Review of *An updated EAWS matrix to determine the avalanche danger level: derivation, usage, and consistency* by Müller et al. (nhess-2024-48)

General Comments

In this study, the authors provided a thorough background on the European Avalanche Danger Scale (EADS) and European Avalanche Warning Services (EAWS) Matrix development and definitions, described the revised EAWS Matrix and associated methods of updating the Matrix, and then presented relevant results of using the newly revised EAWS Matrix for one full season. This manuscript is well written and logically organized. The methods are sound, and the results are supported by sufficient evidence. The interpretation of those results is reasonable with reference to existing literature. The recommendations section of this paper is very useful and points to limitations of consistency by providing solutions. Overall, I think this is a valuable contribution to the literature, fit well within this Special Issue, and should be published. I have a couple of general comments for the authors to consider and a few specific and technical comments as well.

I understand that this study focuses on evaluating the updated (current) EAWS matrix, its use and consistency among forecasters, and compatibility with EADS, and not necessarily an evaluation of the individual three key factors that determine the danger rating. As the authors point out, consistency in forecasters' evaluation or interpretation of the three key factors is crucial for the matrix to be used to its full potential. Indeed, the authors provide recommendations on how to enhance the use of the matrix by improving consistency in the three contributing factors. However, it seems that across the surveyed regions, a variety of input data (in-situ observations, model output, meteorological data, etc.) exist. The authors mention the influence of input data very briefly (lines 485-489), but the quality and quantity of input data plays a crucial role in danger assessment and would potentially influence the assessment variability across forecasters and regions. Can the authors provide information or comment on, generally, what data types each region uses and/or provide evidence on how the assimilation process of various types of data across regions may influence the classification of the three factors and ultimately the danger rating?

The authors present a thorough summary on the evolution of EAWS and the EADS and provide some geographic references to other non-European forecasting tools like the CMAH in North America. However, there are no references to other avalanche sector decision making tools. In other words, this study focuses on public avalanche forecasting operations, but not forecasting in other sectors like transportation corridors, ski areas, natural resource industry, etc. Is the same tool used for those sectors throughout Europe? I suggest being clear that this study focuses on a matrix for public avalanche forecasting or state how using the EAWS matrix in those sectors differs, if at all, from public avalanche forecasting.

Specific and Technical Comments

Figure 2: Consider adding a legend to the proportion scales for b) and c) that easily shows the reader which colors represent higher correlation values.

Figure 6: The use of D^1 and D^2 here is confusing. Is this the same as the median D^1 and D^2 used in Figure 2a and defined in Lines 262 -264? Also, if there was disagreement with D (forecaster derived) and D^1 (Matrix derived), that is indicated in the left column in Figure 6, correct? If the second column represents $D^1 \neq D^2$, then forecasters used D^2 (again the median second selection from Figure 2a)? Please clarify.

Line 88: four? In "Despite for minor changes in 1994…" or do you mean "Except for minor changes…"

Line 104: the way this is written is confusing to me. I read it as 'as stability decreases you need a greater load to trigger an avalanche.' (i.e. inverse relationship). Perhaps 'instability' should be used here instead of stability.

Lines 241 and 244: 60 responses in total in line 241, but in line 244, you state 76 responses. What is the difference?

Line 258: Similar to the comment above (Line 104). I view an increase in stability as the snowpack becoming more stable. Instability?

Line 345: "center of gravity"? Do you mean largest proportion?