Reply to Reviewer #1

We thank the reviewer for the positive feedback and his constructive suggestions.

The paper is of high scientific quality. It brings forward important quantitative results on the occurrence of snow avalanches in the Eastern Alps, and provides an important analysis and discussion of the implications for assessment and forecasting of avalanche danger. The study uses a unique and rich data set on observed avalanches and regional avalanche danger estimates, in order to reduce a critical knowledge gap that has limited the development of objective procedures for determining the avalanche danger level. The publication of the paper will contribute to improved workflows, standards and eventually better avalanche forecasting products in the future.
My recommendation to the editors is to publish the paper, after addressing the points below (minor revisions).
The language, figures and tables are generally of high quality and easy to read. The structure is easy to follow, and the balance between data, results and discussion is well suited for a publication. However, I recommend improving readability by splitting many long complex sentences into shorter sentences.

Thanks for the suggestion; we followed this advice while revising the manuscript.

One aspect I was missing was an analysis and/or discussion of the general transferability of the results to other parts of the world, especially where terrain or climate conditions differ from the Davos region. One could assume that not only the total size of the study area matters, but rather the size of avalanche terrain area. Some terrain is not able to produce larger avalanches, while other types of terrain produce many large avalanches. Some climates produce many large natural avalanches, while others less so. If the authors could add these aspects to the discussion, it may be easier for the readers or future studies to generalise, add to, or test the results.

We added the proportion of avalanche terrain (67%) based on the classification by Harvey et al. (2018). Moreover, the snow climate can be described as transitional (McClung and Schaerer, 2006). As you indicate terrain and snow climate influence avalanche activity. However, we deem it speculative to predict the characteristics of avalanche activity in regions with other snow climates or terrain characteristics. More similar studies are needed. We do refer to some studies from North America and the New Zealand in the Discussion section. We suppose the avalanche size distribution to be fairly independent of terrain and climate characteristics. However, the frequency and intensity of occurrence will certainly vary.
We now mention this issue in the Discussion section (lines 374-377).

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

#7 Add size of study area, number of avalanche observations and number of regional danger assessments.

We added these numbers to the Abstract (lines 7-8).

#13 Could the sentence ending in "…given day" be improved by adding "an danger level" at the end?

Changed as suggested (line 13).

#15 Could the sentence be improved by replacing "may allow revisiting" by "suggest reworking of"?

Changed as suggested (line 16).

Reply to Reviewer #2

We thank the reviewer for the positive feedback, insightful comments and constructive suggestions.

The present manuscript describes an analysis that takes advantage of a large existing dataset of observed and mapped avalanches to explore the relationship between avalanche danger levels and avalanche occurrence in the Davos region of Switzerland. The research is of high quality, and the results contribute valuable insights to the current discussion on avalanche forecasting practices and consistency. The various analyses described in the manuscript offers useful information on the role avalanche size in avalanche forecasting, and the recommendation on the number of expected avalanches at danger rating level High has the potential to be the starting point for making avalanche forecasting more objective by replacing the existing qualitative descriptors in the danger scale with more objective quantitative measures.
The manuscript fits well with the mandate of The Cryosphere journal, and it will offer great value for avalanche safety researchers and practitioners. However, despite the obvious strength of the research, I believe that the manuscript has a few weaknesses that should be addressed before the manuscript is published. My concerns mainly relate to the presentation of the dataset correction procedures and the qualitative description of the results. I hope that my comment below are useful for making the manuscript even more impactful.

GENERAL COMMENTS AND SUGGESTIONS

Correction procedure (Lines 109-140)

Given that the objective of your paper is to examine the relationship between avalanche danger ratings and avalanche activity, manually changing danger ratings based on observed avalanche activity prior to analysis seems risky. While I do not necessarily disagree with the approach, I have the following recommendations for making it more transparent for the reader:

To put the number of days with corrected the danger ratings into perspective, it would be useful to provide readers with counts and proportions of danger ratings prior to correction right at the beginning of this section. This information is currently only available for the corrected danger ratings (Table 2). Having this information upfront would help readers to understand how much of the dataset was modified.

We agree and added the following Table as supplement to the revised manuscript.

Table A: Frequency of danger levels before and after corrections. Also given are the changes per danger level.

Danger level	Number of days before corrections	Change of danger level					Number of days after corrections
		-2	-1	0	+1	+2	
1–Low	306	-	-	303	1	2	303
2–Moderate	1809	-	0	1765	32	12	1766
3–Considerable	1367	0	0	1310	57		1366
4–High	47	0	21	24	2	-	94
5–Very High	4	1	1	2	-	-	4

The description of the correction procedure refers several times to the fact that avalanche activity was 'unusually high' or 'unusually low'. However, you do never explicitly specify what your expectations regarding avalanche activity actually are and how you determined your thresholds (one exception is the recoding of moderate days with AAI > 1.0). Being more explicit about your criteria would make your procedure more transparent.

We agree and regret the confusion. The procedure and the criteria are as follows:

1. We evaluated all days with danger levels *4–High* or *5–Very High* and a value of the AAI ≤ 1 (“zero or unusually low”).
2. We evaluated all days with danger level *3–Considerable* and a value of the AAI > 13.6 (after the first correction step), which was the median AAI for the days with danger level *4–High* or *5–Very High* (line 120).
3. We evaluated all days with danger level *2*–Moderate and a value of the AAI > 1, which was the median AAI for the days with danger level *3–Considerable.*
4. We evaluated all days with danger level *1–Low* and value of the AAI > 1.

I personally found the description of the correction procedure somewhat difficult to follow due to many details described in the text. I wonder whether a diagram (e.g., flow chart) showing which danger ratings were changed to what and for what reason would help the reader to better understand the magnitude of your changes and their potential impact on the subsequent analysis.

Thanks for the suggestion. We hope that we can address your concern with Table A shown above and the correction procedure described above, which we now both provide in the Supplement.

The numbers in your description of the changes applied to danger with High and Very High danger ratings (Lines 111-116) do not seem to add up properly.

Thanks for checking. We checked and the numbers should now be consistent with Table A in the Supplement.

Overall, you changed the danger rating in 122 of 3533 days, which only amounts to 3.5%. This seems like a rather small amount and my initially thought was that the correction procedure was unnecessarily complicated given that it will likely only have a minor impact on the analysis. However, it represents 9% of the days with avalanche activity, and, if I understand your descriptions correct, the number of days with danger ratings High and Very High changed substantially through the correction procedure. The number of days with a High danger rating was first reduced from 44 to 26 (-18) (Line 116) and then increased again to 94 (+68) (Line 141). This means that only 28% of the days with High danger ratings in the analysis dataset were originally assigned a High danger rating. Given that the High danger rating sample plays an important role in the subsequent analysis, I believe that the impact of the correction procedure on the nature of the dataset should be described more clearly.

The main effect of the correction procedure is that the median AAI for days with *4–High* increased from about 10 to about 21. For days with *3–Considerable*, the median AAI was 1 and did not change due to the corrections. Hence, the difference in avalanche activity between *3–Considerable* and *4–High* was already very prominent before the correction procedure (before and after the corrections: *U*-test, $p < 0.001$). With regard to avalanche size, the effects are less prominent. Size 2 avalanches were the most frequent ones at danger levels *1–Low* to *4–High* before and after the corrections.
We added a short paragraph on the impact of the correction procedure in the Discussion section (lines 398-404).

A brief discussion of the potential effect of the correction procedure on the analysis results in the discussion section would further acknowledge its impact. I think it is important to explicitly mention that there is potential for a bit of a circular argument here: You corrected the danger rating levels based on avalanche activity expectations to later analyze exactly this relationship.

We agree, see above, and added a short paragraph in the Discussion section along the lines described above (lines 398-404).

Description of analysis methods

The section titled 'Data and methods' only includes descriptions of the derivation of the avalanche size, the danger rating dataset and the quality control and correction procedure but seems to completely skip a description of the actual analysis approach and statistical methods employed. This seems rather unusual. I believe the manuscript would benefit a short overview of the analysis approach that describes the measures used (e.g., avalanche activity index, proportion of days with avalanches, etc.) and how they relate to the components of avalanche hazard (e.g., snow stability, frequency of locations) in the methods section.

We now describe the analysis methods in more detail (lines 163-173).

Statistical support for qualitative descriptions

Much of the description of the observed patterns are rather qualitative with some statistical tests here and there. I am wondering whether some of the statement could be supported with statistical test statistics. I believe that this would considerable strengthen the power of the manuscript.

We agree and are fully aware that we did not provide many statistical test results. We did so since we believe that the analysis is simple and the data actually speak for themselves.
We now provide more statistical test results to better support some of the main findings (e.g. lines 178, 231, 237).

Description of study area

On line 377, you provide recommendations about the number of expected avalanches at danger rating level 4-High (at least 10 per 100 km2), and on line 416, you suggest that the term "many avalanches" should mean on the order of at least about 10 avalanche per 100 km2. I believe that this is an interesting result. However, while you highlight that avalanche occurrence probability depends on scale as it is a combination of stability, its distribution within the forecast area and the size of the forecast area, it seems to me that the nature of the terrain in the forecast region would also have a substantial impact on the suggested number. I therefore wonder whether a more detailed description of the nature of the avalanche terrain in the study area (e.g., number of avalanche paths of different size, total extent of avalanche terrain) would offer valuable context for understanding the results and recommendations.

We agree that the type of terrain certainly affects avalanche activity. We now provide the proportion of avalanche terrain (lines 79-81).

Insight into avalanche warning practices

In several places in this manuscript, you comment on the somewhat unexpected differences in the observed numbers of dry and wet avalanches at the same danger rating level (e.g., Lines 323-324, Lines 375-376). However, there is no explicit statement in the discussion or conclusion section that points out that these observations might indicate inconsistencies in forecasting practices. I think that a statement like this would fit nicely with the recent literature on avalanche forecasting inconsistency and further contribute to this research.

Thanks for that suggestion; we do actually mention it in the Conclusions (line 500). We also added this point to the Discussion section (lines 395-397).

LINE-SPECIFIC COMMENTS

*1. **Line 95 – Figure 1** The exponential increase presented in Figure 1 seems to be the direct consequence of the classification criteria presented in Table 1. I wonder whether plotting the log of avalanche area versus avalanche size class would be more useful to highlight that the approach* *classifies avalanche in the spirit of the Canadian size classification. The number of avalanches per class shown in the chart do not add up to the total number of avalanches given in the caption.*

We have deliberately chosen the length, as this measure is the one most practitioners in Europe can well relate to. Also, the length is included in the EAWS definition of avalanche sizes. In addition, we provide the median area per size class in Table 1.
Thanks for checking the numbers. There is indeed an error since for some of the very small avalanches a meaningful value of length could not be derived so that the total number reduces to 13,802.
We now mention this in lines 89-91. Moreover, we now provide the number of cases in Table 1.

2. ***Line 100*** *It might be useful to explicitly state that the weight of 0.81 is appropriate because it is highly likely that the avalanches without known triggers were likely natural avalanches.*

We now explicitly state that the value is appropriate (line 113).

3. ***Line 191 – Table 3*** *It might be useful to add row percentages to the columns to better highlight the relationship between avalanche size distribution and danger rating.*

Thanks for the suggestion. We added the proportions in Table 2.

4. ***Line 192*** *The Kruskal-Wallis test only indicates whether there are any differences in the avalanche size distributions among all danger rating levels. You could follow-up with pairwise Wilcoxon rank-sum tests between adjacent danger rating levels to determine where exactly the differences are.*

Thanks for the suggestion. We now describe some of the pairwise comparisons (lines 229-232) and also provide all *p*-values in Table B of the Supplement.

5. ***Line 196 – Figure 3*** *I think it would be best if the proportion scales in all charts would range from 0 to 1 and be styled the same.*

We agree and changed the scale on the y-axis in Figure 3a.

6. ***Line 196 – Figure 3c*** *If I understand your analysis correctly, Fig. 3c depicts average or median number of avalanches of different sizes per day at different avalanche danger levels. A bar chart does not seem an appropriate way to display this information as bar charts are typically used to depict proportions. Populating the same layout grid with a series of box plots instead of vertical bars would represent not only the magnitude difference in number of avalanches between danger ratings and avalanche size, but also the range within the observations.*

Thanks for your suggestion. We prepared the box plots for the number of avalanches, but realize it would not be possible to provide all data in one graph as you suggest. Moreover, the information would be hard to depict since the scale on the y-axis would range from 0 to 250. Therefore, we prefer to keep Figure 3c as presented.

7. ***Line 209*** *Can the statement "The median as well as the 90-percentile avalanche length did not increase for the danger levels 1-Low to 4-High." be substantiated with a statistical test result?*

We now provide more details on the median avalanche length (lines 247-250). Also see Figure A below, which we included into the Supplement.

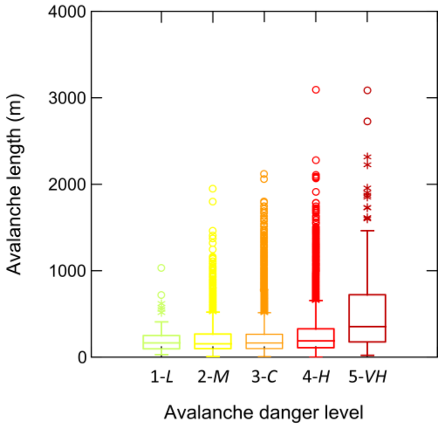

Figure A: Avalanche length per avalanche danger level (N=13'745).

*8. **Line 217 and Figure 4a** It seems odd to combine avalanches with unknown snow conditions with the mixed category in the snow conditions analysis (Section 3.3) as these are very different categories. I think it would be better to leave these avalanches out of the analysis all together.*

Based on our experience the unknown conditions are often related to mixed conditions. The size distribution for unknown and mixed were statistically not different (*U*-test, $p = 0.82$).
We added this information (lines 257-258).

*9. **Line 227 and Figure 4d** The difference in the number of avalanches per day between dry and wet avalanches under avalanche danger level 1-Low shown in Fig. 4d seems minute. Can this statement be supported with a statistical test?*

As stated in lines 226-229 there were distinct differences between days with wet-snow avalanches and days with dry-snow avalanche on days when the danger level was 1–*Low* – though the number of cases is rather small:
On the 10 days with natural wet-snow avalanches, the number of avalanches was 18 and the total AAI was 15.2. Whereas on the 6 days with natural dry-snow avalanches, the number of avalanches was 8 and the total AAI was 0.71. As stated, the AAI was more than ten times larger.
On the other hand, in Figure 4d the average number of avalanches per day is shown, which is 1.3 for dry-snow and 1.8 for wet-snow conditions; this difference is not particularly large as you point out and statistically not significant.
We now provide more details and rephrased that paragraph (lines 274-279).

*10. **Line 233** I assume that this discussion should be referring to the AVERAGE or MEDIAN number of avalanches per day.*

We added the average (line 284).

*11. **Line 241 – Figure 4d** Same comment as for Figure 3c*

Please see our response to point #6 above. At 4–*High* the maximum number of avalanches per day was 255 so that the y-scale would need to range from 0 to 300 and not much could be seen at the lower danger levels.

*12. **Line 253** The statement "… size 4 avalanches were five times more frequent among the natural than the human-triggered avalanche." does not seem to be supported by Figure 5a.*

We agree that our statement was not correct since it refers to the absolute number of size 4 avalanches. Since there are three times more natural than human-triggered avalanches, the proportions are 2.9 % vs. 1.6 %, still a significant difference (proportion test, $p = 0.02$). We corrected the statement (lines 306-307).

*13. **Line 256 – Figure 5c** Same comment as for Figure 3c and Figure 4d*

Please see our response to point #6 above.

*14. **Lines 268-273** Same comment as earlier regarding the average/median number of avalanches per day.*

We added the average (line 329).

*15. **Line 293** I think it would be useful for the reader if the fact that no temporal trends in the avalanche size distribution were detected in the analysis dataset was included and substantiated in the initial description of the dataset.*

We now mention that fewer avalanches were recorded in the first six years of the study period in the Data and Methods section (lines 84-85). Also, we added some more details in the Discussion section (lines 355-359).

*16. **Lines 316-324** It seems to me that some of the explanations of the data correction procedure described in this paragraph should be included in the methods section.*

We now provide more details in the Methods section on the correction procedure (see above) and also added a paragraph to discuss possible effects of the corrections (lines 398-404).

*17. **Lines 340-346** In this section, you refer to the avalanche size distribution of human triggered avalanches, and the topic comes up again on Line 407. However, I did not find this explicit analysis in your manuscript. If I read your manuscript correctly, you only analyzed the number of human triggered avalanches under different danger ratings but not their size distributions. It seems to me that such an analysis would nicely complement your existing analyses.*

Harvey (2002) considered avalanches that caused damage to either people, infrastructure or forest, and reported avalanche size. Hence, we discuss his study to compare his results to our findings.
In fact, we have analysed the size distribution for human-triggered avalanches at the different danger levels. The distributions are very similar to the overall distribution shown in Figure 5a. For instance, at danger level *3–Considerable* and *4–High*, the proportion of size 3 avalanches is 0.74 and 0.73 (proportion test, $p = 0.77$). For size 4 avalanches the corresponding proportions are 0.14 and 0.19 (proportion test, $p = 0.36$). Overall avalanche size did not increase with increasing danger level.
We added this information to section 3.4 of the Results (lines 315-320).

*18. **Lines 355** Same comment as earlier regarding the average/median number of avalanches per day.*

Thanks, we now specify that the number is the average (line 437).

*19. **Lines 378-381** The description of the results of Bründl et al (2019) is a bit confusing to me. What are the five frequency classes and how do they relate to the results presented in this paper?*

We regret the confusion. In fact, the frequency classes by Bründl et al. (2019) are not relevant in our context. They counted the number of size 4 avalanches per 250 km^2 and provided the results in five classes of varying avalanche density. In the lowest class they reported <29 avalanches, whereas the number of avalanche per 250 km^2 in the highest class varied between 122 and 202 avalanches. We simply provide the range from the lowest to the highest class: < 29 to 202.
We simplified the description (lines 465-468).

*20. **Line 389** Add "and terrain choices." at the end of this sentence*

Thanks for the suggestion, we added this additional factor (line 477).

#18 Add "according to our data" after "km^2"

Re-worded as suggested (line 18).

#24 Improve the sentence starting with "For these…"

We revised this statement (lines 25-27).

#35 Improve flow (order of words) of sentence

We rephrased this sentence (lines 36-38).

#35-38: May also use the EAWS description "Avalanche danger is a function of snowpack stability, its spatial distribution and avalanche size" (https://www.avalanches.org/wp-content/uploads/2019/07/general-assembly-oslo_minutes_EAWS.pdf)

Thanks for the suggestion. As the description is not really new, we prefer referring to previously published articles.

#50-51 This description should be updated. The latest EAWS matrix specifically accounts for size (https://www.avalanches.org/standards/eaws-matrix/)

We now mention the present version of the Bavarian matrix (line 57).

#75 Explain if the size of the area of avalanche observations is equal to the size of the forecasting region

The size of the area of avalanche observations corresponds to the typical size of a so-called warning region. The warning region that includes Davos is even a bit smaller, but still representative of the study area. We now specify that the warning region of Davos is slightly smaller but still in the centre of the study region (lines 116-117).

#75 Add a short description of how these observations were obtained. Information is provided in the discussion chapter, but it would be logical for the reader to learn about the data upfront. Did the observations cover the entire 360 km^2? The entire winter season? All seasons with the same rigorousness?

We added some more details on how the observations were made (lines 83-85).

#107 Could replace "In other words, on" by "On average,"

We reworded the start of the sentence (line 120).

#109 Add a sentence at the beginning of the paragraph, about why the work described in the paragraph was carried out. E.g., "The forecasting data were scrutinized, in order to adjust danger levels to the most realistic values."

Thanks for the suggestion. We re-wrote parts of this paragraph (lines 122-129).

#126 Replacing "Moreover, there were also days, 17 in total" with "This was also the case for 17 days" could improve the readability

Changed as suggested (line 147).

#165 Add a descriptor for the values, probably "median values"

Yes, thanks, we added median values (line 195).

#175 Replace "for" by "to"

Changed as suggested (line 205).

#277 Replace "to" by "of"

We changed as suggested (line 338).

#289 Since the Eckerstorfer el al. study, the number of Sentinel-1 satellites has doubled with 1B in orbit. Thus, the statement of too poor temporal resolution is less valid today. I suggest to add this information to the sentence.

We have actually checked this statement last fall and also figured out that S1 now consists of two satellites S1-A and S1-B, which alternately image central Europe every six days from the same orbit. Based on this information we considered the temporal resolution in the Alps as still rather poor, i.e. not sufficient for operational forecasting.
We clarified this point (lines 349-351).

#305 Spell out what is meant by "the potential impact"

We mean that avalanches not causing any damage, or e.g. not reaching the road, are more likely to be not reported.
We now mention this (lines 371-372).

#374 Terrain usage probably also decrease from level 2 to 3, as well as from 3 to 4. It would be useful to add references, if these exist, on the differences in terrain usage between danger levels 1, 2, 3 and 4.

Thanks for this suggestion. We agree that there may be some decrease in usage frequency already from *2–Moderate* to *3–Considerable*. We are aware of two studies that looked into that issue. Techel et al. (2015) analysed avalanche risk based on accident data and usage frequency, they inferred usage frequency by exploring two social media mountaineering websites; their study showed a decrease in

ski touring activity with regard to danger level (*2–Moderate* vs. *3–Considerable*). Wäger and Zweifel (2008) also reported a decrease in touring activity, but no change with regard to off-piste skiing. In the region of Davos, ski touring and off-piste skiing are equally relevant.
We now refer to these two study on usage frequency in relation to the avalanche danger level (lines 456-459).

#377 Explain in more detail how you arrive at the number 10.

This is an informed guess. As described in the paper the average number of natural avalanches at danger level *4–High* was 48. Given the size of our study area (300 km^2), one obtains about 16 avalanches per 100 km^2, hence we wrote “at least 10”. Considering that we do not have full observation coverage in our study area, we could as well suggest about 20 avalanches per 100 km^2. This means the term “many” can be quantified, roughly in the range of 10-20, and it becomes clear that “many” is not just one to three.
We now provide some more background on this estimate (lines 462-464).

#384-389 I find parts of this paragraph unfinished. I would recommend arguing or substantiating why you make the statements “need to” and “should not”. I would also recommend to put the sentence “The actual locations…” into context (e.g. the wordings of the NA/CMAH and EADS wrt. spatial distribution).

Recent discussions in the EAWS and publications (e.g. on the CMAH) have shown that there is potential for confusion with regard to terminology, in particular with regard to avalanche probability. Hence, we simply considered it useful to point out some of the differences since we use some of the terms in the paper as well. There is a similar issue with the spatial distribution of stability as referred to, for instance, in the EAWS definition, you mentioned above. The term spatial distribution can be interpreted in different ways, for instance, spatially and non-spatially. However, the factor contributing to the danger level is the frequency of triggering spots, a non-spatial property. Where the spots in the terrain are located, is not relevant for the definition of the danger levels, only their frequency is relevant. This issue will be clarified in detail in an upcoming manuscript by (Techel et al., 2020, in preparation).
We have partly edited this paragraph (lines 469-477).

References

Harvey, S., Schmudlach, G., Bühler, Y., Dürr, L., Stoffel, A., and Christen, M.: Avalanche terrain maps for backcountry skiing in Switzerland, Proceedings ISSW 2018. International Snow Science Workshop, Innsbruck, Austria, 7-12 October 2018, 1625-1631, 2018.

McClung, D. M., and Schaerer, P.: The Avalanche Handbook, 3rd ed., The Mountaineers Books, Seattle WA, U.S.A., 342 pp., 2006.

Techel, F., Zweifel, B., and Winkler, K.: Analysis of avalanche risk factors in backcountry terrain based on usage frequency and accident data in Switzerland, Nat. Hazards Earth Syst. Sci., 15, 1985-1997, https://doi.org/10.5194/nhess-15-1985-2015, 2015.

Techel, F., Müller, K., and Schweizer, J.: On the importance of snowpack stability, its spatial distribution and avalanche size in assessing the avalanche danger level, in preparation, 2020.

Wäger, P., and Zweifel, B.: 10hochminus5, Bergundsteigen - Zeitschrift für Risikomanagement im Bergsport. Oesterreichischer Alpenverein, Innsbruck, Austria, 17, 22-36, 2008.

On the relation between avalanche occurrence and avalanche danger level

Jürg Schweizer[1], Christoph Mitterer[2], Frank Techel[1], Andreas Stoffel[1], Benjamin Reuter[1,3]

[1] WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland
[2] Avalanche Forecasting Service Tyrol, Innsbruck, Austria
[3] Météo France, CNRS, CNRM, Centre d'Étude de la Neige, Grenoble, France

Correspondence to: Jürg Schweizer (schweizer@slf.ch)

Abstract. In many countries with seasonally snow-covered mountain ranges warnings are issued to alert the public about imminent avalanche danger, mostly employing an ordinal, five-level danger scale. However, as avalanche danger cannot be measured, the characterization of avalanche danger remains qualitative. The probability of avalanche occurrence in combination with the expected avalanche type and size decide on the degree of danger in a given forecast region ($\gtrsim$ 100 km^2). To describe avalanche occurrence probability the snowpack stability and its spatial distribution need to be assessed. To quantify the relation between avalanche occurrence and avalanche danger level we analyzed a large data set of visually observed avalanches (13’918 in total) from the region of Davos (Eastern Swiss Alps, ~ 300 km^2), all with mapped outlines, and compared the avalanche activity to the forecast danger level on the day of occurrence (3533 danger ratings). The number of avalanches per day strongly increased with increasing danger level confirming that not only the release probability but also the frequency of locations with a weakness in the snowpack where avalanches may initiate from, increases within a region. Avalanche size did generally not increase with increasing avalanche danger level, suggesting that avalanche size may be of secondary importance compared to snowpack stability and its distribution when assessing the danger level. Moreover, the frequency of wet-snow avalanches was found to be higher than the frequency of dry-snow avalanches for a given day and danger level; also, wet-snow avalanches tended to be larger. This finding may indicate that the danger scale is not used consistently with regard to avalanche type. Although, observed avalanche occurrence and avalanche danger level are subject to uncertainties, our findings on the characteristics of avalanche activity suggest reworking the definitions of the European avalanche danger scale. The description of the danger levels can be improved, in particular by quantifying some of the many proportional quantifiers. For instance, based on our analyses ‘many avalanches’, expected at danger level *4–High*, means on the order of at least 10 avalanches per 100 km^2. Whereas our data set is one of the most comprehensive, visually observed avalanche records are known to be inherently incomplete so that our results often refer to a lower limit and should be confirmed using other similarly comprehensive data sets.

1 Introduction

Avalanche forecasting was described by McClung (2002) as the prediction of snow instability in space and time relative to a given triggering level. The main sources of uncertainty in forecasting are the unknown temporal evolution and the spatial variations of instability in the snow cover. For these reasons predictability of snow avalanche occurrence is limited; it is inversely related to scale, i.e. a probability of occurrence can be given at the regional scale, but not at the scale of a single avalanche path (Schweizer, 2008). In forecasting of natural systems, in which variations may or may not be random, a distinction is often made between forecasting and prediction. In our case, prediction means precisely defining when and where an avalanche occurs. Forecasting, on the other hand, implies describing the probability of avalanche occurrence within a certain time frame and area. Given these definitions it is obvious that prediction is not possible – even though it would be desirable – whereas forecasting is certainly possible but inherently includes uncertainty as the forecast is probabilistic (Silver, 2012).

Even if avalanche forecasting is probabilistic and includes uncertainty, it should be grounded in clear definitions and uncertainty should not stem from ambiguous definitions but the nature of the problem. In public forecasting, i.e. issuing bulletins describing the avalanche situation, avalanche hazard is described by one of five avalanche danger levels. The danger levels (*1–Low, 2–Moderate, 3–Considerable, 4–High, 5–Very High*) are defined in the avalanche danger scale that was originally agreed by the European avalanche warning services in 1993 (EAWS, 2019a; Meister, 1995). Subsequently, a very similar five-level scale was adopted in North America (Dennis and Moore, 1997), which was later revised with an emphasis on risk communication (Statham et al., 2010). In the original European danger scale, the avalanche danger levels were defined in terms of the release (or triggering) probability, the frequency and location of triggering spots and the potential avalanche size. All three elements are supposed to be combined when assigning a danger level to a given avalanche situation. Moreover, it is assumed that all three elements increase with increasing avalanche hazard. However, the definitions for the different danger levels are short, qualitative descriptions and leave room for widely varying interpretations (Müller et al., 2016a). Not surprisingly, a recent study that looked at forecast differences across borders of contiguous forecast areas suggests that remarkable inconsistencies in the application of the danger levels exist (Techel et al., 2018). Based on a survey among forecasters, Lazar et al. (2016) also found substantial differences in assigning a single danger rating to a given scenario of avalanche conditions. These studies demonstrate that there is a lack of quantification with regard to the three key elements and their links in the avalanche danger scale.

This lack of formal underpinnings, among other reasons, motivated the development of a conceptual model of avalanche hazard in North America, which essentially formalizes the hazard assessment process (Statham et al., 2018). However, the final step on how to derive the danger level is not described. In Europe, the so-called Bavarian matrix was developed to support the decision process in forecasting. It is basically a look-up-table that allows assigning the danger level based on the probability of avalanche release and the frequency of triggering spots (Müller et al., 2016a). Avalanche size is not explicitly considered in the Bavarian matrix. Hence, recent developments in Europe were aiming at including avalanche size and harmonizing the

European with the North American approach. To this end, an approach with two matrices, a so-called likelihood matrix and a danger matrix, was suggested in an attempt to merge the concepts behind the conceptual model of avalanche hazard with the Bavarian matrix (Müller et al., 2016a; Müller et al., 2016b). Also, a version of the Bavarian matrix including avalanche size was suggested (EAWS, 2020).

There are few data-driven studies that link the avalanche danger level to any of the three key elements. Haegeli et al. (2012) analyzed two years of public avalanche forecasts with underlying hazard assessments by Avalanche Canada. They found that the maximum likelihood of triggering had the strongest impact on danger rating selection; the second most important predictor variable was the maximum expected avalanche size. More recent analyses on the relation between the components of the conceptual model of avalanche hazard and the danger ratings showed that identical avalanche scenarios were often rated differently – possibly indicating substantial inconsistencies in the forecasting process. This finding is likely due to the lack of explicitly assigning danger ratings to the various combinations in the likelihood-magnitude chart (Clark and Haegeli, 2018; Clark, 2019).

The avalanche danger levels can also be characterized with observational data related to snow instability. In the context of a verification campaign, Schweizer et al. (2003) established typical stability distributions for the danger levels *1–Low*, *2–Moderate* and *3–Considerable* based on many snow instability tests for single avalanche situations. Likewise, signs of instability such as whumpfs, shooting cracks and recent avalanching were related to the danger levels (Jamieson et al., 2009a; Schweizer, 2010). As shooting cracks were almost ten times more frequent at *3–Considerable* (or higher) than at *2–Moderate* (or lower), they had most predicting power in a simple classification tree. Avalanche activity was only considered as binary variable, which does not allow insight into the avalanche characteristics at a given danger level.

Given the lack of quantitative definitions in the avalanche danger scale, our aim is to characterize avalanche activity with regard to avalanche hazard. We therefore analyzed a large data set of avalanche observations from the region of Davos (Eastern Swiss Alps) and compared the avalanche activity to the avalanche danger forecast. Though both variables are subject to uncertainty, we aim at characterizing the danger levels based on frequency, type and size of avalanche occurrence.

2 Data and methods

We analyzed a 21-year data set of manually observed avalanche occurrences from the region of Davos, an area of about 300 km^2. Large parts of the study area ranging between 1200 and 3200 m a.s.l. are steep mountain terrain. According to the avalanche terrain classification (CAT) by Harvey et al. (2018) about 67 % of our study area are considered avalanche terrain. The snow climate in the region of Davos can be described as transitional (McClung and Schaerer, 2006).

Data cover the winters from 1998-1999 to 2018-2019 and include 13,918 individual avalanches, which were all mapped (Figure A in the Supplement). Avalanches were recorded on a daily basis by SLF as well as by ski resort staff. In the more remote parts of the study area the observations are in general less consistent. Also, during the first six years, fewer avalanches were recorded than during the rest of the study period.

For each avalanche, we derived avalanche length and width from a rectangle of the smallest width enclosing the mapped perimeter using the ‘minimum boundary geometry’ tool in ArcGIS. Based on these values of avalanche length and width we assigned the avalanche size class (1 to 4) according to the Canadian size classification (McClung and Schaerer, 2006). Since avalanches of size class 5 were rare ($N = 11$), we assigned those to class 4. Moreover, 116 avalanches were too small to derive meaningful values of length and width, but were still assigned an avalanche size of 1 (hence $N = 13{,}802$ in Table 1 and Figure 1). Table 1 describes the criteria for size classification. Also given are the resulting median length, width and area per size class for our data set.

Figure 1 shows that our size classification based on mapped outlines well reproduced the exponential increase that underlies the original proposal for the size classification (McClung and Schaerer, 1981). They suggested classifying avalanche size S based on mass and proposed five classes where $S = \log M$ with M the avalanche mass given in tens of tons. Their intention was to derive a classification that is based on destructive power, which in the end is related to volume or length. Figure 1 suggests that estimating avalanche size based on avalanche length seems indeed feasible.

Table 1: Definition of avalanche size based on length and width of avalanche. Resulting median length, width and area per size class ($N = 13{,}802$).

Avalanche size	Class length (m)	Operator	Class width (m)	N	Median length (m)	Median width (m)	Median area (m^2)
1: small	< 50	AND	< 50	501	37	24	654
2: medium	< 300	AND	< 300	9766	144	42	3989
3: large	[300, 999]	OR	[300, 999]	3228	430	91	21,252
4: very large	≥ 1000	OR	≥ 1000	307	1196	256	144,113

In addition, the avalanche records included information on the type of triggering (natural, person, explosives/snow grooming machine, unknown) and the type of snow conditions. The snow conditions relate to the liquid water content in the starting zone (dry, wet, mixed, unknown). Dry and wet refer to dry-snow and wet-snow avalanches, respectively, whereas mixed is less well defined and typically refers to avalanches with dry-snow conditions in the starting zone, but wet-snow conditions in the track or runout zone. Our records of avalanche observations cover 1358 individual days.

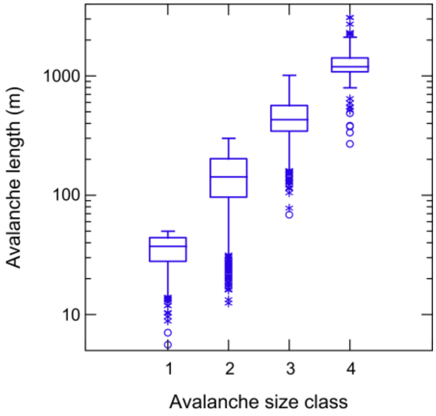

Figure 1: Distribution of avalanche length per avalanche size class for the 21-year data set of avalanche observations from the region of Davos applying the classification criteria given in Table 1. Boxes span the interquartile range from 1st to 3rd quartile with a horizontal line showing the median. Whiskers show the range of observed values that fall within 1.5 times the interquartile range above the 3rd and below the 1st quartile. Asterisks and open circles refer to outliers and far-outliers beyond the fences. Numbers indicate avalanches per class; total number of avalanches: *N* = 13,802.

We calculated the avalanche activity index for each day using the usual weights for size classes 1 to 4, namely 0.01, 0.1, 1, and 10, respectively (Schweizer et al., 2003). Moreover, we considered the type of triggering using weights, namely 1 for natural avalanches, 0.5 for human-triggered avalanches, and 0.2 for the other artificially triggered avalanches (Föhn and Schweizer, 1995). For the avalanches with unknown trigger we assigned a weight of 0.81 since this was the weighted average of the triggering weight considering the frequency of avalanches for the three known triggering classes. In fact, almost all of the avalanches in the unknown triggering class were likely natural avalanches so that a weight of 0.81 was appropriate. We also calculated the individual AAI for the combinations of the various types of triggering and types of snow conditions.

We then merged the data set of avalanche observations with the avalanche danger as forecast in the public bulletin for that day and the region of Davos; the forecasting region is smaller than our study area, but is representative as it is located well in its center where most avalanches were recorded. For a total of 3533 days a danger rating was available. Some of the avalanches occurred outside the period when public forecasts were issued, e.g. in October or late May, and were not included for further analysis. This reduced the total number of observed avalanches to 13,745 and the number of days when at least one avalanche was recorded to 1301. This means, about every third day (37 %) with a danger rating at least one avalanche was observed for the 21-year period we analyzed.

As independent data to verify the issued danger level is not available for the entire data set, we compared the avalanche activity observed on a given day to the forecast danger level on that day. Nevertheless, we did some obvious data checking. For instance, at *4–High* we expect many natural avalanches. Indeed, the two highest danger levels can be verified by avalanche activity. Therefore, we scrutinized the forecast danger levels and adjusted them to most realistic values. All these corrections are summarized in the supplementary material in Table A1. We first checked the days with the highest danger levels, since in these cases erroneous forecasts can most easily be detected: when the danger level was 4–*High* and no avalanches were observed, the forecast was either too high or the avalanche observations were not correctly assigned to the day of occurrence (but rather the day of observation).

We therefore started the correction process by checking the avalanche activity on the 51 days when the danger rating was either *4–High* or *5–Very High* (47 and 4 days, respectively). We found that on 30 out of 51 days the avalanche activity was zero or unusually low. For each of these days, we revisited the weather, snow and avalanche conditions in the relevant period. For 23 out of the 30 days we down-rated the danger. For the remaining 7 days we corrected a temporal mismatch between the date the hazard peaked and the date avalanches were registered. For example, occasionally all avalanche observations from a 3-day storm had been assigned to the first or last day of the storm. For 1 out of these 7 days we increased the danger level from *4–High* to *5–Very High*. This reduced the number of days with rating *4–High* from 47 to 25, and with rating *5–Very High* from 4 to 3. Still, on one day with danger rating *4–High* no avalanches were observed; this seems unlikely, but it was not possible to reconstruct the likely date of occurrence in that well-known storm period in February 1999. Unfortunately, records were in general rather inconsistent during the major storms in January and February 1999. For one day in January 2018, when the forecast danger level was *4–High* and the avalanche activity very prominent (AAI = 158) a detailed verification revealed (Bründl et al., 2019) that the forecast danger level should have been *5–Very High*. This increased the total number of days with danger level *5–Very High* to 4. Hence, after these corrections, the danger rating was 5–*Very High* on 4 days, *4–High* on 25 days and *3–Considerable* on the remaining 22 days.

The median AAI of natural avalanches for the days with danger rating of either *4–High* or *5–Very High* was 13.6, which corresponds to, for instance, only one avalanche of size 4 and a few smaller avalanches. Further quality checking revealed that there were a number of days with higher avalanche activity but lower danger level. In total on 59 days, the avalanche danger was rated *3–Considerable*, but many natural avalanches occurred. This was also the case for 17 days, when danger *2–Moderate* was forecast. Again, we checked all these cases against the weather, snow and avalanche conditions. For 57 of these 59 days we increased the rating from *3–Considerable* to *4–High* since the AAI clearly indicated that the avalanche activity had been underestimated at the time of the forecast. On the remaining two days the number of natural avalanches was too low (<10) to justify a change. For 12 out of 17 days with forecast danger *2–Moderate*, we changed the danger level to *4–High* as many avalanches were observed, in most cases wet-snow avalanches, and the AAI was high. On the remaining 5 days we changed the danger level to *3–Considerable* as the total number of natural avalanches was too low (<10).

Subsequently, we considered the number of cases with *2–Moderate* danger, but an avalanche activity (only naturals) higher than the median index (1.0) for days with *3–Considerable* danger. There were 99 days with AAI > 1.0. In 25 of these cases, the number of avalanches (size 2 and larger) was larger than 10. For these 27 days we changed the danger rating to *3–Considerable*. In 19 out of these 27 cases the avalanches were wet-snow avalanches. Finally, we adjusted the danger level from *1–Low* to *3–Considerable* for 2 days, one with high natural wet-snow avalanche activity and the other with several skier-triggered avalanches.

Overall, we changed 129 out of the 3533 danger ratings (3.7 %), mostly by one danger level, occasionally by two danger levels (12 %); in most cases (106 out of 129: 82 %) we increased the danger rating since there was clearly a rather high activity of natural avalanches. In total, there were finally 94 days (2.7 %) with danger rating *4–High*, still fairly few for 21 winter seasons.

For the analyses, we stratified the data into the danger levels and compared the avalanche activity index (AAI; Schweizer et al., 2003), the proportion of avalanche sizes, the proportion of days when avalanches were observed, and the average number of avalanches per day.

The number of avalanches per day relates to the probability of avalanche occurrence in our study area, i.e. one metric integrating snow stability and its distribution. The lower stability is and the more frequent the triggering locations are, the more avalanches are observed.

To compare distributions, we used the non-parametric Mann-Whitney *U*-test. We selected a level of significance $p = 0.05$ to judge whether the observed differences were significant. We also checked for equality of proportions in 2 x 2 contingency tables, for instance, do decide whether the proportion of size 2 avalanches was equal for two different danger levels. Relations of continuous data such as the AAI with ordinal variables were described with the Spearman rank-order correlation coefficient (Spiegel and Stephens, 1999).

3 Results

3.1 Avalanche activity

Figure 2 shows the avalanche activity index including all avalanches irrespective of triggering type or snow conditions and Table 2 summarizes some key figures on the avalanche activity with respect to danger level. The avalanche activity, expressed as AAI, increased with increasing danger level (Spearman rank order correlation: 0.42, $p < 0.001$). The median values were 0.15, 0.21, 1, 21 and 88 for the avalanche danger levels 1 to 5. The increase was particular prominent from *3–Considerable* to *4–High*. The highest values with AAI > 145 correspond to four distinct, well known avalanche periods in the region of Davos: 23 April 2008, 9 March 2017, 22 January 2018 and 14 January 2019. There were only four days with danger level *5–Very High* so that the corresponding AAI statistics are indicative at best.

The proportion of days when avalanches were observed at a given danger level (AAI > 0), increased from about 9 % at *1–Low* to 99 % for *4–High* (Table 2). If natural avalanches were considered only, these proportions were 6 %, 16 %, 35 % and 95 %. At *1–Low*, natural avalanches were observed at only 1 out of 16 days when this danger level was forecast. At *3–Considerable*, natural avalanches were recorded every third day and at *4–High* on almost all days. This increase of the proportion of avalanche days primarily reflects that snow stability decreases with increasing danger level.

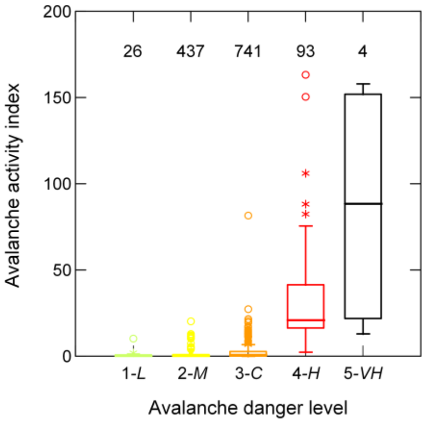

Figure 2: Avalanche activity index AAI per danger level (*1–Low* to *5–Very High*). Only days with AAI > 0 are included. Numbers indicate number of days per danger level; total number of days: *N* = 1301.

Moreover, the number of observed natural avalanches increased with increasing danger level. At the lower danger levels *1–Low* and *2–Moderate*, the median number of avalanches on a day with avalanche activity was 1; at *3–Considerable* the median number increased to 3, with an even stronger increase to 22 natural avalanches per day at *4–High*. As the avalanche records are likely incomplete for two out of four days with danger level *5–Very High* the median number of natural avalanches per day was only 19. This prominent increase of natural avalanche activity with increasing danger level is also evident in the AAI (only considering natural avalanches; median value): 0.1, 0.1, 0.2, 20 and 76 for the danger levels *1–Low* to *5–Very High*, respectively. The increase of the AAI or the number of avalanches per day reflects the increasing avalanche occurrence probability with increasing danger level. The prominent increase is due to decreasing snow stability and at the same time increasing frequency of locations with poor snow stability where avalanches can initiate from.

Table 2: Avalanche activity per danger level. The AAI considers all types of avalanches independent of snow conditions and trigger type; median value per day is given. Moreover, the number of days with either natural or human–triggered avalanches, at least size 2 or larger is shown.

Danger level	Number of days	Number of days with AAI > 0 (proportion in %)	AAI Median	Number of days with natural avalanches (≥ size 2) (proportion in %)	Number of days with human–triggered avalanches (≥ size 2) (proportion in %)
1–Low	303	26 (8.6 %)	0.15	19 (6.3 %)	7 (2.7 %)
2–Moderate	1766	437 (25 %)	0.21	286 (16 %)	144 (8.2 %)
3–Considerable	1366	741 (54 %)	1.0	479 (35 %)	341 (25 %)
4–High	94	93 (99 %)	21	89 (95 %)	36 (38 %)
5–Very High	4	4 (100 %)	88	4 (100 %)	0 (0 %)

Whereas the number of natural avalanches steadily increased with increasing danger level, the relation differed for the human-triggered avalanches – mainly at the higher danger levels. The proportion of days with at least one human-triggered avalanche (≥ size 2) prominently increased from *1–Low* to *3–Considerable* (Table 2), about tripling from one danger level to the next; it less prominently increased to *4–High*, but was 0 at *5–Very High* – indicating that fewer people expose themselves to the hazard. The respective proportions were about 3 %, 8 %, 25 %, 38 % and 0 %.

Hence, triggering at the danger levels *1–Low* and *2–Moderate* was rather rare. In case a human–triggered avalanche was observed (≥ size 2) when the forecast avalanche danger level was *1–Low* or *2–Moderate*, this avalanche was in most cases (67 % and 68 %, respectively) the only one. In the case of natural avalanches, these proportions were lower, 47 % for *1–Low* and 48 % for *2–Moderate*.

The proportion of days with natural avalanches was higher than the proportion of days with human-triggered avalanches at all danger levels – even the lower ones (Table 2). The higher proportions of days with natural avalanches are primarily related to the occurrence of natural wet-snow avalanches at the lower danger levels. Below we will consider avalanche activity with regard to snow conditions and type of triggering in more detail.

3.2 Avalanche size

Smaller avalanches were more common than large avalanches – irrespective of the danger level (Table 3, Figure 3a). The majority of the avalanches recorded (9649 out of 13,745) were size 2 avalanches. This size was the most frequent at all danger levels. Size 3 and size 4 avalanches were less frequent at all danger levels. In other words, for sizes 2 to 4, the frequency of occurrence decreased with increasing avalanche size. The overall frequencies were 4 %, 70 %, 23 % and 2 % for the sizes 1 to 4, respectively (Table 3, bottom row).

Table 3: Frequency of avalanche size per danger level (proportion in %).

Danger level		Number of avalanches				
		Avalanche size class				
	Number of days	1	2	3	4	Total
1–Low	303	3 (5.8 %)	40 (77 %)	8 (15 %)	1 (1.9 %)	52
2–Moderate	1766	73 (5.5 %)	977 (73 %)	262 (20 %)	18 (1.4 %)	1330
3–Considerable	1366	299 (4.8 %)	4608 (74 %)	1277 (20 %)	82 (1.3 %)	6266
4–High	94	220 (3.7 %)	3956 (67 %)	1590 (27 %)	166 (2.8 %)	5932
5–Very High	4	5 (3.0 %)	68 (41 %)	60 (36 %)	32 (19 %)	165
Total	3533	600 (4.4 %)	9649 (70%)	3197 (23 %)	299 (2.2 %)	13,745

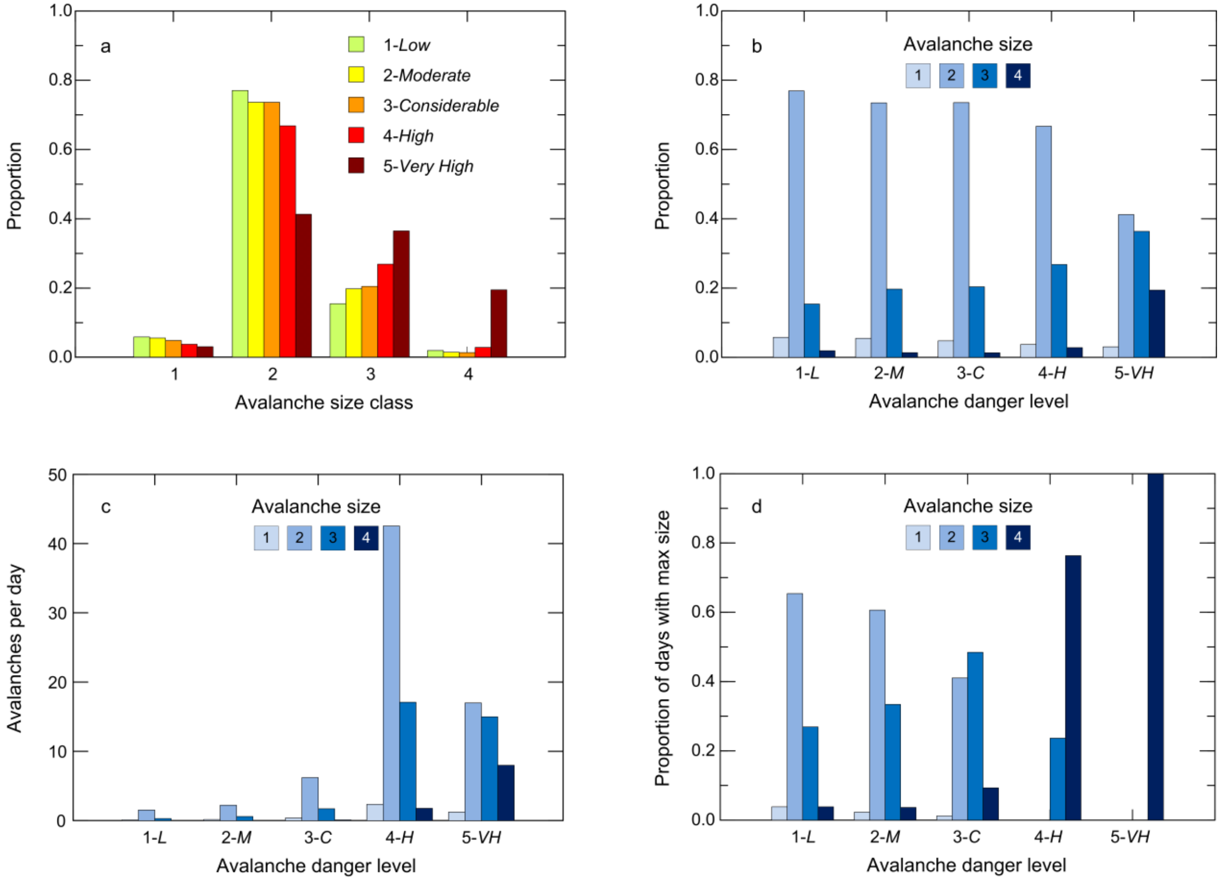

Figure 3: Avalanche size per danger level. (a) Relative frequency of avalanche sizes (1 to 4) at the danger levels *1–Low* to *5–Very High*. (b) Distribution of avalanche size for each danger level (same data as in Fig. 3a). (c) Average number of avalanches per day by size class for each danger level, for days with AAI > 0. (d) Frequency of days when the largest observed avalanche was either size 1, 2, 3 or 4 for each danger level. *N* = 13,745 avalanches on 1301 days.

The distributions of avalanche sizes per danger level (Figure 3b) did not significantly differ for the danger levels 1–*Low* to 3–*Considerable* (see Supplement, Table B). For 4–*High* the distribution looked similar, still the size 2 avalanches were the most frequent followed by size 3 avalanches, but the distribution was statistically different (*U*-test, $p < 0.001$). At danger level 5–*Very High* the distribution was clearly different with relatively more size 4 avalanches and less size 2 avalanches.

Overall, about 80-90 % of the avalanches were size 2 or 3, whereas size 1 and size 4 avalanches were always rare (≲5 %), except at 5–*Very High* when about 19 % of the recorded avalanches were size 4 avalanches. Overall, size 2 avalanches tended to decrease, while size 3 and 4 avalanches tended to increase with increasing danger level (Figure 3b).

The proportion of size 2 avalanches was similar (73 to 77 %) at danger levels *1–Low* to *3–Considerable* (proportion test, $p > 0.57$), but clearly decreased at danger levels *4–High* (67 %; $p < 0.001$) and *5–Very High* (41 %; $p < 0.001$). On the other hand, the proportion of size 3 avalanches increased from 15 % at 1–*Low* to 20 % at both *2–Moderate* and *3–Considerable* to eventually 27 % at *4–High* (Table 3). The latter increase was statistically significant (proportion test, $p < 0.001$). Size 4 avalanches were relatively most frequent at danger level *5–Very High*, about 5 times more frequent than at *4–High* and about 15 times more frequent than at *3–Considerable* (Figure 3c); the increase was statistically significant (proportion test, $p < 0.001$) .

At danger levels *1–Low* and *2–Moderate* avalanches were generally not smaller, but were simply less frequently observed. There was a substantial increase in avalanche occurrence with increasing danger level (Figure 3c) as also reflected in the strong increase of the avalanche activity index (Figure 2). The average total number of avalanches per day was 2, 3, 8, 64 and 41 for the danger levels *1–Low* to *5–Very High*, respectively. Hence, in general, the frequency rather than the size increased with increasing danger level. The median avalanche length was 164, 154, 163 and 198 m for the danger levels *1–Low* to *4–High*, respectively (see Supplement, Figure B). Hence only at danger level 4–*High* the avalanches were about 25 % longer, the difference was statistically significant (*U*-test, $p < 0.001$). Avalanche length was weakly correlated with the danger levels (*1–Low* to *4–High*) (Spearman rank order correlation: 0.13).

Whereas size 2 avalanches were clearly the most frequent at the danger levels *1–Low* to *4–High* and the size distributions looked partly similar, the largest avalanche observed at a given day increased with increasing danger level (Figure 3d). At the danger levels 1–*Low* and 2–*Moderate* the largest avalanche observed was on most days a size 2 avalanche, whereas at the danger levels 4–*High* and 5–*Very High* the largest avalanches were mostly or even always size 4 avalanches, respectively.

3.3 Snow conditions

For about half of all avalanches (56 %) snow conditions were reported as dry, for 30 % as wet and for 6.3 % as mixed; for the remaining 8 % (1072 avalanches) the type of avalanche snow was unknown, i.e. not recorded. As the distributions of avalanche sizes was similar for the mixed and unknown conditions (*U*-test, $p = 0.822$), we merged these two groups. For the three groups (dry, wet, mixed/unknown), again about 70 % were size 2 avalanches and about 20 % size 3 avalanches (Figure 4a). The distributions of avalanche sizes were similar and comparable to the overall distribution (Table 3), but found to be statistically different (*U*-test, $p < 0.001$). Differences in size distribution with regard to snow conditions were however small. For size 2 avalanches, the proportions were even the same (71 %) for both dry and wet-snow conditions (proportion test, $p = 0.54$). The proportion of wet-snow avalanches of size 3 was slightly larger (25 %) than the corresponding proportion of dry-snow avalanches (22 %; $p < 0.001$). Most size 4 avalanches were recorded for mixed or unknown conditions, relatively twice as many as for dry-snow or wet-snow conditions. In general, for mixed or unknown conditions, the avalanche size seems to be somewhat larger. Relatively fewer size 2 and more size 3 and 4 avalanches were reported. In fact, the median avalanche length was 167, 188 and 190 m for dry-snow, wet-snow and mixed/unknown conditions, respectively. The avalanche length distributions were similar for wet-snow and unknown/mixed conditions (*U*-test, $p = 0.55$), but both different from the dry-snow conditions (*U*-test, $p < 0.001$). So far, in Figure 4a, we have considered all avalanches irrespective of the type of triggering. In the following, we will only consider natural avalanches.

Whereas overall the size distributions are not very different (Figure 4a), some differences emerge when considering the size distribution per danger level for dry-snow and wet-snow conditions (Figure 4b,c). For instance, at *1–Low* on a day with wet-snow avalanches also size 3 and size 4 avalanches were recorded. Avalanches tended to be larger and were also more frequent (Figure 4d). On the 10 days with natural wet-snow avalanches, the number of avalanches was 18 and the total AAI was 15.2. Whereas on the 6 days with natural dry-snow avalanches, the number of avalanches was 8 and the total AAI was 0.71. Hence, the AAI was on average more than ten times larger for wet-snow than for dry-snow conditions. On the other hand, as shown in Figure 4d the average number of avalanches per day was 1.3 for dry-snow and 1.8 for wet-snow conditions; this difference is not particularly large. In fact, the difference in the number of avalanches per day is statistically not significant (*U*-test, $p = 0.39$).

At the danger levels *1–Low* to *3–Considerable* there were relatively more size 3 and size 4 wet-snow avalanches recorded than dry-snow avalanches. Whereas for dry-snow conditions the proportion of size 2 avalanches decreased and the proportion of size 3 avalanches increased with increasing danger level, this tendency was not evident for wet-snow conditions.

Natural avalanches under wet-snow conditions tended not only to be larger, but there were also more wet-snow avalanches than dry-snow avalanches observed on a given day. The average number of natural avalanches per day with a given danger level was larger for wet-snow than for dry-snow conditions (Figure 4d). For wet-snow avalanches the average numbers were: 1.8, 2.5, 7.7 and 56 for danger levels *1–Low* to *4–High*, for dry-snow conditions: 1.3, 2.2, 4.8 and 22. For the danger levels *3–Considerable* and *4–High*, the difference in the number of avalanches per day was statistically significant (*U*-test, $p = 0.001$). Hence, the average number of avalanches per day was about 1.6 and 2.6 times larger at *3–Considerable* and *4–High*, respectively, for a day with wet-snow avalanches compared to a day with dry-snow avalanches. Overall, the 3044 natural dry-snow avalanches were recorded on 482 days, the 3955 natural wet-snow avalanches on only 331 days; i.e. on average almost twice as many wet-snow than dry-snow avalanches per day were recorded: 12 vs. 6.3. The difference in the number of avalanches per day between dry- and wet-snow conditions was statistically significant (*U*-test, $p = 0.001$) Hence, under wet-snow conditions, avalanches were not only larger, but also more frequent compared to dry-snow conditions.

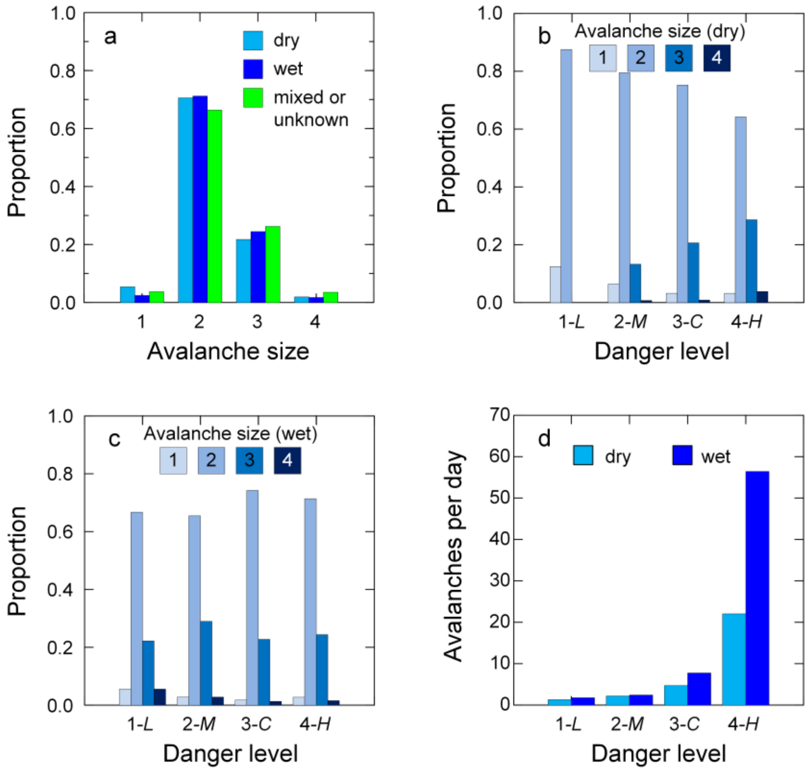

Figure 4: (a) Avalanche size distribution for dry-snow avalanches (*N* = 7748), wet-snow avalanches (*N* = 4056) as well as for avalanches where the snow type was either recorded as mixed or it was not recorded at all (*N* = 1941). Avalanche size distribution per danger level (*1–Low* to *4–High*) for (b) natural dry-snow natural avalanches (*N* = 3044) and (c) natural wet-snow avalanches (*N* = 3955). (d) Number of avalanches on a day with corresponding avalanche activity per danger level for dry-snow and wet-snow conditions.

3.4 Type of triggering

Whereas above we have compared avalanche activity with regard to snow conditions and primarily focused on natural releases, in the following we will only consider dry-snow avalanches and focus on the type of triggering: i.e. compare natural to human-triggered avalanches. For dry-snow conditions, there were about three times more natural (*N* = 3044) than human-triggered avalanches (*N* = 1036). For both natural and human-triggered avalanches, size 2 avalanches were most frequently observed, in 70 % and 74 %, respectively (Figure 5a), followed again by size 3 avalanches (24 % and 15 %). Hence, the frequency distribution was again similar, but statistically significantly different (*U*-test, $p < 0.001$). There were relatively more human-triggered avalanches of size 1 and 2, yet more natural avalanches of size 3 and 4. For instance, size 4 avalanches were 1.7 times more frequent among the natural than the human-triggered avalanches. Still, there were 17 human-triggered size 4 avalanches recorded. In summary, natural dry-snow avalanches tended to be larger than human-triggered dry-snow avalanches.

In total, 27 % of the dry-snow natural avalanches were size 3 and size 4 avalanches, whereas the corresponding proportion was only 17 % among the human-triggered avalanches (proportion test, $p < 0.001$).

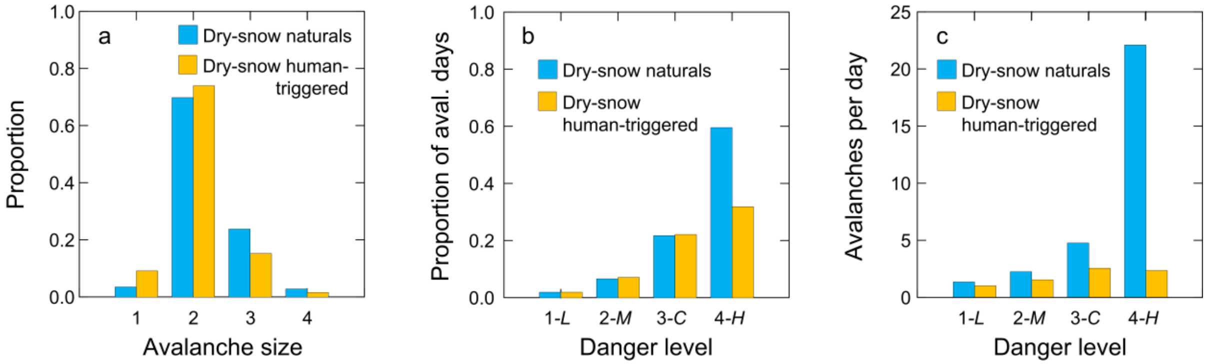

Figure 5: Type of triggering for dry-snow avalanches. (a) Avalanche size distribution for natural and human-triggered avalanches. (b) Proportion of days with either natural or human-triggered avalanches with regard to danger level. (c) Number of avalanches per day vs. danger level.

For the dry-snow human-triggered avalanches the size distributions at the four danger levels *1–Low* to *4–High* were similar to the overall distribution shown in Figure 5a. For instance, at danger level *3–Considerable* and *4–High*, the proportion of size 3 avalanches was 0.74 and 0.73 (proportion test, $p = 0.77$). For size 4 avalanches the corresponding proportions were 0.14 and 0.19 (proportion test, $p = 0.36$). Moreover, the median avalanche length was 154, 139, 128 and 142 m at the danger levels *1–Low* to *4–High*. The corresponding length distributions per danger level were statistically not different (pairwise *U*-test, $p > 0.19$). Hence, avalanche size did not increase with increasing danger level (Spearman rank correlation coefficient: 0.004).

We then considered the relative frequency of avalanche days per danger level (Figure 5b). Human-triggered as well as natural dry-snow avalanches were rare when the danger was rated as *1–Low*. Only in 6 out of 303 days (2 %) with *1–Low* a human-triggered avalanche was recorded, and on 6 days a natural avalanche (considering all size classes). In total there were 6 human-triggered and 8 natural avalanches on 11 individual days, i.e. typically there was one avalanche per day when there were dry-snow avalanches at all at *1–Low*. Hence, avalanche occurrence at *1–Low* is unlikely, irrespective of the type of triggering. The proportion of days with human triggered dry-snow avalanches increased to 7.2 %, 22 % and 32 % for days with forecast danger level of *2–Moderate* to *4–High*, respectively. For natural dry-snow avalanches, the corresponding percentage values were 2 %, 6.7 %, 22 % and 60 %.

Not only the proportion of avalanche days, but also the average number of avalanches per day prominently increased with increasing danger level (Figure 5c). For the human-triggered avalanches at *2–Moderate* the average number per day, when at least one avalanche was recorded, was 1.5, at *3–Considerable* 2.5, but at *4–High* it slightly decreased to 2.3. For the natural dry-snow avalanches, which are more closely related to the release probability and do not depend on the presence of people, the increase was more prominent: 1.3, 2.2, 4.8, and 22 natural avalanches per day with danger rating *1–Low* to *4–High*, respectively. This corresponds to about a 2, 4 and 17 times increase from *1–Low* to the higher levels.

If not only dry-snow natural avalanches are considered, the increase of the average number of avalanches per day is even more prominent. Considering all natural avalanches irrespective of the snow conditions (i.e. dry, wet, mixed and unknown), the number of avalanches per day was 2.1, 2.5, 6.0, 48 and 33 for the danger levels *1–Low* to *5–Very High*, respectively. The higher number of natural avalanches per day, when considering all types of snow conditions instead of dry-snow only, reflects the finding that generally more wet-snow than dry-snow avalanches were observed at a given danger level as shown above in Figure 4d.

4 Discussion

We analyzed a data set of visually observed avalanches from the region of Davos (Switzerland). Even though the data set was collected with the aim to record all (or at least as many as possible) avalanches in the region of Davos, also our data set certainly cannot provide the complete picture of avalanche activity. Small avalanches may be underreported in general. Moreover, avalanche records based on visual observations are known to be biased since during times of poor visibility it is often difficult, and sometimes even impossible, to accurately outline the avalanche extent or record the release date (van Herwijnen and Schweizer, 2011). Only with remote avalanches detection systems the observation bias during storms can be overcome, at least with regard to the temporal resolution (Lacroix et al., 2012; Ulivieri et al., 2011). A good spatial resolution, can only be achieved with remote sensing from satellites (Eckerstorfer et al., 2017). However, presently the temporal resolution over the Alps is still too poor for operational purposes, since only every 6 days images are acquired from the same orbit, which is necessary for change detection. Hence, when we provide the number of avalanches per day observed in our study region, this number should be considered as a lower limit, since with visual observation a full coverage of the study area cannot be achieved.

Moreover, there may be other biases as it is, for instance, easier to record wet-snow than dry-snow avalanches. Also, the level of reporting varied during the 21 winter seasons. There were relatively more observations in the period of 2004-2005 to 2018-2019, than in the first 6 years of our study period. Still, size 2 and size 3 avalanches were the most frequent ones in both periods. The proportion of size 2 avalanches was about 70 % in both periods (proportion test, $p = 0.93$). The proportion of size 3 avalanches was 22 and 23 % during the first and second period, respectively, again not significantly different ($p = 0.28$). Hence, the key characteristics did not change much.

With 13,918 avalanches the data set is extensive and covers many different avalanche situations; small as well as large avalanches, single avalanches as well as records from intense storms with many avalanches – in short, it seems a rather complete data set. A much smaller data set with also mapped avalanche perimeters for the surroundings of the village Zuoz in the lower Engadine (Swiss Alps) was analyzed by Stoffel et al. (1998). In France, the “Enquête permanente sur les avalanches” (EPA) is an extensive data set including the avalanche events in approximately 5,000 major paths in the French Alps and the Pyrenees (e.g., Eckert et al., 2010). Primarily large avalanches that can threaten infrastructure are recorded. Hence this data set seems biased towards very large avalanches and less suited for our purpose. Other extensive data sets were record along mountain passes, e.g. along the Milford road in New Zealand (Hendrikx et al., 2005). They found a significant lack of smaller sizes in their size class distribution, which contained 1842 avalanches. Since only larger avalanches (size ≥ 2.5) are relevant for road safety, smaller avalanches were under-reported. As similar underestimation of size 1 and size 2 avalanches was present in avalanche records observed along the highway crossing Rogers Pass (Canada) (McClung and Schaerer, 1981). Hence, our data set seems to be one of the most comprehensive ones – providing unique insight into avalanche activity. In general, the type of recording and the potential impact (e.g., whether the avalanche hit the road and/or caused damage) may represent the most relevant biases in avalanche data sets.

The avalanche activity as described based on our data set, also depends on terrain characteristics and snow climate. Hence, the characteristics of avalanche activity in our study area cannot simply be transferred to other mountain regions. Nevertheless, we suppose the avalanche size distribution for a given danger level to be fairly robust and rather independent of terrain and climate characteristics. However, the frequency and intensity of avalanche occurrence will certainly vary between regions.

The proposed size classification based on perimeter data had previously been used to study indicator avalanches in the regions of Davos (Schweizer et al., 2012). With the suggested classification criteria, the resulting median length (Table 1) is well in line with the typical values associated with the corresponding size classes provided by the European avalanche warning services (EAWS, 2019b). They indicate typical length categories of <50 m, 50-200 m, several 100 m, 1-2 km for size classes 1 to 4, respectively.

We then compared avalanche activity to the forecast danger level. Again, this is far from perfect as one would need the verified danger level to compare with. Whereas verification at the lower danger levels can be done with numerous snow instability tests (Schweizer et al., 2003), measurements (Reuter et al., 2015) or by expert opinion (Techel and Schweizer, 2017), at the higher danger levels (*4–High* and *5–Very High*) avalanche activity is the most reliable hazard indicator.

As no consistent verification data existed for the entire data set, we had to use the forecast. Still, we tried to remove some obvious outliers such as days with forecast danger level *4–High* but no avalanche records. This quality check should not be considered as comprehensive verification. In total, we changed less than 4 % of the danger ratings and mostly increased the danger level (Table A in the Supplement). In contrast, most verification studies showed a forecast accuracy of about 70-80 % and a trend to over-forecasting (e.g., Techel and Schweizer, 2017). Hence, our avalanche danger data are still biased, yet also reflect some past and recent practice of applying the danger levels. For example, the danger level *4–High* was relatively rarely forecast (on less than 3 % of the days). This may partly be explained by the location of Davos, which is somewhat protected from major storms, but also relates to forecast practice in Switzerland (Techel et al., 2018). However, it is also remarkable that similar avalanche activity was often differently rated for dry-snow and wet-snow conditions – at all danger levels. This finding adds to the list of inconsistencies in avalanche warning as recently reported in several studies (Clark and Haegeli, 2018; Lazar et al., 2016; Techel et al., 2018).

Correcting the forecast danger level based on avalanche activity, while analyzing the relation between danger level and avalanche activity, may seem questionable. The main effect of the correction procedure was that the median AAI for days with 4–*High* increased from about 10 to about 21. For days with 3–*Considerable*, the median AAI was 1 and did not change due to the corrections. Hence, the difference in avalanche activity between 3–*Considerable* and 4–*High* was already very prominent before the correction procedure (before and after the corrections: *U*-test, $p < 0.001$). With regard to avalanche size, the effects are less prominent. Size 2 avalanches were the most frequent ones at danger levels 1–*Low* to 4–*High* before and after the corrections.

The avalanche size distribution we found was remarkably robust with regard to different data stratifications. In particular, the size distribution did not change substantially with the danger level (Figure 3). In other words, for our data set, avalanche size did not prominently increase with increasing danger level, at least for the danger levels *1–Low* to *4–High*; size 2 avalanches were the most frequent (Figure 3b). This finding seems somewhat surprising, given that the avalanche danger level is characterized by a combination of the probability of avalanche occurrence and expected avalanche size (Meister, 1995); it suggests that avalanche size may be of secondary importance compared to snowpack stability and its distribution when assessing the danger level (Techel et al., 2020). Hence, it seems unlikely that the typical (or most frequent) avalanche size is decisive for choosing between four different danger levels for one given snow stability scenario as suggested recently by Müller et al. (2016a). Also, in the conceptual model of avalanche hazard (CMAH) a frequency-magnitude (or likelihood-size) matrix was suggested to estimate avalanche hazard (Statham et al., 2018).

On the other hand, considering the largest avalanche recorded on a given day as suggested by Techel et al. (2020) showed more prominent differences between the danger levels *1–Low* to *2–Moderate* compared to *4–High* to *5–Very High* (Figure 3d). Hence, the maximum expected avalanche size may be useful to differentiate between some hazard situations (Techel et al., 2020). Clark and Haegeli (2018) also reported maximum expected size to be the second most relevant factor for selecting a hazard rating.

Moreover, our findings on avalanche size are in line with the results of a study on avalanche incidents in relation to the danger rating. Harvey (2002) reported that length, width and fracture depth of human-triggered avalanches were very similar at the danger levels *1–Low* to *4–High*, hence did not increase with increasing danger level; the median length was 200 to 250 m, the width 50 to 60 m, which corresponds to avalanche size 2, in agreement with our analysis. However, when he considered all avalanches that had caused damage, i.e. not only human-triggered avalanches, but also avalanches that had destroyed trees or infrastructure, he found avalanche size to be larger at the danger levels *4–High* and *5–Very High* than at the lower danger levels. Logan and Greene (2018) recently also related avalanche size to danger level. They reported that size 2 avalanches (destructive size) were the most frequent size at the danger levels *2–Moderate*, *3–Considerable* and *4–High* – in agreement with our findings.

The relative frequency of the avalanche size classes 2, 3 and 4 were 70 %, 23 % and 2 %, respectively. Hence the frequency of avalanche sizes 2 to 4 decreased with increasing size (Table 3, bottom row). Stoffel et al. (1998) also reported decreasing frequency of occurrence for their avalanche size classes medium, large and very large. These findings are in line with the magnitude-frequency relation of most natural hazard events including earthquakes for which the relation is known as Gutenberg-Richter law (Jentsch et al., 2006). Several other studies have shown frequency-size power-laws for snow avalanches (e.g., Birkeland and Landry, 2002; Faillettaz et al., 2004). The fact that size 1 avalanches were not the most frequent, as theoretically should be the case, is probably related to an observation bias: small avalanches may often not be mapped, in particular when other larger avalanches occur, or cannot be mapped at all if they are too small.

The average number of observed natural avalanches strongly increased with increasing danger level: 2.1, 2.5, 6 and 48 for the danger levels *1–Low* to *4–High*, respectively. This suggests that the differentiation between the lower two danger levels cannot be based on avalanche occurrence. Also, the relative increase from one danger level to the other is increasing – suggesting an exponential increase of the hazard. Previously, it was suggested that the hazard would double from one level to the other (Munter, 2003). Using accident data a 2- to 3-fold increase was shown (Pfeifer, 2009; Techel et al., 2015), whereas a survey among North American avalanche professionals suggested a 10-fold increase of triggering probability when the regional danger increases by one level (Jamieson et al., 2009b). Based on avalanche observations from Colorado, Elder and Armstrong (1987) assigned avalanche frequencies per day to the four danger levels that were in use at those times: 0-3, 4-9, 10-20 and ≥21.

The prominent, non-linear increase of avalanche occurrence with increasing danger level reflects that, according to its definition, the avalanche danger level increases with decreasing snow stability. With decreasing snow stability, the frequency of locations with a potential weakness where an avalanche may be released increases. However, the suggested increase of weaknesses in the snowpack can only be assessed with spatial variability studies (e.g., Reuter et al., 2016), which are most appropriate to determine the spatial distribution of instabilities. For example, Schweizer et al. (2003) reported an increase of the proportion of poorly rated profiles from virtually 0 % to 24 % to 53 % for the danger levels *1–Low* to *3–Considerable*, respectively. Their observations correspond to our finding that the number of natural dry-snow avalanches doubled from *2–Moderate* to *3–Considerable*, and increased almost three times for wet-snow avalanches. Whereas the number of natural dry-snow avalanches consistently increased with increasing danger level, this was not the case for the human-triggered avalanches. The frequency of human-triggered avalanches did not increase from *3–Considerable* to *4–High*. This finding does not mean that triggering becomes less likely but rather reflects terrain usage and the effect of avalanche warnings. In fact, Techel et al. (2015) showed a decrease in ski touring activity already from *2–Moderate* to *3–Considerable*, and even more prominently from *3–Considerable* to *4–High*. On the other hand, (Wäger and Zweifel, 2008) found no decrease in frequency usage rom *2–Moderate* to *3–Considerable* when considering off-piste skiing.

At danger level 4–*High*, the average number of natural avalanches per day in our study region was about 20 for dry-snow avalanches and about 50 for wet-snow avalanches. If we assume that '*many'* natural avalanches are typically observed at the danger level *4–High* (EAWS, 2019a), we may conclude that 10-20 avalanches per 100 km^2 have to be expected, considering some underreporting in our study area of about 300 km^2, where about two thirds are avalanche terrain. Hence, our analyses suggest that in the terrain typical of our study area at least about 10 natural avalanches have to be observed for verifying the danger level *4–High*. Based on satellite images, Bründl et al. (2019) recently analyzed the avalanche activity during a major avalanche cycle in January 2018 when the danger level was 5–*Very High*. They analyzed the number of size 4 avalanches per area; for an area of 250 km^2 the number of size 4 avalanches ranged from less than 29 to up to 202 – roughly consistent with our observations.

With regard to the definition of the avalanche danger by snow stability, its spatial distribution and avalanche size, we would like to point out that the spatial distribution does only refer to the frequency of locations with very poor snow stability where avalanches can initiate from. In other words, where the points with very poor snow stability are located in space is irrelevant, only their frequency counts, when deciding on a given danger level. Moreover, there is a difference between the probability of avalanche release, which is a local property related to local snow instability as recently re-visited by Reuter and Schweizer (2018), and the avalanche occurrence probability, which depends on scale and is the result of stability and its distribution (frequency of triggering spots) for a given area. Moreover, there is a third probability, not to be confused with the two previously mentioned, namely the probability of triggering an avalanche faced by an individual who travels in avalanche terrain on a given day with a given danger level, which also depends on scale (terrain travelled) and terrain choices.

Finally, when the definition for *2–Moderate* danger states that "*Large natural avalanches are unlikely*" (EAWS, 2019a), this definition could as well be modified to "*Natural avalanches are unlikely*" since the probability for any size of natural avalanche at *2–Moderate* is less than 5 %, which according to the IPCC (IPCC, 2014) is best described by '*very unlikely*'. Likewise, the formulation in the definition of *1–Low* "*Only small and medium avalanches are possible*." is not appropriate and should be modified. Hence, our analyses of the avalanche activity may be used to improve the descriptions in the avalanche danger scale.

5 Conclusions

We quantified some of the key characteristics such as frequency and size of avalanches at a given danger level. To this end, we analyzed a unique data set of 21 years of visually observed avalanche records from the surroundings of Davos (Eastern Swiss Alps, 300 km^2), consisting of the mapped outlines of 13,745 avalanches, and compared the characteristics of avalanche activity to the regional danger level as forecast on 3533 days.

The proportion of days with natural avalanches at a given danger level substantially increased with increasing danger level. Also, the overall number of avalanches per day prominently increased, which reflects that with increasing danger level snow stability decreases and the frequency of locations with a potentially critical weakness in the snowpack increases. The recorded number of avalanches per day in our study region (300 km^2) was 2, 3, 8 and 64 for the danger levels *1–Low* to *4–High*, respectively.

The relative frequency of the four avalanche size classes did not substantially change with increasing danger level, neither for human-triggered nor for natural avalanches, except for danger level *5–Very High*. In other words, avalanche size did not increase with increasing danger level: the most frequent avalanches were size 2 avalanches at any danger level. This suggests that avalanche size may be of secondary importance compared to snowpack stability and its distribution when assessing the danger level. Only in certain situations avalanche size may be decisive – and rather by considering the maximum expected size. Still, the absolute number of very large avalanches (size 4) per day prominently increased, namely by 20 times from *3–Considerable* to *4–High*.

At a given danger level the frequency of natural avalanches was typically larger for wet-snow conditions than for dry-snow conditions and wet-snow avalanches tended to be larger – potentially reflecting inconsistence usage of the danger scale in Switzerland. Based on our findings, we propose revisiting the definitions of the danger scale and possibly quantifying some of the descriptions. For example, we suggest that '*many avalanches*' may mean on the order of at least about 10 avalanches per 100 km^2.

We are aware that visual observations are notoriously incomplete. Hence, our results should be challenged by similar analyzes with similarly extensive data sets. In future, more comprehensive data sets based on remotely-sensed data and results from avalanche detection systems may allow better founded analyses.

Finally, our analyses confirm that different avalanche situations are typically condensed into one specific danger level, which results in a loss of information. Hence, avalanche warning services are encouraged to describe the danger as best as they can, and not only provide the danger level. Likewise, risk assessment in avalanche education should not focus solely on the danger level, or the release probability, but as well include the potential consequences.

Data availability

The data sets of avalanche observations and avalanche danger will become available at www.envidat.ch.

Author contributions

JS, CM and BR designed the study. AS, FT, JS and CM extracted and curated the data. JS analyzed the data and prepared the manuscript with contributions from all co-authors.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We would like to thank everyone who contributed to the avalanche occurrence data set, in particular the interns of the avalanche warning service who mapped most of the avalanches. We are grateful for the constructive comments by the two reviewers and the editor that helped to improve the paper.

References

Birkeland, K., and Landry, C. C.: Power-laws and snow avalanches, Geophys. Res. Lett., 29, 1554, https://doi.org/10.1029/2001GL014623, 2002.

Bründl, M., Hafner, E., Bebi, P., Bühler, Y., Margreth, S., Marty, C., Schaer, M., Stoffel, L., Techel, F., Winkler, K., Zweifel, B., and Schweizer, J.: Ereignisanalyse Lawinensituation im Januar 2018, WSL Bericht, 76, edited by: Steffen, K., Federal Research Institute WSL, Birmensdorf, Switzerland, 162 pp., 2019.

Clark, T., and Haegeli, P.: Establishing the link between the conceptual model of avalanche hazard and the North American public avalanche danger scale: initial explorations from Canada, Proceedings ISSW 2018. International Snow Science Workshop, Innsbruck, Austria, 7-12 October 2018, 1116-1120, 2018.

Clark, T.: Exploring the link between the conceptual model of avalanche hazard and the North American public avalanche danger scale, M.Sc., Faculty of Environement, Simon Fraser University, Burnaby BC, Canada, 104 pp., 2019.

Dennis, A., and Moore, M.: Evolution of public avalanche information: The north American experience with avalanche danger rating levels, Proceedings International Snow Science Workshop, Banff, Alberta, Canada, 6-10 October 1996, 60-66, 1997.

EAWS: Avalanche Danger Scale, https://www.avalanches.org/standards/avalanche-danger-scale/, access: 27 July 2019, 2019a.

EAWS: Avalanche Size, https://www.avalanches.org/standards/avalanche-size/, access: 27 July 2019, 2019b.

EAWS: EAWS Matrix, https://www.avalanches.org/standards/eaws-matrix/, access: 17 January 2020, 2020.

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 Reg. Sci. Technol., 144, 39-51, https://doi.org/10.1016/j.coldregions.2017.08.004, 2017.

Eckert, N., Parent, E., Kies, R., and Baya, H.: A spatio-temporal modelling framework for assessing the fluctuations of avalanche occurrence resulting from climate change: application to 60 years of data in the northern French Alps, Climatic Change, 101, 515-553, https://doi.org/10.1007/s10584-009-9718-8, 2010.

Elder, K., and Armstrong, B.: A quantitative approach for verifying avalanche hazard ratings, Symposium at Davos 1986 - Avalanche Formation, Movement and Effects, IAHS Publ., 162, 593-601, 1987.

Faillettaz, J., Louchet, F., and Grasso, J. R.: Two-threshold model for scaling laws of noninteracting snow avalanches, Phys. Rev. Lett., 93, 208001/208001-208004, https://doi.org/10.1103/PhysRevLett.93.208001, 2004.

Föhn, P. M. B., and Schweizer, J.: Verification of avalanche danger with respect to avalanche forecasting, in: Les apports de la recherche scientifique à la sécurite neige, glace et avalanche. Actes de Colloque, Chamonix, 30 mai-3 juin 1995, edited by: Sivardière, F., ANENA, Grenoble, France, 151-156, 1995.

Haegeli, P., Falk, M., and Klassen, K.: Linking avalanche problems to avalanche danger - A first statistical examination of the conceptual model of avalanche hazard, Proceedings ISSW 2012. International Snow Science Workshop ISSW 2012, Anchorage AK, U.S.A., 16-21 September 2012, 859-865, 2012.

Harvey, S.: Avalanche incidents in Switzerland in relation to the predicted danger degree, Proceedings ISSW 2002. International Snow Science Workshop, Penticton BC, Canada, 29 September-4 October 2002, 443-448, 2002.

Harvey, S., Schmudlach, G., Bühler, Y., Dürr, L., Stoffel, A., and Christen, M.: Avalanche terrain maps for backcountry skiing in Switzerland, Proceedings ISSW 2018. International Snow Science Workshop, Innsbruck, Austria, 7-12 October 2018, 1625-1631, 2018.

Hendrikx, J., Owens, I., Carran, W., and Carran, A.: Avalanche activity in an extreme maritime climate: The application of classification trees for forecasting, Cold Reg. Sci. Technol., 43, 104-116, https://doi.org/10.1016/j.coldregions.2005.05.006, 2005.

IPCC: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, IPCC, Geneva, Switzerland, 151 pp., 2014.

Jamieson, B., Haegeli, P., and Schweizer, J.: Field observations for estimating the local avalanche danger in the Columbia Mountains of Canada, Cold Reg. Sci. Technol., 58, 84-91, https://doi.org/10.1016/j.coldregions.2009.03.005, 2009a.

Jamieson, B., Schweizer, J., and Shea, C.: Simple calculations of avalanche risk for backcountry skiing, Proceedings ISSW 2009. International Snow Science Workshop, Davos, Switzerland, 27 September - 2 October 2009, 336-340, 2009b.

Jentsch, V., Kantz, H., and Albeverio, S.: Extreme events: magic, mysteries and challenges, in: Extreme events in nature and society, edited by: Jentsch, V., Kantz, H., and Albeverio, S., Springer, Berlin, Germany, 1-18, 2006.

Lacroix, P., Grasso, J. R., Roulle, J., Giraud, G., Goetz, D., Morin, S., and Helmstetter, A.: Monitoring of snow avalanches using a seismic array: Location, speed estimation, and relationships to meteorological variables, J. Geophys. Res.-Earth Surf., 117, F01034, https://doi.org/10.1029/2011jf002106, 2012.

Lazar, B., Trautman, S., Cooperstein, M., Greene, E., and Birkeland, K.: North American avalanche danger scale: are public backcountry forecasters applying it consistently?, Proceedings ISSW 2016. International Snow Science Workshop, Breckenridge CO, U.S.A., 3-7 October 2016, 457-465, 2016.

Logan, S., and Greene, E.: Patterns in avalanche events and regional scale avalanche forecasts in Colorado, USA, Proceedings ISSW 2018. International Snow Science Workshop, Innsbruck, Austria, 7-12 October 2018, 1059-1062, 2018.

McClung, D. M., and Schaerer, P.: Snow avalanche size classification, Proceedings of a Avalanche Workshop, Vancouver BC, Canada, 3-5 November 1980, 12-27, 1981.

McClung, D. M.: The elements of applied avalanche forecasting - Part I: The human issues, Nat. Hazards, 26, 111-129, https://doi.org/10.1023/a:1015665432221, 2002.

McClung, D. M., and Schaerer, P.: The Avalanche Handbook, 3rd ed., The Mountaineers Books, Seattle WA, U.S.A., 342 pp., 2006.

Meister, R.: Country-wide avalanche warning in Switzerland, Proceedings International Snow Science Workshop, Snowbird, Utah, U.S.A., 30 October-3 November 1994, 1995.

Müller, K., Mitterer, C., Engeset, R. V., Ekker, R., and Kosberg, S. Ø.: Combining the conceptual model of avalanche hazard with the Bavarian matrix, Proceedings ISSW 2016. International Snow Science Workshop, Breckenridge CO, U.S.A., 3-7 October 2016, 472-479, 2016a.

Müller, K., Stucki, T., Mitterer, C., Nairz, P., Konetschny, H., Feistl, T., Coléou, C., Berbenni, G., and Chiambretti, I.: Towards an improved European auxiliary matrix for assessing avalanche danger levels, Proceedings ISSW 2016. International Snow Science Workshop, Breckenridge CO, U.S.A., 3-7 October 2016, 1229-1231, 2016b.

Munter, W.: 3x3 Lawinen - Risikomanagement im Wintersport, Pohl&Schellhammer, Garmisch-Partenkirchen, Germany, 223 pp., 2003.

Pfeifer, C.: On probabilities of avalanches triggered by alpine skiers. An empirically driven decision strategy for backcountry skiers based on these probabilities, Nat. Hazards, 48, 425-438, https://doi.org/10.1007/s11069-008-9270-2, 2009.

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.

Reuter, B., Richter, B., and Schweizer, J.: Snow instability patterns at the scale of a small basin, J. Geophys. Res.-Earth Surf., 121, 257-282, https://doi.org/10.1002/2015JF003700, 2016.

Reuter, B., and Schweizer, J.: Describing snow instability by failure initiation, crack propagation, and slab tensile support, Geophys. Res. Lett., 45, 7019-7027, https://doi.org/10.1029/2018GL078069, 2018.

Schweizer, J., Kronholm, K., and Wiesinger, T.: Verification of regional snowpack stability and avalanche danger, Cold Reg. Sci. Technol., 37, 277-288, https://doi.org/10.1016/S0165-232X(03)00070-3, 2003.

Schweizer, J.: On the predictability of snow avalanches, Proceedings ISSW 2008. International Snow Science Workshop, Whistler, Canada, 21-27 September 2008, 688-692, 2008.

Schweizer, J.: Predicting the avalanche danger level from field observations, Proceedings ISSW 2010. International Snow Science Workshop, Lake Tahoe CA, U.S.A., 17-22 October 2010, 162-165, 2010.

Schweizer, J., Alig, C., and Mitterer, C.: On indicator path avalanches for local avalanche forecasting, Proceedings ISSW 2012. International Snow Science Workshop, Anchorage AK, U.S.A., 16-21 September 2012, 51-54, 2012.

Silver, N.: The Signal and the Noise - The Art and Science of Prediction, Pinguin Books, 544 pp., 2012.

Spiegel, M. R., and Stephens, L. J.: Schaum's Outline of Theory and Problems of Statistics, 3rd ed., Schaum's outline series, McGraw-Hill, New York, 538 pp., 1999.

Statham, G., Haegeli, P., Birkeland, K. W., Greene, E., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., and Kelly, J.: The North American public avalanche danger scale, Proceedings ISSW 2010. International Snow Science Workshop, Lake Tahoe CA, U.S.A., 17-22 October 2010, 117-123, 2010.

Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., and Kelly, J.: A conceptual model of avalanche hazard, Nat. Hazards, 90, 663-691, https://doi.org/10.1007/s11069-017-3070-5, 2018.

Stoffel, A., Meister, R., and Schweizer, J.: Spatial characteristics of avalanche activity in an alpine valley - a GIS approach, Ann. Glaciol., 26, 329-336, https://doi.org/10.3189/1998AoG26-1-329-336, 1998.

Techel, F., Zweifel, B., and Winkler, K.: Analysis of avalanche risk factors in backcountry terrain based on usage frequency and accident data in Switzerland, Nat. Hazards Earth Syst. Sci., 15, 1985-1997, https://doi.org/10.5194/nhess-15-1985-2015, 2015.

Techel, F., and Schweizer, J.: On using local avalanche danger level estimates for regional forecast verification, Cold Reg. Sci. Technol., 144, 52-62, https://doi.org/10.1016/j.coldregions.2017.07.012, 2017.

Techel, F., Mitterer, C., Ceaglio, E., Coléou, C., Morin, S., Rastelli, F., and Purves, R. S.: Spatial consistency and bias in avalanche forecasts – a case study in the European Alps, Nat. Hazards Earth Syst. Sci., 18, 2697-2716, https://doi.org/10.5194/nhess-18-2697-2018, 2018.

Techel, F., Müller, K., and Schweizer, J.: On the importance of snowpack stability, its spatial distribution and avalanche size in assessing the avalanche danger level, in preparation, 2020.

Ulivieri, G., Marchetti, E., Ripepe, M., Chiambretti, I., Rosa, G. D., and Segor, V.: Monitoring snow avalanches in Northwestern Italian Alps using an infrasound array, Cold Reg. Sci. Technol., 69, 177-183, https://doi.org/10.1016/j.coldregions.2011.09.006, 2011.

van Herwijnen, A., and Schweizer, J.: Monitoring avalanche activity using a seismic sensor, Cold Reg. Sci. Technol., 69, 165-176, https://doi.org/10.1016/j.coldregions.2011.06.008, 2011.

Wäger, P., and Zweifel, B.: 10hochminus5, Bergundsteigen - Zeitschrift für Risikomanagement im Bergsport. Oesterreichischer Alpenverein, Innsbruck, Austria, 17, 22-36, 2008.