

We thank the reviewer for the careful assessment of our manuscript and the constructive comments. Our point-by-point responses are provided below. For clarity, the reviewer's comments are shown in blue italics, followed by our responses in black.

General comments

The manuscript describes a study where the avalanche probability is forecast based on three variables. The study area is the Galongla Valley, which has a snow climate that was characterised as maritime. In maritime snow climates, avalanche activity is predominantly controlled by precipitation and snowmelt. The data used for the study are daily weather and snow data (from 10 monitoring stations) and avalanche records (13 dry-snow avalanche events, 24 wet-snow avalanche events) covering two winter seasons. The predictor variables were snowfall intensity, snowfall duration, and snow depth prior to the snowfall period. For dry-snow and wet-snow avalanche events, corresponding non-events were selected, three times as many as events. With these data, a logistic regression model was developed. The model performed well. The authors compare model performance with and without using the snow depth as input, and report an improvement in model performance when using snow depth.

Overall, the study is mostly well described; results are easy to follow and are adequately discussed. The scientific significance, however, is rather modest. It is hard to identify a substantial contribution to the understanding of avalanche forecasting (new concepts, ideas, methods, or data). Still, the study may be helpful for local authorities responsible for road safety, but such application is not shown.

Response: We appreciate the reviewer's assessment and the important comment regarding the scientific contribution of the study. We have clarified that the contribution of this study lies in an observation-based assessment of dry- and wet-snow avalanche occurrence in a maritime mountain region with limited in-situ observations. We model the two avalanche types separately by incorporating the intensity and duration of the corresponding snowfall or snowmelt processes together with pre-event snow depth, and quantify the additional discriminatory information provided by snow depth through model comparison. In addition, the models revealed negative conditional associations between snowfall or snowmelt duration and avalanche occurrence in the study area, providing new regional observational evidence on the relationship between process duration and dry- and wet-snow avalanche occurrence. Avalanches in the study area pose a direct threat to local transportation routes. Although the present study does not demonstrate a road-safety application, its findings provide a quantitative basis for further assessment of avalanche conditions relevant to local road management.

Specific comments

Background snow depth

The term "background" snow depth is not particularly clear. I suggest omitting the qualifier "background" and explaining that you use the snow depth prior to the snowfall period, and then rather use a qualifier in that sense, e.g., "prior" snow depth (or "pre-event" snow depth). It is partly misleading to refer to snow depth as pre-event snowpack conditions. I suggest replacing such terms with simply snow depth, since no information on the quality of the snowpack (stratigraphy) is provided.

The importance of snow depth prior to snowfall has been explained in previous studies (Schweizer et al., 2009) that a certain amount of snow on the ground is needed for large avalanches to reach infrastructure. If there is not enough snow prior to the snowfall, the terrain is very high and often prevents far-reaching avalanches. This seems the most relevant reason, in particular in a region with a maritime climate, where the presence of persistent weak layers in the snowpack is typically not a major concern.

By the way, the snow depth prior to the snowfall event has traditionally been considered (e.g., Schweizer and Föhn, 1996; Schweizer et al., 2009; Stoffel et al., 1998). However, in these studies, the snow depth 3 days before the actual day was used, since also the 3-day-sum of new snow height was considered (in addition to new snow height of 24 hours).

In this context, it would be interesting to see statistics of the variable snowfall duration. This might indicate that it is advantageous to use the effective snowfall duration rather than the often used 3-day period.

Response: We thank the reviewer for this helpful clarification. We agree that "background snow depth" may be ambiguous and that snow depth alone should not be interpreted as information on snowpack stratigraphy or stability. We therefore use "pre-event snow depth" consistently for this variable and have revised expressions that previously referred more broadly to pre-event snowpack conditions. We have also incorporated the references suggested by the reviewer.

Regarding the snowfall-duration statistics, the median duration for the dry-snow avalanche samples was 4 h (Q1–Q3: 2–8 h; range: 1–15 h). These results show that the identified snowfall processes varied in duration

and were considerably shorter than the conventional 3-day period. This supports the use of the identified snowfall-process duration as a more process-specific temporal descriptor than a fixed 3-day period.

Technical corrections

Lines 59-60: As far as I understand, you refer to spatial variability of the snowpack. If so, the reference to Monti et al. (2016) does not make sense to me. I suggest referring to, for instance, a review paper on spatial variability.

Response: Thank you for pointing this out. We agree that Monti et al. (2016) does not directly support the statement on the spatial variability of snowpack conditions. We have therefore replaced this citation with the review by Schweizer et al. (2008), which specifically addresses the spatial variability of snowpack properties and its relevance to avalanche formation.

Lines 64-65: The value of snowpack information (stratigraphy, stability) for avalanche forecasting has been known for a long time, conceptually (LaChapelle, 1980), but also for forecasting models (Schweizer and Föhn, 1996).

Response: Thank you for pointing this out. We agree that the importance of snowpack information in avalanche forecasting has long been recognized and should not be presented as a recent development. We have therefore revised this part to acknowledge the earlier conceptual work of LaChapelle (1980) and the forecasting-model application of Schweizer and Föhn (1996), while retaining the more recent study as an example of subsequent developments.

Line 80: I suggest changing the mean slope angle to 34°.

Response: Thank you. We have changed the reported mean slope angle to 34°.

Line 81: ... and the adjacent transportation corridor.

Response: Thank you. We have added the definite article.

Line 99: avalanche release areas

Response: Thank you. We have used “avalanche release areas” as suggested.

Line 102: Are several transportation corridors threatened? If so, use plural also above in line 81.

Response: Yes. Several transportation corridors are threatened by avalanche activity in the study area. We have therefore changed “adjacent transportation corridor” in Line 81 to the plural “adjacent transportation corridors” for consistency.

Figure 2b I suggest also labeling the instrumentation in the lower right corner of the monitoring station.

Response: Thank you for the suggestion. We have added a label to the instrumentation in the lower-right corner of Fig. 2b.

Line 114-115: What is the temporal resolution of the predictor variables? Days? Hours?

Response: Thank you for requesting this clarification. The predictor variables were calculated from hourly records. We have clarified this in the predictor-variable description.

Line 116: For clarification, I suggest replacing truncated with terminated or cut off, and adding snowfall to event.

Response: Thank you for the clarification. We have revised the sentence accordingly.

Line 117: I suggest using lower fonts (not capitalized), even though the abbreviation is capitalized. This is the common practice in NHESS. Here, and elsewhere in the manuscript. <https://www.natural-hazards-and-earth-system-sciences.net/submission.html#english>

Response: Thank you. We have changed the full variable names to lower case throughout the manuscript while retaining the capitalized abbreviations (SFD, SFI, SMD, SMI, and SD), following NHESS style.

Figure 3: deposit instead of deposition, as well as in Figure 4. Figures 3 and 4 are nice but also very basic.

Response: Thank you. We have replaced “deposition” with “deposit” consistently in Figs. 3 and 4 as suggested.

Lines 143-154: This paragraph is rather hard to follow, but it is important to understand how the negative samples were selected. Maybe an illustration showing the snow depth and the avalanche event, and then indicate the time block of the positive event, and any negative event blocks.

Response: Thank you for pointing out that the sample-selection procedure requires a clearer explanation. We have revised this section to clarify the retrospective selection procedure. For each recorded avalanche, candidate snowfall or snowmelt process blocks were identified within the preceding 3-day window. The most recent qualifying process before the avalanche occurrence was selected as the positive sample, while earlier qualifying process blocks were considered candidate negative samples. Candidate negative blocks ending less

than 12 h before the start of the positive block were excluded. From the remaining candidates, up to three temporally separated negative samples were retained. We have also added a time-axis illustration based on an actual example to show the avalanche occurrence time, the positive sample block, the retained negative blocks, the 3-day retrospective window, and the 12 h exclusion interval.

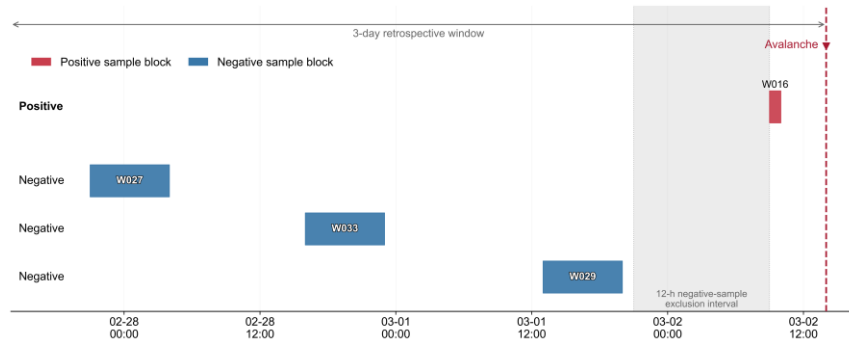


Figure R1. Schematic of positive and negative sample selection. The red and blue blocks represent the positive and retained negative samples, respectively. The red dashed line marks the avalanche occurrence time. The arrow indicates the 3-day retrospective window, and the grey shading denotes the 12 h exclusion interval.

Finally, it is not clear to me whether the predictor variables were selected from the monitoring station closest to the avalanche event, or whether another approach was selected.

Response: Thank you for requesting this clarification. For each recorded avalanche, the predictor variables were derived from observations at the geographically closest monitoring station.

Line 155: Here or in the Discussion, I suggest referring to previous forecasting models employing logistic regression, e.g., Gauthier et al. (2017).

Response: Thank you for this suggestion. We have included a reference to previous avalanche forecasting models using logistic regression.

Line 158: I find it somewhat misleading to refer to “meteorological and snowpack predictors” but only two meteorological variables and one snowpack variable were used. I suggest explicitly stating that you use three predictor variables.

Response: Thank you for pointing this out. We have revised the description to state explicitly that SMLM uses three predictor variables, whereas MOLM uses two.

Line 175: In avalanche forecasting where the history of the winter is important, one usually does not use leave-one-out cross-validation but rather tests one season against the other. I agree that this is difficult with only two winters of data, but doing so would still indicate a realistic alternative level of performance. This could also be discussed in the Discussion chapter.

Response: Thank you for this helpful suggestion. We conducted an additional exploratory cross-season validation in chronological order, using the 2024 season for model fitting and the subsequent 2025 season for validation. For threshold-dependent metrics, the classification threshold was determined using only the 2024 training data and was then applied to the 2025 validation data. For dry-snow avalanches, the AUC values were 0.74 for MOLM and 0.75 for SMLM, with a TSS of 0.37 for both models. For wet-snow avalanches, the corresponding AUC values were 0.82 and 0.86, and the TSS values were 0.56 and 0.58, respectively. Thus, both model formulations retained discriminatory ability when evaluated on the subsequent snow season. We have also added these results to the Discussion. Given that the observations cover only two snow seasons, we interpret them as exploratory evidence of temporal transferability between the two observed seasons, without extrapolating to long-term forecasting performance.

Lines 185ff: If models are trained on unbalanced datasets, some metrics are biased by the abundance of non-events, including the F1 score. One metric which is known to be insensitive to this bias is the true skill statistic (TSS), also known as Peirce skill score (PSS), or Hanssen–Kuipers discriminant (HKD), see, e.g., Ebert and Milne (2022).

Response: Thank you for this helpful suggestion. We have added TSS to the model evaluation to complement the existing performance metrics. For dry-snow avalanches, the TSS values were 0.51 for MOLM and 0.56 for SMLM; for wet-snow avalanches, the corresponding values were 0.58 and 0.79.

Line 205: I suggest replacing increased with were.

Response: Thank you. We have replaced “increased” with “were” as suggested.

Figure 5: Please use the international day format (1 January 2024, or 1 Jan) as specified in <https://www.natural-hazards-and-earth-system-sciences.net/submission.html#english>; it is also specified there that in axis labels only the first word is capitalized.

Response: Thank you. We have revised the figure to use the day–month date format and sentence-style capitalization for the axis labels, as suggested.

Line 246: I suggest omitting basal.

Response: Thank you. We have omitted “basal”.

Figure 7: The use of the rainbow color bar is strongly discouraged for data visualization in science.

Response: Thank you for this suggestion. We have revised the figure accordingly.

Line 279: It would be interesting, as mentioned before, to have graphs with summary statistics for the variables, maybe separated for events and non-events.

Response: Thank you for this suggestion. We have added Figure R2 showing the distributions of the predictor variables separately for avalanche events and non-events. The figure presents SFI, SFD, and SD for dry-snow avalanches, and SMI, SMD, and SD for wet-snow avalanches, allowing the differences in predictor distributions between the two sample groups to be compared directly.

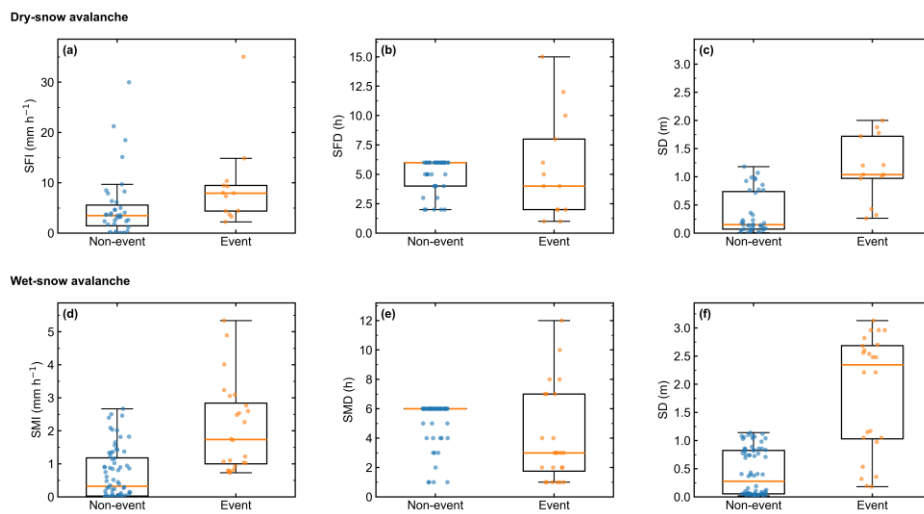


Figure R2. Distributions of predictor variables for avalanche events and non-events. Panels (a–c) show SFI, SFD, and SD for dry-snow avalanches, and panels (d–f) show SMI, SMD, and SD for wet-snow avalanches. Boxes indicate the interquartile range, horizontal lines indicate the median, whiskers extend to 1.5 times the interquartile range, and points represent individual observations.

Lines 304-305: I suggest replacing information on snowpack conditions with (pre-event, or prior) snow depth.

Response: Thank you. We have replaced “information on snowpack conditions” with “pre-event snow depth” as suggested.

Lines 326-328: I doubt that the explanation given here is very relevant. A deeper snowpack can only increase stress if weak layers exist within the snowpack prior to the snowfall event. However, this is not typical in mountain ranges with a maritime climate. In maritime snow climates, the failure most often occurs at the interface between the new snow and the old snowpack.

Response: Thank you for this comment. We agree that our original interpretation placed too much emphasis on loading of pre-existing weak layers. We have revised this discussion accordingly. The positive association between pre-event snow depth and avalanche occurrence remains an empirical finding. However, because our observations do not resolve snow stratigraphy or the actual failure interface, we do not attribute this association to a specific failure mechanism.

Line 336: I suggest deleting “newly accumulated snow” since it is redundant, and not sure whether McClung (2021) is the most appropriate reference for this rather basic statement.

Response: Thank you. We have removed the redundant phrase “newly accumulated snow” and revised the citation.

Line 340: I would expect SFD and SFI to be correlated. I suggest providing the details.

Response: Thank you for this careful observation. We examined the relationship between SFD and SFI in the final dry-snow dataset. Across all 52 dry-snow samples, the Pearson correlation was $r = -0.03$ ($p = 0.83$), and the Spearman rank correlation was $\rho = 0.06$ ($p = 0.66$). Both results indicate little evidence of either a linear or monotonic association between SFD and SFI.

Line 360: ... when the pre-event snow depth is sufficiently deep.

Response: Thank you. We have revised the relevant passage to refer explicitly to pre-event snow depth.

Line 361: I don't know how you demonstrated the complex coupling between meteorological forcing and snowpack conditions given that you used three predictor variables.

Response: Thank you for this clarification. We agree that the term “complex coupling” was too strong given the three predictor variables used in the model. We have revised the text.

Figure 10: I suggest changing the color bar.

Response: Thank you. We have revised the figure accordingly.

Line 424: What do you mean with “integrated”?

Response: Thank you. By “integrated”, we meant that each type-specific model combines pre-event snow depth with the corresponding snowfall- or snowmelt-process variables. We have revised the text to make this clearer.

Line 426: What does "state-constrained framework" mean?

Response: Thank you. By “state-constrained framework”, we meant that snowfall- or snowmelt-process variables were considered together with observed pre-event snow depth in each model. We have replaced this term with a clearer description.

Line 432-433: I suggest rewording: These gains were primarily associated with the incorporation of pre-event snow depth into the forecasting model. Compared with baseline models that excluded snow depth...

Response: Thank you. We have revised the text in line with this suggestion.

Assets / datasets

I have finally looked at the assets; thank you for providing those. I am surprised to see values of snow depth with 6 significant digits such as: 1.045633.

Response: Thank you for pointing this out. The additional decimal places in the snow-depth values resulted from numerical processing of the observations and do not represent the precision of the underlying measurements. We have therefore rounded the reported snow-depth values in the revised dataset to avoid implying unwarranted precision.