Reply to Grant Statham

We sincerely thank Grant Statham for his thorough and constructive review. Please find our answers to each of the points raised below.

General comments

Scale

I think a discussion on scale and scale issues is necessary for this manuscript. The concept of scale, and in particular the “point scale” is missing, yet measurements at the point scale are fundamental to this proposed EAWS matrix. The proposed “frequency of snowpack stability” method is based on the frequency of point scale measurements, yet this is unclear because snow stability is traditionally thought of as a “slope stability” estimate, and well-established systems exist in different countries for estimating it (CAA, 2016; AAA, 2022).

Reuter et al. (2015) state that “snow slope stability describes the mechanical state of the snow cover on an inclined slope” then further add that “the link between point observations of snow stability and snow slope stability is not clear” and that “the point stability scale is not even well defined”. Reuter and Schweizer (2018) state “a description of snow instability at the scale of the snow cover or a point in the terrain is much needed, yet presently lacking” and present a first framework for doing so. The Conceptual Model of Avalanche Hazard (Statham et al., 2018) describes spatial scales used in avalanche forecasting, but the point scale is missing at the lowest end. Techel et al. (2020) state “the probability of avalanche release refers to a specific location and relates to the local (or point) snow instability”. Thus, according to this, the probability of avalanche release can only be determined in a single spot, until these spots are combined using the frequency method proposed here. Schweizer et al. (2023) suggest a definition (redefinition) of snow stability as “point snow stability refers to snowpack layering, propensity for failure initiation and onset of crack propagation” and distinguishes this from “slope stability”. All of this to say that although the concept of snow stability as a point scale measurement has been in development for close to a decade, this is not well established in practice. The concept is crucial and needs to be included here as it is the basis for this method. Scale issues are fundamental to avalanche forecasting and thus some explanation is necessary in this manuscript and I suggest a section dedicated to scale issues with an explanation and definition of snow stability as a point scale measurement.

I think a discussion on scale and scale issues is necessary for this manuscript. The concept of scale, and in particular the “point scale” is missing, yet measurements at the point scale are fundamental to this proposed EAWS matrix. The proposed “frequency of snowpack stability” method is based on the frequency of point scale measurements, yet this is unclear because snow stability is traditionally thought of as a “slope stability” estimate, and well-established systems exist in different countries for estimating it (CAA, 2016; AAA, 2022).

We agree that a standard definition of the point scale is currently lacking. However, we have the impression that *point scale* is the commonly understood term when referring to the evaluation of snowpack stability. We will provide or reference a definition of the *point scale* and incorporate it into the concept of *frequency [distribution] of snowpack stability*. Avalanche danger can be assessed across different scales. Regional public avalanche warnings are typically provided for areas of 100-300 km2 and larger. Site-specific warnings are generally issued for one or several avalanche paths, e.g., along a road section. Snowpack stability, however, is assessed at a *single point* - usually by digging a snow pit to record a snow profile or conducting stability tests (e.g. ECT, Rutschblock). The extent of this assessment is typically no more than about 2x2 m, which we define as the point scale. Several studies validating avalanche danger levels rely on the frequency distribution of snowpack stability at the point scale (e.g., Schweizer et al., 2003, Bakermans et al., 2010). In theory, the EAWS workflow requires estimating stability assessments at all points within a warning region. A higher frequency of unstable points corresponds to a higher likelihood of triggering avalanches. This frequency distribution of stability classes (representing likelihood), combined with the potential avalanche size, determines the danger level for the region — a concept similar to the hazard chart in the CMAH. In practice it is impossible to assess the snowpack stability at all points within a region of several 100 km2. Forecasters must extrapolate and *imagine* the frequency distribution of unstable points based on available observations and models. While the CMAH uses the *spatial distribution of avalanche problems* to assess danger, an avalanche problem does not necessarily correspond to a fixed stability class in space or time. Therefore, the presence of an avalanche problem is not directly equivalent to the snowpack stability at a specific location.

Frequency of snowpack stability

This concept is awkward, and the method is mismatched with actual avalanche forecasting as is done today. This is in-part due to the issues described above, where snow stability is not well defined as a point scale observation and without that, the concept of frequency of snow stability is hard to grasp. Frequency of snowpack stability is a quantitative method using the "frequency distribution of points of snowpack stability" as a statistical measure of probability [Line 185: f (i) = ni /n]. This method assumes one knows both ni and n, both of which are impossible to know without snowpack modelling that is distributed across the landscape. This method is well suited to compute a danger rating but requires snowpack modelling to supply the necessary data and seems distant from actual avalanche forecasting as practiced currently, which uses observational data and requires answering the question of "where" along with developing a mental image of how snow stability is distributed across the terrain. "Distribution" is a much better term for this concept in the English language and for avalanche forecasting specifically, where the question of how snow stability is distributed across the landscape is one of the most fundamental questions to be solved. Frequency of snowpack stability tries to partially solve this statistically, but the EAWS Matrix is being applied by forecasters using observational data. The f formula is sound, but it will better resonate with avalanche forecasters if the concept is termed "Distribution of Snowpack Stability". I suggest lines 181 – 186 be revised in such a way:

1. Snowpack stability (point scale)
2. Distribution of snowpack stability (estimated using Table 3, or calculated using frequency)
3. Avalanche size

We would like to begin by addressing the terminology. We agree that frequency distribution is more appropriate in this context than frequency alone. Accordingly, we will update the manuscript to use frequency distribution of snowpack stability when referring to the overall distribution across all stability classes and frequency when discussing a single class, such as the frequency of stability class poor. While frequency of snowpack stability is the current official term within the EAWS, we will clarify in the text that we have updated this terminology for precision and context.

We acknowledge that determining the exact frequency distribution of snowpack stability for a region — especially for areas larger than a single mountain slope — is practically impossible. However, it remains the ideal goal, and regional avalanche forecasters strive to assess it as realistically as possible. While modeling may not completely resolve this challenge, it offers the potential to provide more realistic spatial-temporal estimates than individual field observations alone. We believe that the professional framework for determining regional avalanche danger levels should reflect what theoretically needs to be assessed rather than obscuring these requirements. For this reason, we will retain the technical and statistical definition of frequency distribution and clarify that the sum of the frequencies for the four stability classes equals 1. Additionally, we will emphasize the gap between what we aim to measure and what is practically achievable, highlighting the irreducible uncertainties inherent in avalanche forecasting. In this context, clear guidelines on managing incomplete information and addressing resulting uncertainties, can help forecasters to tackle these challenges.

Lines 27-29: describe the preparation of regional avalanche forecasts as involving two steps: 1) assessment of current and future avalanche danger, and 2) communication of future avalanche danger. This is true for the assessment of avalanche danger, but not for the preparation of a regional avalanche forecast, which also includes (depending on region) analysis of snowpack, terrain, locations and avalanche problems. Avalanche danger is only one component of a regional avalanche forecast and this paper focusses on the assessment of avalanche danger ratings. Please revise this section to reflect a broader view of what preparing a regional avalanche forecast entails or narrow the scope of the preparation down to avalanche danger ratings only. This same issue is found on Line 67, where "regional avalanche forecasting" should be narrowed in scope to "the assessment of avalanche danger ratings" and Line 72 where "the process of regional avalanche forecasting" should also be narrowed in scope to "the standards for assessing avalanche danger ratings".

We re-phrased this section and narrowed the scope to assessing regional avalanche danger ratings.

Lines 47-49: describe the importance of reliable avalanche forecasts in reducing damage and loss, enhancing safety and mitigating risks. This is all correct but focuses solely on negative consequences and ignores positive outcomes, which is half of the risk equation and a major reason for public avalanche forecasts. It is not all about loss. In addition to preventing loss, avalanche forecasts enable backcountry experiences by highlighting good conditions and appropriate terrain choices, improving the experience for backcountry users. It is essential to not only on focus on mitigation and loss, but also on gain and enabling the backcountry experience. I would like to see this theme applied throughout this paper where appropriate.

We will add a reference to the positive effects of avalanche warnings.

Line 177: a better heading would be "Factors and workflow to determine avalanche danger levels"

Changed as suggested to "Factors and workflow to determine avalanche danger levels".

Tables 2,3 & 4: I realize these are EAWS standards and perhaps not easily changed, but I suggest you consider the following to make these crucial reference tables more applicable in practice.

These tables should provide enough reference to make assessments based on observational evidence.

Table 2 - In general I find this table insufficient for making a proper assessment of snow stability. EAWS 2022b Appendix A has much more detail and I suggest improving this paper's Table 2 by adding additional columns to aid in assessment (i.e., stability test results). – Make it clear that this is a point scale assessment. – Change the Description of Very Poor to "very easy to trigger (e.g., natural)" to be consistent

We added the reference to the full table in the caption, but only reprint the short version here for brevity. The updated caption is "Snowpack stability classes refering to the point scale, and the type of triggering typically associated with these classes. For the full table, including typical observations related to each class see EAWS (2022b, Table 1)". We changed the description as suggested to "very easy to trigger (e.g., natural)".

Table 3 – Consider showing the frequency equation f (i) = ni /n and add a column on the right showing typical ni /n values.

Within the working group, we discussed in depth whether typical values should be shown together with these definitions. In the end, we decided including some examples in the EAWS-document describing the Matrix (Appendix B in EAWS, 2022b). However, the results shown there are based on one or two services and from very confined regions. In addition the values diverge substantially between stability tests and avalanche activity. We therefore concluded that they may be more confusing than clarifying. Therefore, we prefer not to show exemplary frequency values in this table.

Table 4 – Rephrase Size 1 be more concise, "Unlikely to bury or injure a person except in terrain traps".

This is taken from a standard. We only recite it here and will not change it. But we will take this comment to the appropriate forum.

Line 430 – The SWI forecasters were instructed to assess the three contributing factors, but to then ignore the EAWS Matrix and this led to a 46% deviation from the EAWS Matrix. This is a substantial deviation and makes me wonder why. Please explain why there is such a large deviation. Also Line 433 says that "forecasters largely concurred with the danger level proposed by the EAWS Matrix", except for SWI, which is the only service not linked to the matrix, where they disagreed 46% of the time. I think this needs more explanation please.

In most of the analyzed warning services, forecasters utilized the EAWS Matrix to determine the danger level (DL), typically aligning their choice with the first DL suggested by the Matrix. Deviations from this initial DL were relatively rare in these services (see also Table 6). In Switzerland, however, forecasters assessed factors and DL independently, without directly referencing the EAWS Matrix to determine the DL. This approach is somewhat like the North American system, where factors (e.g., *likelihood* and *avalanche size*) are evaluated for each avalanche problem, but where the final DL decision remains with the forecaster as no clear assessment rules are defined (Clark and Haegeli, 2018). While Swiss forecasters frequently deviated from the first DL shown in the EAWS Matrix, absolute differences were generally minor as forecasters most frequently selected the sub-level closest to this DL (for description of sub-levels used in Switzerland refer to Techel et al. 2022). These findings are described in greater detail, analyzing the second season after introducing the Matrix, in Techel et al. (2024). We will provide these explanations in the Discussion section.

Line 443: This is just a comment, but the fact that "All warning services described stability most often as poor when giving "D=2" is interesting because typically at D2, avalanche conditions are generally stable except for a few locations, thus backcountry travel can be done in many locations. This scenario shows that forecasters are assessing this D2 based on the frequency of poor stability. Although this Matrix applies the concept of snow stability, the assessments of avalanche danger ratings are actually determined and driven by the frequency of instability. This links to my earlier comments about Lines 47-49 whereby this model is driven by negative consequences, which is a bias towards instability and fails to appreciate stable conditions. No change requested, just an observation which has followed me through this review and has made me wonder why stability is not instead referred to as instability, when this is what truly drives the D assessment?

This issue has been discussed within the Matrix workgroup and among EAWS members. The term stability works better in communication and direct speech, which to our knowledge also true for the North American avalanche community. We would like to stick to the term stability. However, we generally agree with your observation and concerns. We hope and agree that an avalanche forecast has positive effects as you mention. However, the main task of an avalanche forecaster is to identify the danger and where/when it is located.

Technical corrections:

Line 6: Changed to "promote" as suggested

Line 19: Changed to "...to improve the reliability, credibility and timeliness of avalanche forecasts,..."

Line 44: changed to "forecaster's" **Line 88:** changed to "Except" **Line 108:** changed to "frequency" **Line 131:** replaced reference a suggested

Line 158: fixed

Line 220: fixed **Line 298:** We added information on terrain: "The workflow is specifically designed for regional avalanche forecasters. It assumes that the forecast area is large enough to encompass multiple mountains, elevation zones, all aspects, and varied terrain features, such as ridges, gullies, and open slopes. Consequently, terrain is not treated as an independent factor."

Figure 3 caption: We replaced "the most vulnerable locations" by "the lowest stability class".

Line 306: Rephrased to "...we aggregated bordering warning services to groups of services when using the same workflow and forecasting software."

Line 561: We replaced "imminent" by "crucial". **Line 635:** Changed as suggested.

References

Clark, T. and Haegeli, P.; ESTABLISHING THE LINK BETWEEN THE CONCEPTUAL MODEL OF AVA LANCHE HAZARD AND THE NORTH AMERICAN PUBLIC AVALANCHE DANGER SCALE:

INITIAL EXPLORATIONS FROM CANADA; Proceedings, International Snow Science Workshop, Innsbruck, Austria, 2018

EAWS: Determination of the avalanche danger level in regional avalanche forecasting, European Avalanche Warning Services, https://www.690 avalanches.org/wp-content/uploads/2022/12/EAWS_matrix_definitions_EN.pdf, last access: 21 June 2023, 2022b.

Techel, F., Mayer, S., Pérez-Guillén, C., Schmudlach, G., and Winkler, K.: On the correlation between a sub-level qualifier refining the danger level with observations and models relating to the contributing factors of avalanche danger, Natural Hazards and Earth System Sciences, pp. 1911–1930, https://doi.org/10.5194/nhess-22-1911-2022, 2022

Techel, F., Lucas, C., Müller, K., Pielmeier, C., and Morreau, M.; UNRELIABILITY IN EXPERT ESTIMATES OF FACTORS DETERMINING AVALANCHE DANGER AND IMPACT ON DANGER LEVEL ESTIMATION USING THE MATRIX; Proceedings, International Snow Science Workshop, Tromsø, Norway, 2024

Schweizer, J., Kronholm, K., Wiesinger, T.; Verification of regional snowpack stability and avalanche danger; Cold Regions Science and Technology, Volume 37, Issue 3, Pages 277-288, https://doi.org/10.1016/S0165-232X(03)00070-3; 2003.

Bakermans L, Jamieson B, Schweizer J, Haegeli P. Using stability tests and regional avalanche danger to estimate the local avalanche danger. Annals of Glaciology. 2010;51(54):176-186. doi:10.3189/172756410791386616