explain why you derived three stability classes her, while using five classes above (#95) and even 4 classes in the appendix*

Traditionally, five classes were used as described by Schweizer and Wiesinger (2001). In various subsequent studies simplifications were sought, often to ease interpretation or to make it amenable for automatization. As with any classification the choice of classes is partly arbitrary. We will provide some more details in the revised manuscript.

- *#171-174 Please simplify this description*

We will revise this description.

- *#242 Please explain the acronym "RF" (presumably regional forecast)*

We will introduce RF in the Data and Methods section of the revised manuscript.

- *#245 Please clarify: Did the observer estimate the LN for the same region / area as the RF or is LN valid for a smaller area than RF and thus probably biased towards lower danger levels (presuming that RF is the highest danger level in the region)?*

Many thanks for pointing out this scale issue, which we will discuss it in the revised manuscript. We also provide a figure in the Reply to Reviewer #2 that illustrates this issue (Figure R2).

- *#277 I presume "even" could be deleted (or is the sentence incomplete?)*

We will revise this sentence.

- *#367 Could you improve the sentence "the split variables and values were plausible"?*

We will improve the sentence.

- *#436 Could you improve the sentence "only in 1 out 5 cases differences in snow depth were indicative of snow instability"?*

We will clarify this sentence and provide more detail on the study by Reuter et al. (2015).

- *#523 Please explain how/why this poses a challenge to forecasting?*

Persistent weak layers can linger for a long time in the snowpack and represent the failure layer in skier-triggered avalanches. However, it is challenging to communicate a danger (or the avalanche problem type: persistent weak layer) that is “invisible” and often only sporadically present.

References

Reuter, B., van Herwijnen, A., Veitinger, J., and Schweizer, J.: Relating simple drivers to snow instability, Cold Reg. Sci. Technol., 120, 168-178, https://doi.org/10.1016/j.coldregions.2015.06.016, 2015.

Schweizer, J., and Wiesinger, T.: Snow profile interpretation for stability evaluation, Cold Reg. Sci. Technol., 33, 179-188, https://doi.org/10.1016/S0165-232X(01)00036-2, 2001.

Schweizer, J., and Jamieson, J. B.: Snowpack properties for snow profile analysis, Cold Reg. Sci. Technol., 37, 233-241, https://doi.org/10.1016/S0165-232X(03)00067-3, 2003.

Schweizer, J., McCammon, I., and Jamieson, J. B.: Snowpack observations and fracture concepts for skier-triggering of dry-snow slab avalanches, Cold Reg. Sci. Technol., 51, 112-121, https://doi.org/10.1016/j.coldregions.2007.04.019, 2008.

Schweizer, J., Mitterer, C., Techel, F., Stoffel, A., and Reuter, B.: On the relation between avalanche occurrence and avalanche danger level, Cryosphere, 14, 737-750, https://doi.org/10.5194/tc-14-737-2020, 2020.

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, Cryosphere, 14, 3503-3521, https://doi.org/10.5194/tc-14-3503-2020, 2020.