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 will follow 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 will discuss this issue in the revised manuscript at the beginning of the Discussion section where we already discuss limitations of our dataset.

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 will add these numbers to the Abstract.

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

Yes, thanks, we will change as suggested.

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

Yes, of course, it makes it a stronger case.

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

We will add this restriction.

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

We will elaborate this statement in the revised manuscript.

#35 Improve flow (order of words) of sentence

We will consider rephrasing this statement in the revised manuscript.

#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 refer to recent developments, which are not finished yet and are well described in the references we provide. The version of the Bavarian matrix you refer to is an intermediate step, we may refer to as you wish.

#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.

#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 will add some more details on how the observations were made in the revised manuscript.

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

Yes, one could. We will consider rephrasing the start of the sentence in the revised manuscript.

#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 will add a sentence at the beginning of the paragraph.

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

We will improve readability of this sentence.

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

Yes, thanks, we provide indeed the median values. We will specify this in the revised manuscript.

#175 Replace "for" by "to"

We will change as suggested.

#277 Replace "to" by "of"

We will change as suggested.

#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 will clarify this point in the revised manuscript.

#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.

#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.

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

This is an informed guess. The number may as well be 20. 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,

360 km^2, one obtains 13 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.

#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).