Review of Müller, K., Techel, F., and Mitterer, C.: An updated EAWS matrix to determine the avalanche danger level: derivation, usage, and consistency, Nat. Hazards Earth Syst. Sci. Discuss. [preprint], https://doi.org/10.5194/nhess-2024-48, in review, 2024.

By Grant Statham

General comments

This is an important paper that describes in detail the process used by the European Avalanche Warning Services (EAWS) to assess an avalanche danger rating, with the objective of adding more objective support to the assessment in order to improve consistency between forecasters and forecast regions. For years the Bavarian Matrix, then the EAWS Matrix were background documents used in Europe and hard to follow and understand. With this manuscript, it is now possible to closely study the EAWS method and thus the topic is relevant, an important contribution to the field and appropriate for publication in NHESS.

The method described is specifically for determining an avalanche danger rating using the European Avalanche Danger Scale (EADS), an ordinal scale of 1-5. Determining a danger rating is one specific niche of avalanche forecasting (public avalanche forecasting), and does not apply to the broader scope of avalanche forecasting which includes avalanche forecasting for ski areas, transportation, infrastructure, guided recreation, etc. It is important to recognize this specific focus of the EAWS matrix and that this method will not easily apply to other types of avalanche forecasting as it is independent of terrain, and terrain is fundamental to avalanche forecasting. This EAWS Matrix assesses avalanche danger ratings, not avalanches themselves.

The EAWS Matrix is intended to bring consistency to the process, which is necessary as currently there is inconsistency in the application of danger ratings between forecasters and forecast regions. Avalanche forecasting is done from observational evidence and has little to no quantitative support, thus the process is mostly judgement based. This paper implements the quantitative concept of “frequency of snow stability” which has been evolving over the past decade and is an excellent step forward that advances the challenge of assessing the probability of avalanche release.

As a general comment, I find the EAWS Matrix method a bit awkward and not intuitive to follow in practice. It seems that the method of frequency of snow stability is well suited to a future state of using snowpack models in order to model avalanche danger ratings – which is an excellent and necessary step. But in terms of using this method in day-to-day avalanche forecasting and applying observational evidence it seems the flow is awkward as shown in the manuscript. I think this is because the focus is on determining a danger rating between 1-5, rather than understanding the distribution of avalanche hazard across the landscape. It certainly works for improving consistency by having all users follow the same process to land on the same danger rating, and this is important and necessary.

Specific 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.

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: $f(i) = n_i/n$]. This method assumes one knows both n_i and n, both of which are impossible to know without snowpack modelling that is <u>distributed across the landscape</u>. 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

This is consistent with using observational data, which is incompatible with the frequency formula. When the data is available, distribution can be calculated as a frequency using $f(i) = n_i/n$.

At the very least, using the full term "frequency distribution" will help as is how the concept is termed in Techel et al., 2020.

Lines 616-623 describe a future state where subjectivity is reduced in avalanche forecasting through the increased use of highly resolved models, which is an excellent objective but still under development. When the use of snowpack modelling becomes implemented operationally, then then $f(i) = n_i/n$ can be properly utilized.

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".

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.

Line 117 a better heading would be "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

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

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

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.

Line 443 – This is just a comment, but the fact that "All warning services described stability most often as *poor* when giving D=2" in 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?

Technical corrections

Line 6 – replace *"ensure"* with "promote" or "improve" as these look-up tables <u>will not</u> ensure or guarantee consistency.

Line 19 – rephrase "*further efforts are required to develop and implement regional avalanche forecasting standards to reach the goal of forecasts being a reliable, credible, and timely source of information of expected avalanche conditions…"*. This reads like the goal of being reliable, credible and timely has not yet been achieved, which is incorrect. Consider "*further efforts are required to develop and implement regional avalanche forecasting standards to improve the reliability, credibility and timeliness of avalanche forecasts, regardless of forecaster or warning service behind the product".*

Line 44 – "*forecasters"* should be possessive (i.e.: forecaster's).

Line 88 – replace *"Despite"* with "Except".

Line 108 – spatial distribution should be "frequency"

Line 131 – Replace reference to Statham et al., 2010 with: https://arc.lib.montana.edu/snow-science/objects/issw-1996-060-066.pdf

Line 158 – typo "allowed <u>forecasters</u> to adjust…"

Line 220 – delete "In case that avalanches release" and just say "Avalanches can reach up to *size 3*.

Line 279 – add "ratings" to the references to avalanche danger. This matrix determines avalanche danger <u>ratings</u>, not avalanche danger.

Line 298 – This workflow description appears to be independent of terrain. Please indicated where is terrain (location) is assessed and how does this fit into the workflow?

Figure 3 – This is just a comment and not a request to change anything, but I did find the matrix easier to use when I inserted the avalanche activity examples from the stability classes shown in Table 2. The figure below shows what I mean (red text). When I make an assessment, and consider the stability class as a point, then having my estimate of what could trigger an avalanche at that point is helpful and easy to cross reference with the terms many, some, a few, none. Simple, but it makes the assessment more intuitive and easier for me.

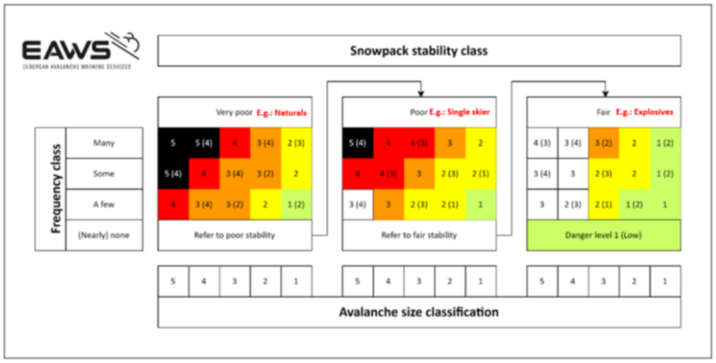

Figure 3 caption – "first assess the frequency of the most vulnerable locations". The most vulnerable locations is not consistent Line 285 that says "they begin by considering the lowest stability class". What do you mean by "the most vulnerable locations"? This is even harder to resolve when stability is defined as a point. This is also the only reference I have seen with regards to locations (i.e., terrain) and this model seems independent of terrain.

Line 306 – rephrase " and when spatially continuous" to better explain what this means.

Line 561 – change the word "imminent". Do you mean "important"?

Line 635 – change the start of the sentence to "Public avalanche forecasting …" as some kinds of avalanche forecasting do not involve public communication.

Conclusion

Thanks for the opportunity to review this manuscript. I think it's an excellent development and advances some of the concepts that the public avalanche forecasting community have been struggling with. Most importantly I am encouraged by the implementation of snow stability as a point measurement, which leads to the concept of frequency distribution as a method to determine avalanche release probability. Probability of avalanche release is a vexing problem to solve, and this work positions the concept well for a future of determining avalanche danger ratings using models. This is an important development.

Sincerely,

Grant Statham
June 24, 2024