Dear Editor,

We thank you and both referees for the thoughful assessment of our manuscript and for their constructive comments during the interactive discussion. Their feedback has helped us sharpen the focus of the study, clarify key methodological choices, and better align our interpretation of the findings with the available observational evidence.

In the revised manuscript, we place greater emphasis on the value of local observations from Galongla Valley and on the additional information provided by pre-event snow depth when combined with snowfall- or snowmelt-process variables. We have clarified the construction of positive and negative samples, refined the interpretation of the logistic regression outputs, expanded the presentation of model performance and uncertainty, and revised the Discussion to better distinguish statistical associations from possible physical interpretations. We have also clarified the limitations of the retrospective evaluation and refined the discussion of potential local applications.

Detailed responses to each referee's comments are provided below, together with references to the relevant revisions in the manuscript. Reviewer comments are shown in black italics, our responses in blue, and quoted passages from the revised manuscript in red.

Referee 1

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 thank the reviewer for this assessment. In the revised manuscript, we have highlighted the scientific value of a new integrated observational dataset from Galongla Valley, where avalanche activity threatens transportation routes and local observational evidence remains limited. Using these observations, we examine the relationships between snowfall- or snowmelt-process variables, pre-event snow depth, and recorded dry- and wet-snow avalanche occurrence.

By comparing type-specific logistic regression models with and without pre-event snow depth, we quantify its additional contribution to distinguishing avalanche-event from non-event conditions. The models also suggest negative conditional associations between forcing duration and classification scores. Together, these findings provide new local evidence on how pre-event snow depth and forcing characteristics are associated with dry- and wet-snow avalanche occurrence.

We have revised the Introduction, Discussion, and Conclusions to better explain these findings and their relevance to avalanche assessment in Galongla Valley. In particular, we emphasise the potential value of an observation-based approach for the further development of locally adapted monitoring and early-warning methods that could support avalanche assessment along threatened transportation corridors.

Changes in the revised manuscript:

In the Introduction, we have clarified the study objectives and the contribution of the local observational dataset and type-specific model comparisons. More details are shown in lines 64-71 of the revised manuscript:

"Focusing on an avalanche-prone high mountain valley in the southeastern Tibetan Plateau, this study compiles an event database using a dedicated in situ monitoring network that integrates meteorological, snow depth, and avalanche observations. Pre-event snow depth is defined as the snow depth at the onset of the preceding snowfall or snowmelt process. Separate logistic regression models are developed for dry- and wet-snow avalanches to quantify the relationships between forcing intensity, forcing duration, pre-event snow depth, and avalanche occurrence. These models are compared with meteorology-only baselines to assess whether the inclusion of pre-event snow depth improves event discrimination and reduces false-alarm rates. By incorporating observed pre-event snow depth into type-specific logistic regression models, this study provides an interpretable quantitative basis for regional avalanche monitoring and forecasting." (Introduction, lines 64–71)

In Sect. 4.3, we explain how the continuous classification scores describe combinations of local environmental conditions and can inform the development of candidate early-warning criteria. More details are shown in lines 380–387 of the revised manuscript:

"The SMLM provides continuous classification scores based on pre-event snow depth and the intensity and duration of snowfall or snowmelt processes. These scores allow the model response to different combinations of predictors to be examined beyond a single binary classification threshold. For illustrative purposes, the model-derived scores were divided into five equal-width intervals (Fig. 11). The resulting three-dimensional patterns illustrate how model scores vary with meteorological forcing and pre-event snow depth, and how different combinations of predictors can produce similar scores. Higher scores indicate predictor combinations more strongly associated with recorded avalanche events within the sampled data. These patterns provide an interpretable quantitative basis for identifying conditions associated with avalanche occurrence and informing the development of locally adapted avalanche early-warning criteria." (Sect. 4.3, lines 380–387)

The Conclusions now summarise the observed relationships, quantified model gains, and implications for local monitoring and further early-warning development. More details are shown in lines 406–419 of the revised manuscript:

"In this study, we developed separate logistic regression models for dry- and wet-snow avalanches in Galongla Valley, southeastern Tibetan Plateau, combining pre-event snow depth with snowfall- or snowmelt-process variables. Using observations from local monitoring stations during the 2024 and 2025 snow seasons, we examined avalanche-event and non-event conditions separately for the two avalanche types to account for their distinct meteorological forcing processes. The findings show that pre-event snow depth (SD) and forcing intensity (SFI/SMI) were positively associated with the model-derived relative likelihood of avalanche occurrence, whereas forcing duration (SFD/SMD) showed negative associations when intensity and SD were held constant." (Conclusions, lines 406–412)

"Incorporating SD improved model discrimination for both avalanche types under LOOCV. Compared with the MOLM, the SMLM achieved a TPR of 0.92 for both dry- and wet-snow avalanches at the selected classification thresholds, increased the AUC from 0.72 to 0.83 for dry-snow avalanches and from 0.85 to 0.94 for wet-snow avalanches, and reduced the FPR from 0.41 to 0.36 and from 0.21 to 0.13, respectively. These results show that incorporating pre-event snow depth alongside snowfall- and snowmelt-process variables improves the identification of avalanche-event conditions for both avalanche types. Based on local observations from a region where avalanche studies remain limited, our findings provide a practical basis for refining avalanche monitoring and informing the development of locally adapted early-warning methods in Galongla Valley." (Conclusions, lines 413–419)

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 referee for these helpful suggestions. We agree that "background snow depth" may be ambiguous and that snow depth alone cannot characterise snowpack stratigraphy or stability. We have therefore consistently adopted the term "pre-event snow depth" (SD) and clarified that it refers to the snow depth measured at the onset of the preceding snowfall or snowmelt process. We have also revised expressions that previously described SD more broadly as representing snowpack conditions.

Our results show that greater pre-event snow depth is associated with a higher model-derived relative likelihood of both dry- and wet-snow avalanche occurrence, even after accounting for forcing intensity and duration. This finding is consistent with previous research suggesting that sufficient antecedent snow cover can reduce the influence of terrain roughness on avalanche formation (Schweizer et al., 2003a). The relevance of antecedent snow depth has also been recognised in studies of large avalanches affecting transportation corridors (Schweizer et al., 2009). Our findings extend these observations to Galongla Valley, demonstrating that directly measured pre-event snow depth provides useful information for distinguishing locally recorded avalanche-event and non-event conditions when combined with snowfall- or snowmelt-process variables.

We have further revised the Introduction to acknowledge earlier avalanche-forecasting research, including Schweizer and Föhn (1996). Regarding snowfall duration, the 13 dry-snow avalanche-event samples had a median SFD of 4 h (Q1–Q3: 2–8 h; range: 1–15 h). These results show that the identified snowfall processes had variable durations, all substantially shorter than the conventional 3-day period. By characterising the duration and mean intensity of each identified snowfall process separately, our approach retains information on the temporal characteristics of snowfall. The negative conditional association between SFD and avalanche occurrence further indicates that longer snowfall duration does not necessarily correspond to a higher model-derived likelihood of avalanche occurrence when snowfall intensity and pre-event snow depth are held constant. This highlights the importance of considering the characteristics of individual snowfall processes.

Changes in the revised manuscript:

In the Introduction, we acknowledge the established role of snowpack information in forecasting and define the specific snow-depth quantity used here. More details are shown in lines 59–62 and 66 of the revised manuscript:

> "The importance of snowpack information for avalanche forecasting has long been recognised (LaChapelle, 1980), including in earlier forecasting models (Schweizer and Föhn, 1996). More recent studies have shown that combining meteorological observations with snowpack process simulations can improve avalanche prediction relative to meteorology-only approaches (Pérez-Guillén et al., 2022)." (Introduction, lines 59–62)
>
> "Pre-event snow depth is defined as the snow depth at the onset of the preceding snowfall or snowmelt process." (Introduction, line 66)

In Sect. 2.3, we specify the duration and intensity calculations and define SD at the onset of each identified process. More details are shown in lines 116–118 and 132–135 of the revised manuscript:

> "Using hourly records, snowfall duration (SFD) was calculated as the duration of this event segment, and snowfall intensity (SFI) as the corresponding cumulative snowfall divided by SFD, representing the mean rate of snow loading. In parallel, the snow depth recorded at the onset of the event was defined as pre-event snow depth (SD)." (Sect. 2.3, lines 116–118)
>
> "Snowmelt duration (SMD) was calculated from hourly records as the duration of this event segment, whereas snowmelt intensity (SMI) was calculated as the corresponding cumulative snow-depth loss divided by SMD. Similarly, the snow depth recorded at the onset of the event was defined as pre-event snow depth (SD)." (Sect. 2.3, lines 132–135)

In Sect. 4.1, we relate our findings on pre-event snow depth to previous research on terrain roughness and avalanches affecting transportation corridors, and clarify the conditional associations between forcing duration and avalanche occurrence. More details are shown in lines 333–340 and 353–356 of the revised manuscript:

> "Avalanche activity is influenced by the combined effects of meteorological forcing and antecedent snow cover under different terrain conditions. The model results show that SD, representing pre-event snow depth,

is positively associated with the model-derived relative likelihood of avalanche occurrence for both dry- and wet-snow avalanches. This finding is consistent with previous studies indicating that sufficient pre-existing snow cover can reduce the influence of terrain roughness on avalanche formation and affect the amount of new snow required for avalanche release (Schweizer et al., 2003a). Previous investigations of large avalanches affecting transportation corridors have also identified antecedent snow depth as a relevant factor in avalanche forecasting (Schweizer et al., 2009). These findings highlight the value of considering pre-event snow depth together with meteorological forcing when assessing recorded avalanche occurrence." (Sect. 4.1, lines 333–340)

"Our modelling results indicate that both SFD for dry-snow avalanches and SMD for wet-snow avalanches are negatively associated with the model-derived relative likelihood of avalanche occurrence when forcing intensity and pre-event snow depth are held constant." (Sect. 4.1, lines 353–356)

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.

Changes in the revised manuscript:

The Introduction now cites the review of spatial variability by Schweizer et al. (2008). More details are shown in lines 56–57 of the revised manuscript:

"This limitation arises because avalanche release depends not only on meteorological forcing but also on snowpack conditions, which vary substantially across space and time (Schweizer et al., 2003a, 2008)." (Introduction, lines 56–57)

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.

Changes in the revised manuscript:

The Introduction now explicitly acknowledges the earlier conceptual and modelling work. More details are shown in lines 59–62 of the revised manuscript:

"The importance of snowpack information for avalanche forecasting has long been recognised (LaChapelle, 1980), including in earlier forecasting models (Schweizer and Föhn, 1996). More recent studies have shown that combining meteorological observations with snowpack process simulations can improve avalanche prediction relative to meteorology-only approaches (Pérez-Guillén et al., 2022)." (Introduction, lines 59–62)

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

Response:

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

Changes in the revised manuscript:

Sect. 2.1 now reports the mean slope angle as 34°. More details are shown in lines 77–78 of the revised manuscript:

"The terrain is steep and deeply dissected, with a mean slope angle of approximately 34° and maximum values exceeding 70°." (Sect. 2.1, lines 77–78)

Line 81: … and the adjacent transportation corridor.

Response:

Thank you. We have revised the phrase to "adjacent transportation corridors" to reflect the multiple corridors in the study area.

Changes in the revised manuscript:

Sect. 2.1 now uses the plural form . More details are shown in lines 78–80 of the revised manuscript:

> "Numerous confined gullies connect high-elevation snow accumulation areas with the valley floor and adjacent transportation corridors, providing favourable topographic conditions for avalanche release and channelised downslope movement." (Sect. 2.1, lines 78–80)

Line 99: avalanche release areas

Response:

Thank you. We have used "avalanche release areas" as suggested.

Changes in the revised manuscript:

Sect. 2.2 now uses "avalanche release areas". More details are shown in lines 95–97 of the revised manuscript:

> "The stations are strategically located near the potential release zones of major avalanche-prone slopes at elevations of 3600–3700 m, enabling the effective capture of key meteorological forcing and snowpack evolution processes within avalanche release areas." (Sect. 2.2, lines 95–97)

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 "the adjacent transportation corridor" in Line 81 to the plural "adjacent transportation corridors" for consistency.

Changes in the revised manuscript:

In Sect. 2.1, the corresponding wording has been revised to use the plural form. More details are shown in lines 78–80 of the revised manuscript:

> "Numerous confined gullies connect high-elevation snow accumulation areas with the valley floor and adjacent transportation corridors, providing favourable topographic conditions for avalanche release and channelised downslope movement." (Sect. 2.1, lines 78–80)

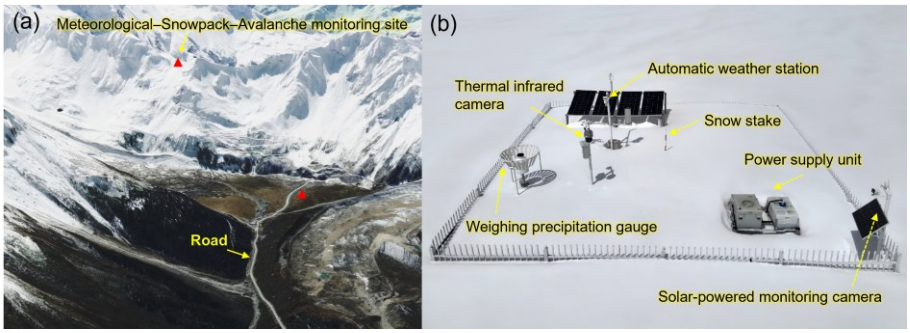
Figure 2b I suggest also labeling the instrumentation in the lower right corner of the monitoring station.

Response:

Thank you for the suggestion. We have added labels to the equipment in the lower-right part of Fig. 2b, identifying the power supply unit and the solar-powered monitoring camera.

Changes in the revised manuscript:

The revised Fig. 2 identifies both the power supply unit and the solar-powered monitoring camera in the lower-right part of Fig. 2b. The complete figure and its manuscript caption are reproduced here:



> "Figure 2. Integrated meteorological–snowpack–avalanche monitoring system in Galongla Valley. (a) Locations of representative monitoring stations; (b) Field configuration of the monitoring network." (Fig. 2, lines 104–105)

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

Response:

Thank you for requesting this clarification. The monitoring stations recorded meteorological and snow-depth variables at 1 min resolution. The predictor variables were calculated from hourly records derived from these observations.

Changes in the revised manuscript:

Sect. 2.2 now explicitly distinguishes the temporal resolution of the original observations from that of the records used to calculate the predictors. More details are shown in lines 97–100 of the revised manuscript:

> "Each station consists of an automatic weather station, a weighing precipitation gauge, a snow stake, and a thermal infrared camera (Fig. 2b), with meteorological and snow-depth variables recorded at a temporal resolution of 1 min. The predictor variables were calculated from hourly records derived from these observations." (Sect. 2.2, lines 97–100)

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

Response:

Thank you for the suggestion. We have replaced "truncated" with "cut off" and specified "snowfall event" to clarify the description of the preceding snowfall process.

Changes in the revised manuscript:

The revised process description in Sect. 2.3 specifies the snowfall event and uses "cut off" for an ongoing process. More details are shown in lines 114–115 of the revised manuscript:

> "For avalanche samples, the triggering process was defined as the last snowfall event preceding the recorded avalanche occurrence time (t_{event}), and any ongoing snowfall event was cut off at t_{event}." (Sect. 2.3, lines 114–115)

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.

Changes in the revised manuscript:

The revised variable definitions in Sect. 2.3 use lower-case full names while retaining the abbreviations. More details are shown in lines 116–117 and 132–134 of the revised manuscript:

> "Using hourly records, snowfall duration (SFD) was calculated as the duration of this event segment, and snowfall intensity (SFI) as the corresponding cumulative snowfall divided by SFD, representing the mean rate of snow loading." (Sect. 2.3, lines 116–117)

> "Snowmelt duration (SMD) was calculated from hourly records as the duration of this event segment, whereas snowmelt intensity (SMI) was calculated as the corresponding cumulative snow-depth loss divided by SMD." (Sect. 2.3, lines 132–134)

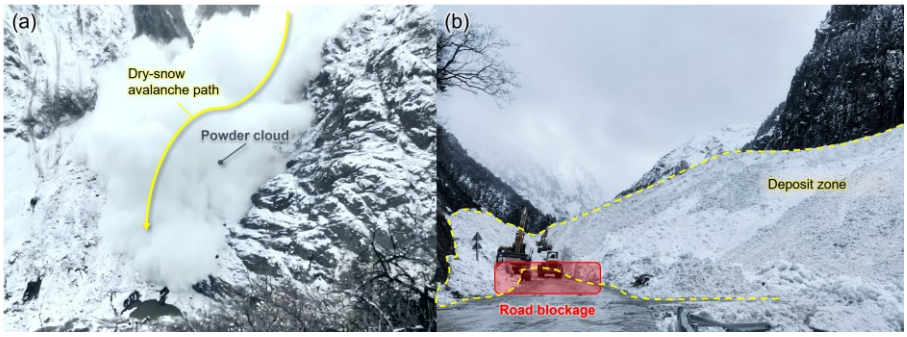
Figure 3: deposit instead of deposition, as well as in Figure 4. Figures 3 and 4 are nice but also very basic.

Response:

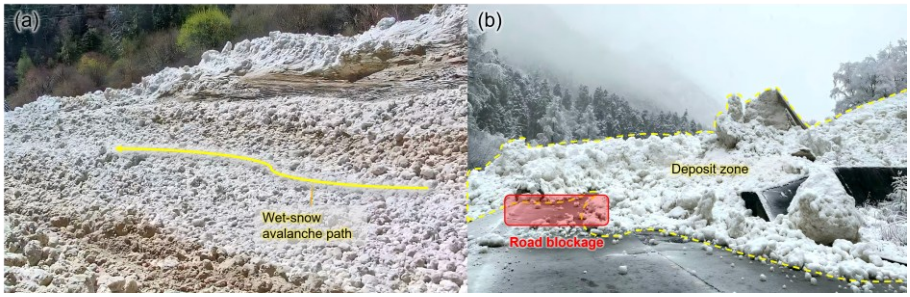
Thank you. We have replaced "deposition" with "deposit" or "deposits", as appropriate, in the captions of Figs. 3 and 4, and revised the labels in panels (b) from "Deposition zone" to "Deposit zone", as suggested.

Changes in the revised manuscript:

The revised captions now use "deposits" and "deposit", and the labels in Figs. 3b and 4b have been changed to "Deposit zone". The complete revised figures and their captions are reproduced below. More details are shown in lines 120–121 and 137–138 of the revised manuscript:



"Figure 3. Typical movement and deposits of dry-snow avalanches. (a) Dry-snow avalanche path with a powder cloud; (b) Avalanche deposit causing road blockage." (Fig. 3, lines 120–121)



"Figure 4. Typical movement and deposits of wet-snow avalanches. (a) Wet-snow avalanche path; (b) Avalanche deposit causing road blockage." (Fig. 4, lines 137–138)

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:

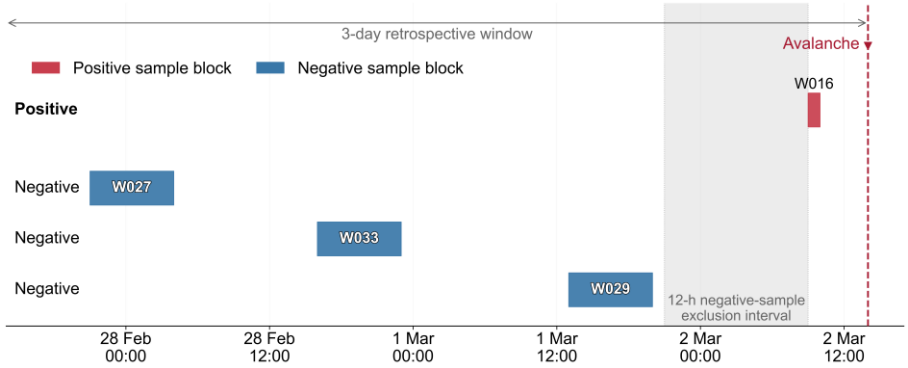
Thank you for pointing out that the sample-selection procedure requires a clearer explanation. We have revised this section to clarify the retrospective selection procedure. For each recorded avalanche, candidate snowfall or snowmelt-related process blocks were identified within the preceding 3-day window. 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 sample intervals were limited to a maximum duration of 6 h, even when the underlying processes lasted longer. Candidate negative blocks overlapping the 12 h immediately preceding the start of the positive block were excluded. Candidate blocks overlapping any known avalanche occurrence at the same station and of the same avalanche type were also excluded. From the remaining candidates, up to three temporally separated negative samples were retained. Each candidate process block contributed at most one retained sample, with an overall positive-to-negative ratio of 1:3 for each avalanche type. To make this procedure easier to follow, we have also added Fig. 5, a schematic time-axis illustration based on an observed wet-snow avalanche case, showing the retrospective window, the positive block, the retained negative blocks, the exclusion interval, and the recorded avalanche time. This revision makes the sample-selection procedure more transparent and helps clarify how the retained negative samples were selected while preserving comparability with the positive samples.

Changes in the revised manuscript:

Sect. 2.4 now sets out the sample construction explicitly, and Fig. 5 illustrates the selected blocks and exclusion interval. More details are shown in lines 140–154 and 156–158 of the revised manuscript:

"For each recorded dry- or wet-snow avalanche, candidate snowfall or snowmelt-related process blocks were identified within a 3-d retrospective window at the monitoring station geographically closest to the avalanche location using the same process-identification criteria described above. The most recent qualifying process block preceding the recorded avalanche time was treated as the positive sample, with all variables calculated only up to the avalanche occurrence time. Earlier process blocks within the same retrospective window were considered candidate negative samples. Candidate negative sample intervals were limited to a maximum duration of 6 h, even when the underlying processes lasted longer. Intervals overlapping the 12 h immediately preceding the start of the positive block were excluded, and eligible intervals closer to the positive block were prioritised. Candidate blocks overlapping any known avalanche occurrence at the same

station and of the same avalanche type were also removed using the full avalanche-time inventory. Eligible negative samples were retained to achieve an overall positive-to-negative sample ratio of 1:3 for each avalanche type. Each candidate process block contributed at most one retained sample, although multiple blocks could originate from the same broader snowfall or snowmelt episode. The predictor variables for positive and negative samples were calculated over their respective retained intervals using identical definitions to ensure direct comparability. After screening, the dry-snow dataset comprised 13 positive and 39 negative samples, and the wet-snow dataset comprised 24 positive and 72 negative samples. Each retained process block was treated as one classification unit. An example of the sample-construction procedure is shown in Fig. 5." (Sect. 2.4, lines 140–154)



"Figure 5. Example of the temporal construction of positive and negative samples using an observed wet-snow avalanche case. The red and blue blocks indicate the positive and retained negative samples, respectively. The grey shading represents the 12 h exclusion interval preceding the start of the positive block, and the red dashed line marks the recorded avalanche time." (Fig. 5, lines 156–158)

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.

Changes in the revised manuscript:

The first sentence of Sect. 2.4 now specifies the station assignment. More details are shown in lines 140–142 of the revised manuscript:

"For each recorded dry- or wet-snow avalanche, candidate snowfall or snowmelt-related process blocks were identified within a 3-d retrospective window at the monitoring station geographically closest to the avalanche location using the same process-identification criteria described above." (Sect. 2.4, lines 140–142)

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.

Changes in the revised manuscript:

Sect. 2.5 now acknowledges the previous use of logistic regression in avalanche forecasting. More details are shown in lines 160–162 of the revised manuscript:

"Within a statistical modelling framework, this study employs logistic regression (LR), which has previously been applied to avalanche forecasting (Gauthier et al., 2017), to develop an avalanche forecasting model that combines snowfall- or snowmelt-process variables with pre-event snow depth, hereafter referred to as the Snowpack–Meteorology Logistic Model (SMLM)." (Sect. 2.5, lines 160–162)

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.

Changes in the revised manuscript:

Sect. 2.5 now specifies the SMLM inputs and the corresponding MOLM baseline. More details are shown in lines 163–164 and 203–205 of the revised manuscript:

> "For each avalanche type, the SMLM uses three predictor variables: forcing intensity, forcing duration, and pre-event snow depth, with avalanche occurrence or non-occurrence treated as the binary response variable (Fu et al., 2026)." (Sect. 2.5, lines 163–164)
>
> "Specifically, the MOLM included the snowfall- or snowmelt-process predictors: SFI and SFD for dry-snow avalanches and SMI and SMD for wet-snow avalanches. The SMLM additionally included pre-event snow depth." (Sect. 2.5, lines 203–205)

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, although the additional benefit of SD was smaller than in LOOCV.

This single chronological split cannot establish stable performance across winters, especially with few events available for training. The results therefore provide an exploratory temporal assessment rather than a robust estimate of cross-season performance. Given this limitation and the manuscript's primary focus on locally observed avalanche characteristics and the relationships between avalanche occurrence, snowfall/snowmelt conditions, and pre-event snow depth, we have retained the cross-season results as supplementary analysis in this response rather than incorporating them into the manuscript.

More broadly, the contribution of this study lies in the new field observations and the quantitative characterisation of avalanche-related conditions in the southeastern Tibetan Plateau, where systematic avalanche records remain limited. By examining dry- and wet-snow avalanches separately, we identified positive associations of forcing intensity and pre-event snow depth with avalanche-event classification, together with negative conditional associations for forcing duration. These findings suggest that longer snowfall or snowmelt processes do not necessarily correspond to a higher likelihood of avalanche occurrence, highlighting the importance of considering forcing characteristics together with pre-event snowpack conditions. The observations and statistical relationships established here provide a foundation for refining local avalanche assessment and forecasting as longer monitoring records become available.

Changes in the revised manuscript:

The cross-season results above are supplementary to this response. In the manuscript, Sect. 4.4 now defines the scope of the current evaluation and the priorities for further validation. More details are shown in lines 392–404 of the revised manuscript:

> "The improved discrimination achieved by the SMLM highlights the value of combining pre-event snow depth with meteorological forcing, while the scope of the available observations places some limits on its broader application. The two-season record and the monitoring stations in Galongla Valley provide valuable local evidence, but do not encompass the full variability of snowpack and meteorological conditions across southeastern Tibet. Moreover, because model performance was assessed using LOOCV, its applicability to other locations and snow seasons remains to be established. At the methodological level, the LR framework

estimates statistical associations between avalanche-event status and the predictors, rather than explicitly modelling the physical processes governing snowpack stability." (Sect. 4.4, lines 392–398)

"Future research should focus on extending the observational record and examining model performance across a wider range of locations and snow seasons. Incorporating snowpack stability indices or mechanistic constraints within the existing statistical structure could strengthen the physical interpretation of the identified relationships. Furthermore, spatially distributed meteorological and topographical information could help account for local environmental differences and support the adaptation of this observation-based approach to other mountain regions. Together, these developments would provide a basis for more widely applicable, locally adapted avalanche assessment methods." (Sect. 4.4, lines 399–404)

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. The paired SMLM−MOLM TSS difference was 0.05 for dry snow (95% CI: −0.08 to 0.19) and 0.21 for wet snow (95% CI: 0.01 to 0.38). Thus, the threshold-dependent improvement is more clearly supported for wet snow; the dry-snow TSS gain remains uncertain.

Changes in the revised manuscript:

Sect. 2.5 now includes TSS in the evaluation, and Sect. 3.3 reports its values and paired uncertainty. More details are shown in lines 187–191 and 318–324 of the revised manuscript:

"The classification scores were then converted into binary classifications using this optimal threshold, and a confusion matrix was constructed (Table 1). Based on this matrix, a comprehensive set of performance metrics was calculated to evaluate classification performance, including TPR, FPR, true negative rate (TNR), false negative rate (FNR), precision, the true skill statistic (TSS), and F1-score (Fawcett, 2006). The mathematical formulations for these metrics are detailed in Eqs. (3) – (9)." (Sect. 2.5, lines 187–191)

"The SMLM demonstrates a more effective trade-off between false alarms and missed events at the selected classification thresholds (Table 2). For wet-snow avalanches, the MOLM yields an FPR of 0.21 and a precision of 0.56. In contrast, the SMLM reduces the FPR to 0.13 while increasing precision to 0.71 and TPR from 0.79 to 0.92. For dry-snow avalanches, the FPR decreased from 0.41 to 0.36, while precision increased from 0.43 to 0.46. The TPR remained unchanged at 0.92. TSS increased from 0.58 to 0.79 for wet-snow avalanches and from 0.51 to 0.56 for dry-snow avalanches. The paired bootstrap TSS differences (SMLM − MOLM) were 0.21 (95% CI: 0.01–0.38) for wet-snow avalanches and 0.05 (95% CI: −0.08–0.19) for dry-snow avalanches." (Sect. 3.3, lines 318–324)

The revised Table 2 provides the corresponding classification metrics, bootstrap confidence intervals, and confusion-matrix counts (Table 2, lines 326–329):

Table 2. Comparison of classification performance metrics between the MOLM and SMLM.

	Wet MOLM	Dry MOLM	Wet SMLM	Dry SMLM
TPR (95% CI)	0.79 (0.64–0.95)	0.92 (0.75–1.00)	0.92 (0.77–1.00)	0.92 (0.75–1.00)
FPR (95% CI)	0.21 (0.12–0.31)	0.41 (0.25–0.57)	0.13 (0.06–0.21)	0.36 (0.20–0.51)
TSS (95% CI)	0.58 (0.39–0.77)	0.51 (0.27–0.71)	0.79 (0.64–0.91)	0.56 (0.33–0.77)
TNR	0.79	0.59	0.88	0.64
FNR	0.21	0.08	0.08	0.08
Precision	0.56	0.43	0.71	0.46
F1	0.66	0.59	0.80	0.62
TN/FP/FN/TP	57/15/5/19	23/16/1/12	63/9/2/22	25/14/1/12

Line 205: I suggest replacing increased with were.

Response:

Thank you. We have replaced "increased" with "were" as suggested.

Changes in the revised manuscript:

More details are shown in lines 211–213 of the revised manuscript:

> "During the 2024 snow season, the mean snow depth and mean air temperature were 0.17 m and -0.27 °C, respectively. In 2025, the corresponding values were 0.24 m and 1.16 °C." (Sect. 3.1, lines 211–213)

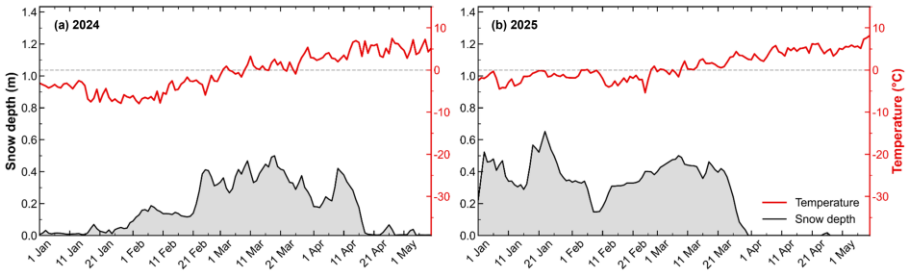
Figure 5: Please use the international day format (1 January 2024, or 1 Jan) as specified in https://www.natural-hazards-and-earth-system-sciences.net/submission.html#english; it is also specified there that in axis labels only the first word is capitalized.

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. The original Fig. 5 is Fig. 6 in the revised manuscript and is reproduced below.

Changes in the revised manuscript:

The revised Fig. 6 uses the corrected dates and axis labels. More details are shown in Fig. 6 and lines 217–218 of the revised manuscript:



> "Figure 6. Time series of daily air temperature and snow depth in the Galongla Valley during the 2024 and 2025 snow seasons. The red lines represent air temperature (°C), and the grey shaded areas indicate snow depth (m)." (Fig. 6, lines 217–218)

Line 246: I suggest omitting basal.

Response:

Thank you. We have omitted "basal".

Changes in the revised manuscript:

The revised sentence in Sect. 3.2.1 omits "basal":

> "Under conditions of a deep snowpack (high SD), even a marginal increment in SFI can drive the model prediction across the critical boundary into the above-threshold region." (Sect. 3.2.1, lines 254–256)

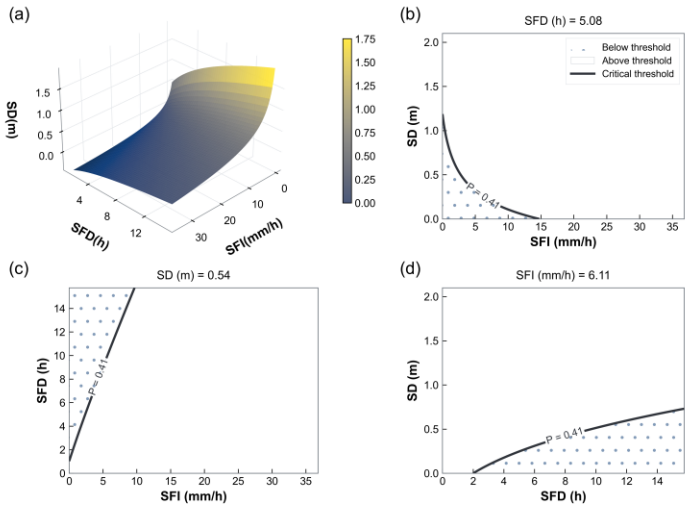
Figure 7: The use of the rainbow color bar is strongly discouraged for data visualization in science.

Response:

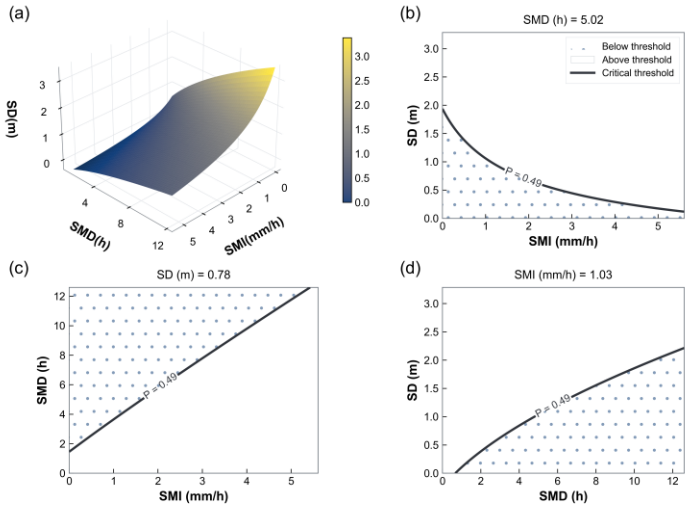
Thank you for this suggestion. We have replaced the rainbow colour scale in the original Fig. 7 (now Fig. 8) and made the same adjustment to the wet-snow model in Fig. 9. Both revised figures and their captions are reproduced below.

Changes in the revised manuscript:

More details are shown in Figs. 8 and 9 (lines 265–270 and 299–304) of the revised manuscript:



"Figure 8. Three-dimensional (3D) visualization of the dry-snow avalanche SMLM classification boundary and its two-dimensional (2D) cross-sections. (a) The 3D surface represents the combinations of SFI, SFD, and SD at which the model-derived classification score equals the classification threshold (P = 0.41); the surface colour denotes SD. (b–d) 2D cross-sections of the model-derived threshold surface, with the third variable fixed at its mean value, as indicated above each panel. The black solid lines represent the classification boundary. The unpatterned white and blue-dotted regions indicate model-derived classification scores above and below the classification threshold, respectively." (Fig. 8, lines 265–270)



"Figure 9. Three-dimensional (3D) visualization of the wet-snow avalanche SMLM classification boundary and its two-dimensional (2D) cross-sections. (a) The 3D surface represents the combinations of SMI, SMD, and SD at which the model-derived classification score equals the classification threshold (P = 0.49); the surface colour denotes SD. (b–d) 2D cross-sections of the model-derived threshold surface, with the third variable fixed at its mean value, as indicated above each panel. The black solid lines represent the classification boundary. The unpatterned white and blue-dotted regions indicate model-derived classification scores above and below the classification threshold, respectively." (Fig. 9, lines 299–304)

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. Figure R1 below shows 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. In dry-snow samples, event medians were 7.94 mm h^{-1} for SFI and 1.04 m for SD, compared with 3.49 mm h^{-1} and 0.15 m for non-events. For wet snow, the corresponding event medians were 1.74 mm h^{-1} for SMI and 2.35 m for SD, compared with 0.32 mm h^{-1} and 0.28 m for non-events.

For avalanche-event samples, median SFD and SMD were 4 h and 3 h, respectively, compared with 6 h in the corresponding non-event groups.

Across both avalanche types, event samples have higher median forcing intensity and pre-event snow depth than non-events, although their distributions overlap. This supports examining the predictors jointly rather than using one variable as a deterministic indicator. These descriptive comparisons complement the multivariable analyses presented in the manuscript.

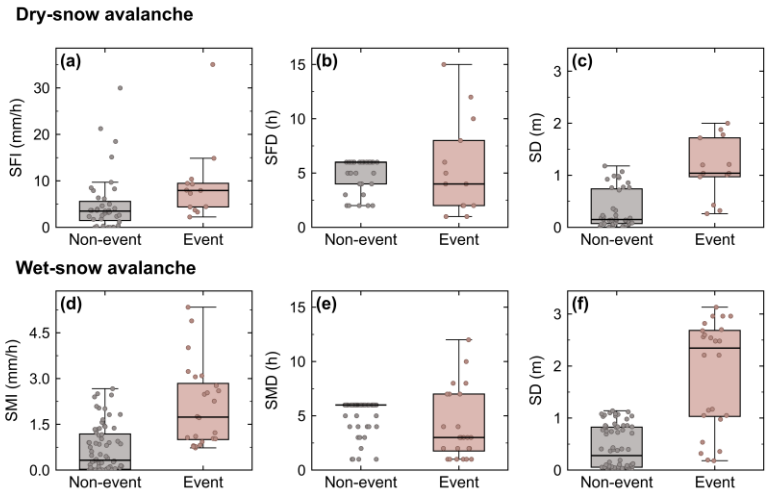


Figure R1. Distributions of predictor variables for avalanche events and non-events. Panels (a–c) show SFI, SFD, and SD for dry-snow samples, and panels (d–f) show SMI, SMD, and SD for wet-snow samples. Boxes indicate the interquartile range, horizontal lines indicate the median, whiskers extend to the most extreme observations within 1.5 times the interquartile range of the lower and upper quartiles, 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 revised the wording to refer specifically to "pre-event snow depth" rather than the broader term "information on snowpack conditions," as suggested.

Changes in the revised manuscript:

Sect. 4.2 now refers specifically to directly observed pre-event snow depth. More details are shown in Sect. 4.2 (lines 374–375) of the revised manuscript:

> "In contrast to approaches that rely on detailed snowpack simulations, our results show that directly observed pre-event snow depth can provide additional discriminatory information when combined with forcing intensity and duration." (Sect. 4.2, lines 374–375)

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. In the SMLMs, the SD coefficients are positive for dry snow (1.93; 95% CI: 1.07–2.71) and wet snow (2.68; 95% CI: 1.83–3.40). Together with the model comparisons, this supports the discriminatory value of measured snow depth. The revised discussion relates this empirical result to antecedent snow cover and terrain roughness, and to previous work on avalanches affecting transportation corridors, without inferring an unobserved weak-layer failure.

Changes in the revised manuscript:

Sect. 4.1 now discusses the positive association between pre-event snow depth and avalanche occurrence in relation to antecedent snow cover, terrain roughness, and previous studies of avalanches affecting transportation corridors, without attributing the association to a specific failure mechanism. More details are shown in Sect. 4.1 (lines 333–340) of the revised manuscript:

> "Avalanche activity is influenced by the combined effects of meteorological forcing and antecedent snow cover under different terrain conditions. The model results show that SD, representing pre-event snow depth, is positively associated with the model-derived relative likelihood of avalanche occurrence for both dry- and wet-snow avalanches. This finding is consistent with previous studies indicating that sufficient pre-existing snow cover can reduce the influence of terrain roughness on avalanche formation and affect the amount of new snow required for avalanche release (Schweizer et al., 2003a). Previous investigations of large avalanches affecting transportation corridors have also identified antecedent snow depth as a relevant factor in avalanche forecasting (Schweizer et al., 2009). These findings highlight the value of considering pre-event snow depth together with meteorological forcing when assessing recorded avalanche occurrence." (Sect. 4.1, lines 333–340)

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 replaced McClung (2021) with Schweizer et al. (2003a).

Changes in the revised manuscript:

Sect. 4.1 now describes the loading effect of snowfall without the redundant wording and cites Schweizer et al. (2003a) in place of McClung (2021). More details are shown in Sect. 4.1 (lines 348–350) of the revised manuscript:

> "Snowfall increases the gravitational load on the existing snowpack, and rapid loading can promote instability when the rate of loading exceeds the rate of snowpack strengthening (Schweizer et al., 2003a)." (Sect. 4.1, lines 348–350)

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 correlation coefficients were close to zero, and neither correlation was statistically significant, providing no evidence of a linear or monotonic association between SFD and SFI in these samples.

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

Response:

Thank you. The original statement has been removed from the revised Discussion, which now describes the observed associations.

Changes in the revised manuscript:

More details are shown in Sect. 4.1 (lines 353–356) of the revised manuscript:

> "Our modelling results indicate that both SFD for dry-snow avalanches and SMD for wet-snow avalanches are negatively associated with the model-derived relative likelihood of avalanche occurrence when forcing intensity and pre-event snow depth are held constant." (Sect. 4.1, lines 353–356)

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 pointing this out. We agree that the term "complex coupling" overstated what our analysis can establish. The logistic regression models include additive terms for the log-transformed predictors and do not estimate interaction terms or resolve physical coupling mechanisms. We have revised the text to describe the conditional statistical associations rather than imply that the model demonstrates physical coupling.

Changes in the revised manuscript:

The revised text replaces the claim of complex coupling with a conditional statistical interpretation of process duration and clarifies that the model estimates statistical associations rather than physical mechanisms. More details are shown in Sects. 4.1 and 4.4 (lines 353–356 and 396–398, respectively) of the revised manuscript:

> "Our modelling results indicate that both SFD for dry-snow avalanches and SMD for wet-snow avalanches are negatively associated with the model-derived relative likelihood of avalanche occurrence when forcing intensity and pre-event snow depth are held constant." (Sect. 4.1, lines 353–356)

> "At the methodological level, the LR framework estimates statistical associations between avalanche-event status and the predictors, rather than explicitly modelling the physical processes governing snowpack stability." (Sect. 4.4, lines 396–398)

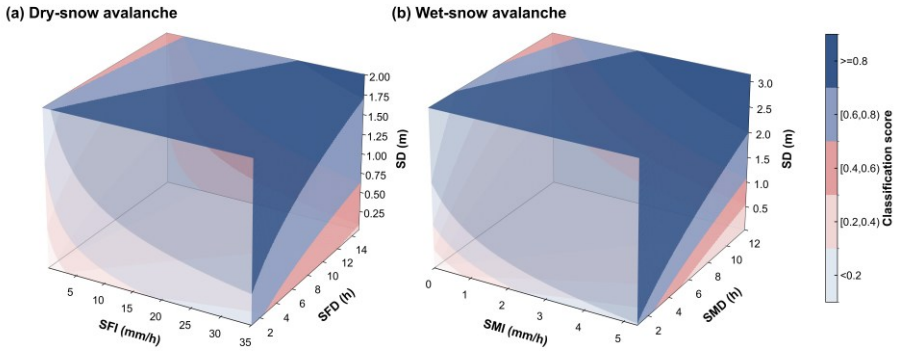
Figure 10: I suggest changing the color bar.

Response:

Thank you for the suggestion. The original Fig. 10 is Fig. 11 in the revised manuscript. The colour bar has been revised, and the updated figure is shown below.

Changes in the revised manuscript:

The colour bar in Fig. 11 has been revised. More details are shown in Fig. 11 (lines 389–390) of the revised manuscript:



> "Figure 11. Illustrative three-dimensional classification-score intervals derived from the SMLM for (a) dry- and (b) wet-snow avalanches." (Fig. 11, lines 389–390)

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 replaced the ambiguous term with an explicit description of the variables included in each model.

Changes in the revised manuscript:

The Conclusions now explicitly identify the variables combined in the dry- and wet-snow models, replacing the ambiguous term "integrated". More details are shown in the Conclusions (lines 406–407) of the revised manuscript:

> "In this study, we developed separate logistic regression models for dry- and wet-snow avalanches in Galongla Valley, southeastern Tibetan Plateau, combining pre-event snow depth with snowfall- or snowmelt-process variables." (Conclusions, lines 406–407)

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 an explicit description of the variables included in each model.

Changes in the revised manuscript:

The Conclusions replace the ambiguous term "state-constrained framework" with a direct description of the variables combined in the models. More details are shown in the Conclusions (lines 406–407) of the revised manuscript:

> "In this study, we developed separate logistic regression models for dry- and wet-snow avalanches in Galongla Valley, southeastern Tibetan Plateau, combining pre-event snow depth with snowfall- or snowmelt-process variables." (Conclusions, lines 406–407)

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 for the suggestion. We have revised the wording to clarify the contribution of pre-event snow depth to the observed gains in model performance.

Changes in the revised manuscript:

The revised Conclusions now state more directly the contribution of pre-event snow depth to model discrimination. More details are shown in the Conclusions (lines 413–419) of the revised manuscript:

> "Incorporating SD improved model discrimination for both avalanche types under LOOCV. Compared with the MOLM, the SMLM achieved a TPR of 0.92 for both dry- and wet-snow avalanches at the selected classification thresholds, increased the AUC from 0.72 to 0.83 for dry-snow avalanches and from 0.85 to 0.94 for wet-snow avalanches, and reduced the FPR from 0.41 to 0.36 and from 0.21 to 0.13, respectively. These results show that incorporating pre-event snow depth alongside snowfall- and snowmelt-process variables improves the identification of avalanche-event conditions for both avalanche types. Based on local observations from a region where avalanche studies remain limited, our findings provide a practical basis for refining avalanche monitoring and informing the development of locally adapted early-warning methods in Galongla Valley." (Conclusions, lines 413–419)

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 dsataset to avoid implying unwarranted precision. The updated dataset is available in Version 2.0 of the Zenodo repository.

Changes in the revised manuscript:

The precision change concerns the deposited dataset. The manuscript now gives the updated repository reference in its Code and data availability statement (lines 421–423):

> "The processed dataset used in this study, including meteorological indicators, snow depth, avalanche-type labels, and binary avalanche occurrence records, and the Python scripts used for the SMLM and MOLM modelling have been deposited in Zenodo at https://doi.org/10.5281/zenodo.23391913." (Code and data availability, lines 421–423)

Referee 2

General comments

This manuscript presents meteorological, snow-depth, and avalanche observations from the Galongla Valley, a region where systematic avalanche data remain relatively scarce. The distinction between dry- and wet-snow avalanches is appropriate, and the study could be useful for local avalanche risk management.

The results suggest that snow depth prior to a snowfall or snowmelt episode contains useful information. However, some aspects of the sampling design and validation currently limit the interpretation of the reported probabilities and forecasting performance.

The manuscript would also benefit from a clearer statement of its original contribution to avalanche forecasting. Moreover, the observations cover only two snow seasons. This relatively short period may limit the generalisability of the conclusions, as the results may partly reflect snowpack characteristics specific to those seasons. This issue is particularly relevant because snow depth is the only snowpack variable considered in the analysis.

Response:

Thank you for this assessment and for recognising the value of the observations from Galongla Valley. We agree that the two-season record and the sampling design constrain the scope of the conclusions, and that the scientific contribution of the study needs to be more clearly articulated.

The main contribution is a region-specific observational baseline and a quantitative assessment of the additional discriminatory information provided by directly observed pre-event snow depth. The ten-station network documented 37 avalanche events during the 2024 and 2025 snow seasons, including 13 dry-snow and 24 wet-snow avalanches. Of the 23 mapped avalanche paths, 12 directly threaten transportation routes. The observations provide an initial characterisation of local avalanche activity: dry-snow events were recorded from late February to the end of March, whereas wet-snow events continued into mid-April. In a region where systematic avalanche records remain scarce, these observations provide an important empirical basis for understanding local avalanche occurrence and its seasonal characteristics, particularly given the need to assess avalanche hazards under a changing climate.

Comparing type-specific logistic regression models with and without pre-event snow depth provides an interpretable assessment of its incremental discriminatory value. Under leave-one-out cross-validation (LOOCV), incorporating pre-event snow depth increased the area under the receiver operating characteristic curve (AUC) from 0.72 to 0.83 for dry-snow avalanches and from 0.85 to 0.94 for wet-snow avalanches. At the selected classification thresholds, the false-positive rate decreased from 0.41 to 0.36 and from 0.21 to 0.13, respectively. These results provide local empirical evidence that directly observed pre-event snow depth adds discriminatory information to snowfall- or snowmelt-process intensity and duration.

To address the methodological concerns, we have clarified the construction of positive and negative samples, including the retrospective selection criteria, exclusion intervals, and the 1:3 positive-to-negative sample ratio. Given this sampling design, the model outputs are interpreted as model-derived relative classification scores rather than calibrated avalanche-occurrence probabilities. We have also clarified that LOOCV assesses internal discrimination and does not establish independent forecasting performance across different snow seasons or locations.

We acknowledge that two snow seasons cannot capture the full variability of meteorological and snowpack conditions, and that the observed relationships may partly reflect snowpack characteristics specific to those seasons. Pre-event snow depth alone does not characterise snowpack stratigraphy or stability. The findings therefore describe local observational relationships and preliminary type-specific modelling, whose applicability to other snow seasons and locations remains to be established. Their value lies in establishing a local event record, identifying basic patterns of avalanche occurrence, and quantifying the additional discriminatory information associated with measured pre-event snow depth. As observations accumulate over subsequent seasons, future research can extend this foundation to broader validation and more detailed investigations of avalanche motion and physical mechanisms.

Changes in the revised manuscript:

The Introduction clarifies the observational contribution and the model comparison used to assess the additional information provided by pre-event snow depth. More details are shown in lines 64–71 of the revised manuscript:

> "Focusing on an avalanche-prone high mountain valley in the southeastern Tibetan Plateau, this study compiles an event database using a dedicated in situ monitoring network that integrates meteorological, snow depth, and avalanche observations. Pre-event snow depth is defined as the snow depth at the onset of the preceding snowfall or snowmelt process. Separate logistic regression models are developed for dry- and wet-snow avalanches to quantify the relationships between forcing intensity, forcing duration, pre-event snow depth, and avalanche occurrence. These models are compared with meteorology-only baselines to assess whether the inclusion of pre-event snow depth improves event discrimination and reduces false-alarm rates. By incorporating observed pre-event snow depth into type-specific logistic regression models, this study provides an interpretable quantitative basis for regional avalanche monitoring and forecasting." (Introduction, lines 64–71)

Sect. 2.4 specifies the maximum negative-sample duration, exclusion intervals, and overall sampling ratio. More details are shown in Sect. 2.4 (lines 144–149) of the revised manuscript:

> "Candidate negative sample intervals were limited to a maximum duration of 6 h, even when the underlying processes lasted longer. Intervals overlapping the 12 h immediately preceding the start of the positive block were excluded, and eligible intervals closer to the positive block were prioritised. Candidate blocks overlapping any known avalanche occurrence at the same station and of the same avalanche type were also removed using the full avalanche-time inventory. Eligible negative samples were retained to achieve an overall positive-to-negative sample ratio of 1:3 for each avalanche type." (Sect. 2.4, lines 144–149)

Sect. 4.2 explains the contribution of directly observed pre-event snow depth to model discrimination in data-limited mountain regions. More details are shown in Sect. 4.2 (lines 374–378) of the revised manuscript:

> "In contrast to approaches that rely on detailed snowpack simulations, our results show that directly observed pre-event snow depth can provide additional discriminatory information when combined with forcing intensity and duration. By using a small set of readily available variables, the SMLM offers an observation-based approach to distinguishing avalanche-event and non-event conditions in data-limited mountain regions, providing a basis for the further development of locally adapted avalanche assessment methods." (Sect. 4.2, lines 374–378)

Sect. 4.4 clarifies the limitations of the two-season record and internal validation, while identifying priorities for extending the observations and strengthening the physical interpretation. More details are shown in Sect. 4.4 (lines 392–404) of the revised manuscript:

> "The improved discrimination achieved by the SMLM highlights the value of combining pre-event snow depth with meteorological forcing, while the scope of the available observations places some limits on its broader application. The two-season record and the monitoring stations in Galongla Valley provide valuable local evidence, but do not encompass the full variability of snowpack and meteorological conditions across southeastern Tibet. Moreover, because model performance was assessed using LOOCV, its applicability to other locations and snow seasons remains to be established. At the methodological level, the LR framework estimates statistical associations between avalanche-event status and the predictors, rather than explicitly modelling the physical processes governing snowpack stability.
>
> Future research should focus on extending the observational record and examining model performance across a wider range of locations and snow seasons. Incorporating snowpack stability indices or mechanistic constraints within the existing statistical structure could strengthen the physical interpretation of the identified relationships. Furthermore, spatially distributed meteorological and topographical information could help account for local environmental differences and support the adaptation of this observation-based approach to other mountain regions. Together, these developments would provide a basis for more widely applicable, locally adapted avalanche assessment methods." (Sect. 4.4, lines 392–404)

Major comments

1. Sample construction and interpretation of probabilities

Section 2.4 is central to the study but remains difficult to reconstruct. Please clarify how avalanches were assigned to monitoring stations, how positive and negative process blocks were identified, whether several samples could belong to the same meteorological episode, and whether the same process could be selected more than once. A schematic time series showing the positive block, excluded periods, buffer, and retained negative blocks would be very helpful.

The datasets contain three selected negative samples per positive sample. Avalanche prevalence is therefore determined by the sampling design and does not represent prevalence among all snowfall or snowmelt processes.

The authors should define the intended prediction unit and explain the consequences of this sampling design. If possible, the model should be calibrated using all eligible processes or corrected for the sampling fractions. Otherwise, the outputs should be described as relative classification scores rather than operational occurrence probabilities.

Response:

Thank you for this important comment. We agree that the sample construction, prediction unit, and interpretation of the model outputs should be stated explicitly. For each recorded avalanche, meteorological and snow-depth observations were obtained from the geographically closest monitoring station. Within the preceding 3-day window, qualifying snowfall or snowmelt-related process blocks were identified. The most recent qualifying process before the recorded avalanche time was selected as the positive block, while earlier qualifying processes were considered candidate negative blocks. For positive samples, predictor variables were calculated only up to the recorded avalanche time; any process still ongoing at that time was truncated there. Candidate negative sample intervals were limited to a maximum duration of 6 h, even when the underlying processes lasted longer. Candidate negative blocks overlapping the 12 h immediately preceding the start of the positive block were excluded. Candidate blocks overlapping any known avalanche occurrence at the same station and of the same avalanche type were also removed. Eligible negative samples were retained to achieve an overall positive-to-negative sample ratio of 1:3 for each avalanche type, with those closer to the positive block prioritised. Each segmented candidate block contributed at most one retained sample, so no individual block was selected repeatedly. However, multiple retained blocks could originate from the same broader snowfall or snowmelt episode.

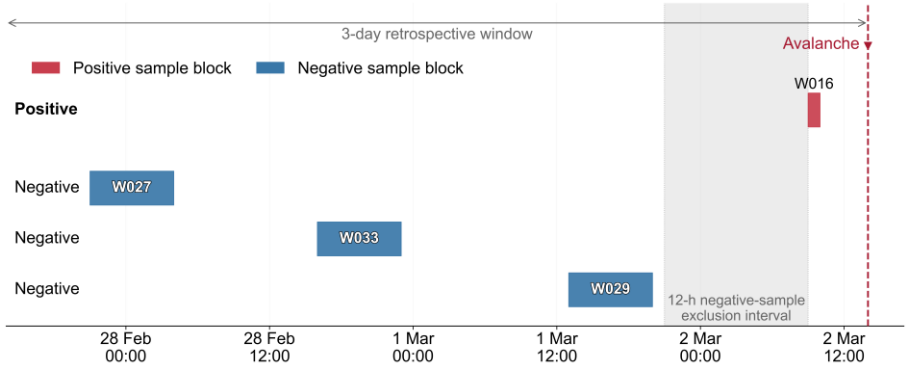
The prediction unit in this retrospective classification analysis is an identified snowfall or snowmelt process block, characterised by its intensity, duration, and pre-event snow depth and assigned an event/non-event binary label. Because negative samples were retained to achieve an overall positive-to-negative ratio of 1:3, this ratio reflects the sampling design and does not represent the natural avalanche prevalence among all eligible snowfall or snowmelt processes. Accordingly, the model outputs are interpreted as relative classification scores for the sampled process blocks and not as calibrated absolute avalanche occurrence probabilities. The sample-selection procedure is illustrated in Fig. 5 of the revised manuscript. Using retained samples from the modelling dataset, the figure shows the positive block, retained negative blocks, 3-day retrospective window, 12 h exclusion buffer, and recorded avalanche occurrence time on a common time axis.

Changes in the revised manuscript:

Sect. 2.4 specifies how the nearest monitoring station was assigned to each avalanche, how positive and candidate negative process blocks were identified, the exclusion criteria and sampling ratio, and the classification unit. More details are shown in Sect. 2.4 (lines 140–154) of the revised manuscript:

> "For each recorded dry- or wet-snow avalanche, candidate snowfall or snowmelt-related process blocks were identified within a 3-d retrospective window at the monitoring station geographically closest to the avalanche location using the same process-identification criteria described above. The most recent qualifying process block preceding the recorded avalanche time was treated as the positive sample, with all variables calculated only up to the avalanche occurrence time. Earlier process blocks within the same retrospective window were considered candidate negative samples. Candidate negative sample intervals were limited to a maximum duration of 6 h, even when the underlying processes lasted longer. Intervals overlapping the 12 h immediately preceding the start of the positive block were excluded, and eligible intervals closer to the positive block were prioritised. Candidate blocks overlapping any known avalanche occurrence at the same station and of the same avalanche type were also removed using the full avalanche-time inventory. Eligible negative samples were retained to achieve an overall positive-to-negative sample ratio of 1:3 for each

avalanche type. Each candidate process block contributed at most one retained sample, although multiple blocks could originate from the same broader snowfall or snowmelt episode. The predictor variables for positive and negative samples were calculated over their respective retained intervals using identical definitions to ensure direct comparability. After screening, the dry-snow dataset comprised 13 positive and 39 negative samples, and the wet-snow dataset comprised 24 positive and 72 negative samples. Each retained process block was treated as one classification unit. An example of the sample-construction procedure is shown in Fig. 5." (Sect. 2.4, lines 140–154)



"Figure 5. Example of the temporal construction of positive and negative samples using an observed wet-snow avalanche case. The red and blue blocks indicate the positive and retained negative samples, respectively. The grey shading represents the 12 h exclusion interval preceding the start of the positive block, and the red dashed line marks the recorded avalanche time." (Fig. 5, lines 156–158)

Sect. 2.5 defines the logistic-regression output as a classification score for a sampled process block. More details are shown in Sect. 2.5 (lines 171–173) of the revised manuscript:

"where P(Y = 1 | x) represents the model-derived avalanche classification score for a sampled process block, w_i is the regression coefficient for the i-th input variable, b is the intercept term, and k is the number of input variables. Model parameters are estimated using maximum likelihood estimation." (Sect. 2.5, lines 171–173)

Sect. 4.3 clarifies that the scores describe predictor combinations associated with recorded avalanche events within the sampled data and discusses their potential relevance to locally adapted early-warning criteria. More details are shown in Sect. 4.3 (lines 380–387) of the revised manuscript:

"The SMLM provides continuous classification scores based on pre-event snow depth and the intensity and duration of snowfall or snowmelt processes. These scores allow the model response to different combinations of predictors to be examined beyond a single binary classification threshold. For illustrative purposes, the model-derived scores were divided into five equal-width intervals (Fig. 11). The resulting three-dimensional patterns illustrate how model scores vary with meteorological forcing and pre-event snow depth, and how different combinations of predictors can produce similar scores. Higher scores indicate predictor combinations more strongly associated with recorded avalanche events within the sampled data. These patterns provide an interpretable quantitative basis for identifying conditions associated with avalanche occurrence and informing the development of locally adapted avalanche early-warning criteria." (Sect. 4.3, lines 380–387)

2. Interpretation of the forcing-duration coefficients

The negative coefficients for SFD and SMD carry considerable interpretive weight. Section 4.1 discusses them at length in terms of settlement, sintering, viscous creep, preferential flow, and restabilisation, while l. 340 presents the role of forcing duration as a relatively unexplored aspect of avalanche initiation. I have three concerns about this interpretation.

First, intensity and duration are mathematically related in the present framework: SFI and SMI are defined as cumulative amount divided by the corresponding duration (l. 117–118 and l. 134–136). Consequently, interpreting the fitted coefficients as fully separate effects of forcing rate and duration is not straightforward. The `ln(1 + x)` transformation further means that there is no simple exact equivalence between the current formulation and a model expressed in terms of cumulative amount and rate. Nevertheless, the interpretation of the duration coefficient may depend on the chosen parameterisation. It would therefore be helpful if the authors could examine this relationship and compare the present model with an alternative formulation using

cumulative amount and rate as predictors. This would clarify whether the negative duration effect is robust or partly dependent on how the predictors are defined.

Second, the definitions at l. 114–115 and l. 132–133 do not appear to impose a maximum process duration, whereas the non-event durations reported in the deposited dataset span a comparatively narrow range. Could the authors clarify whether a fixed window or maximum duration was applied when extracting candidate non-event blocks? If event and non-event durations were determined using different observation rules, part of the fitted duration effect could reflect the sampling procedure rather than snowpack behaviour.

Third, the statement at l. 340 that the role of forcing duration has rarely been examined would benefit from qualification. The relationship between loading rate and strength development in buried weak layers is discussed, for example, by Schweizer et al. (2003a) and within the stability-index approach of Conway and Wilbour (1999). I suggest placing the present results more explicitly within this literature and clarifying how the treatment of duration in this study differs from previous approaches.

Finally, the negative association between forcing duration and avalanche occurrence may be specific to the maritime snow climate and snowpack conditions represented in this study. It may not necessarily apply to snowpacks containing persistent weak layers, where cumulative loading over longer periods may play a different role. Marienthal et al. (2015), already cited in the manuscript, report an association between deep-slab avalanche days and precipitation accumulated over the preceding seven days. This provides a useful contrast and could help define the limits within which the present results may be interpreted. Because the response variable is binary and contains no information on avalanche size, the conclusions should remain limited to occurrence probability and should not be extended to event magnitude.

Response:

Thank you for raising these closely related concerns. To examine the dependence on predictor parameterisation, we compared the current rate–duration formulation with an alternative cumulative amount–rate formulation using exactly the same samples, labels, modelling procedure, and LOOCV scheme. Cumulative amount was calculated as mean intensity multiplied by duration. Because both formulations use the $\ln(1 + x)$ transformation, this comparison is a sensitivity analysis using alternative predictors, rather than an exact algebraic re-expression of the current model. The two formulations showed broadly comparable discrimination (Table R1). When duration was replaced by cumulative amount, the fitted cumulative-amount coefficients were also negative; however, their bootstrap 95% confidence intervals included zero in three of the four model formulations. Thus, the coefficient direction was consistent across the tested formulations, but the statistical support for a negative association was less consistent in the amount–rate models. We therefore interpret the negative SFD and SMD coefficients as conditional empirical associations between the temporal characteristics of the identified processes and avalanche occurrence, rather than as evidence of an independent causal effect of duration.

Regarding the comparatively narrow range of non-event durations in the dataset, this pattern partly reflects a 6 h segmentation criterion applied when constructing candidate non-event process blocks. This criterion was used to segment candidate negative blocks and did not represent a physical maximum duration of the underlying snowfall or snowmelt process. To assess its influence, we reconstructed longer negative process blocks without the maximum-duration restriction, retaining the original positive samples, and repeated the modelling as a diagnostic sensitivity analysis. The median durations of the reconstructed negative processes were 10 h for dry snow and 13.5 h for wet snow. Reconstruction changes the number and duration of negative blocks, as well as their boundaries and mean forcing rates. We therefore do not use this analysis for a direct comparison of predictive performance. Nevertheless, the fitted duration coefficients remained negative, with bootstrap 95% confidence intervals excluding zero, in both the SMLM and MOLM for both avalanche types (Table R1). The negative coefficient direction therefore persists under these reconstruction rules.

We also agree that the statement that the role of forcing duration has rarely been examined is too broad. We have replaced this statement with an explicit acknowledgement of earlier work on time-dependent snow stability. Conway and Wilbour (1999) and Schweizer et al. (2003a) describe the competition between loading and snow-strength development. Our study uses duration as a predictor in a statistical classification of identified process blocks; it does not explicitly calculate the evolution of weak-layer strength or a time-dependent stability index.

Finally, we agree that the negative association between forcing duration and avalanche occurrence should be interpreted within the snow-climate conditions represented by the present observations. Marienthal et al. (2015) found that days with deep-slab avalanches on persistent weak layers were often preceded by greater

precipitation over the previous seven days than days without such events. Their fixed-window accumulation measure and deep-slab setting differ from the conditional process-duration associations examined here. Because the present analysis contains no direct information on weak-layer properties, it cannot establish whether persistent weak layers contributed to individual events. We therefore interpret the negative coefficients as conditional associations between snowfall or snowmelt process duration and avalanche-event status among the sampled process blocks. These findings provide region-specific observational evidence of duration-related differences between event and non-event conditions.

The sensitivity analyses and Table R1 are supplementary results provided in this response; the manuscript passages quoted below present the sampling definitions and the revised conditional interpretation.

Table R1. Sensitivity of model discrimination and fitted coefficients to predictor parameterisation and negative-process reconstruction.

Avalanche type	Model	Model discrimination (LOOCV AUC)		Fitted coefficient, β (bootstrap 95% CI)		
		Duration-based	Amount-based	Duration (Duration-based)	Amount (Amount-based)	Duration (Reconstructed negatives)
Dry	SMLM	0.83	0.85	−0.62 (−1.36, −0.02)	−0.19 (−0.77, 0.35)	−1.18 (−1.57, −0.76)
	MOLM	0.72	0.71	−0.56 (−1.36, 0.12)	−0.12 (−0.71, 0.51)	−1.22 (−1.61, −0.79)
Wet	SMLM	0.94	0.94	−1.49 (−1.97, −1.10)	−0.57 (−1.05, −0.11)	−1.17 (−1.46, −0.83)
	MOLM	0.85	0.82	−0.93 (−1.63, −0.44)	−0.01 (−0.56, 0.47)	−1.02 (−1.37, −0.69)

Note: Values in parentheses are bootstrap 95% confidence intervals. All coefficients refer to ln(1 + x)-transformed predictors. Duration-based models include forcing rate and duration, whereas amount-based models include forcing rate and cumulative amount. SMLM additionally includes pre-event snow depth (SD), whereas MOLM does not. Cumulative amount equals mean rate multiplied by duration and represents cumulative snowfall for dry snow and cumulative snow-depth loss for wet snow. Reconstructed negatives denote longer negative process blocks obtained without the maximum-duration restriction.

Changes in the revised manuscript:

Sect. 2.4 now explicitly states the segmentation and exclusion rules used in the existing sample construction. More details are shown in Sect. 2.4 (lines 143–150) of the revised manuscript:

> "Earlier process blocks within the same retrospective window were considered candidate negative samples. Candidate negative sample intervals were limited to a maximum duration of 6 h, even when the underlying processes lasted longer. Intervals overlapping the 12 h immediately preceding the start of the positive block were excluded, and eligible intervals closer to the positive block were prioritised. Candidate blocks overlapping any known avalanche occurrence at the same station and of the same avalanche type were also removed using the full avalanche-time inventory. Eligible negative samples were retained to achieve an overall positive-to-negative sample ratio of 1:3 for each avalanche type. Each candidate process block contributed at most one retained sample, although multiple blocks could originate from the same broader snowfall or snowmelt episode." (Sect. 2.4, lines 143–150)

Sect. 3.2.1 describes the dry-snow duration result as a conditional association in the fitted model. More details are shown in Sect. 3.2.1 (lines 246–248) of the revised manuscript:

> "This result indicates that, after accounting for pre-event snow depth, longer snowfall duration is associated with a lower model-derived classification score for dry-snow avalanches." (Sect. 3.2.1, lines 246–248)

Sect. 3.2.2 describes the wet-snow duration result while explicitly holding SMI and SD constant. More details are shown in Sect. 3.2.2 (lines 284–285) of the revised manuscript:

> "Furthermore, the regression coefficient for SMD is negative (w_2 = −1.49; 95% CI [−1.97, −1.10]), indicating a negative association between snowmelt duration and wet-snow avalanche occurrence when SMI and SD are held constant." (Sect. 3.2.2, lines 284–285)

Sect. 4.1 now acknowledges earlier time-dependent stability research and distinguishes the observed relationships from their possible physical context. More details are shown in Sect. 4.1 (lines 352–362) of the revised manuscript:

> "Previous studies have examined the time-dependent evolution of snowpack stability under meteorological loading, particularly the balance between loading rates and snow strength development (Conway and Wilbour, 1999; Schweizer et al., 2003a). Our modelling results indicate that both SFD for dry-snow avalanches and SMD for wet-snow avalanches are negatively associated with the model-derived relative likelihood of avalanche occurrence when forcing intensity and pre-event snow depth are held constant. For

dry-snow avalanches, snowfall loading is accompanied by time-dependent processes, including settlement, compaction, sintering, and stress redistribution, which influence the evolution of snowpack stability during a snowfall episode (Schweizer et al., 2003a). For wet-snow avalanches, liquid-water infiltration and redistribution evolve over time, with snow stratigraphy, preferential flow paths, and local water accumulation influencing the development of wet-snow instability (Mitterer et al., 2011; Wever et al., 2016). These processes provide a physical context for the observed duration-related associations, highlighting the relevance of snowfall and snowmelt duration in distinguishing avalanche-event from non-event conditions." (Sect. 4.1, lines 352–362)

Sect. 4.4 now limits the interpretation to statistical associations supported by the available observations. More details are shown in Sect. 4.4 (lines 392–398) of the revised manuscript:

"The improved discrimination achieved by the SMLM highlights the value of combining pre-event snow depth with meteorological forcing, while the scope of the available observations places some limits on its broader application. The two-season record and the monitoring stations in Galongla Valley provide valuable local evidence, but do not encompass the full variability of snowpack and meteorological conditions across southeastern Tibet. Moreover, because model performance was assessed using LOOCV, its applicability to other locations and snow seasons remains to be established. At the methodological level, the LR framework estimates statistical associations between avalanche-event status and the predictors, rather than explicitly modelling the physical processes governing snowpack stability." (Sect. 4.4, lines 392–398)

3. Validation and uncertainty

Leave-one-observation-out cross-validation treats each row as independent. However, samples may be grouped by station, season, avalanche path, or meteorological episode. Closely related observations could therefore occur in both the training and test sets, potentially leading to optimistic performance estimates.

I recognise that the limited number of avalanches and the availability of only two snow seasons make fully independent validation difficult. This limitation should be stated clearly, and the results should be described as internal discrimination rather than generalisable forecasting performance. If feasible, an exploratory comparison in which one winter is used for training and the other for testing could provide some indication of temporal transferability, although the outcome would need to be interpreted cautiously because two seasons are insufficient for a robust validation.

The dry-snow model contains only 13 positive events and three predictors. Confidence intervals should therefore be reported for the AUC and threshold-dependent metrics, as well as for differences between the models. Confusion-matrix counts would also help readers interpret the reported rates.

Response:

Thank you for this important comment. We first conducted an exploratory cross-season validation, using the 2024 snow season for training and the 2025 snow season for validation. Classification thresholds were selected using only the 2024 training data by maximising the Youden index ($J = TPR - FPR$) within internal LOOCV, and were then applied to the 2025 predictions. For dry-snow avalanches, the MOLM and SMLM achieved AUC values of 0.74 and 0.75, respectively, 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. Thus, adding pre-event snow depth yielded a small increase in AUC for both avalanche types, while the threshold-dependent improvement was limited to wet-snow avalanches. Given the limited observational record, comprising only two snow seasons and a relatively small number of avalanche events, this single chronological split provides preliminary evidence of temporal transferability but cannot establish robust performance across years. Moreover, the primary aim of this study is to quantify the additional discriminatory information provided by measured pre-event snow depth beyond snowfall or snowmelt intensity and duration. We therefore agreed to present the cross-season results only as an exploratory analysis in this response, rather than incorporate them into the manuscript.

As the reviewer notes, LOOCV does not account for potential dependence among samples from the same station or meteorological episode. We therefore interpret the main results as internal discrimination within the constructed dataset, rather than as evidence of generalisable forecasting performance. To quantify uncertainty within this evaluation framework, we used 1,000 paired bootstrap resamples of the LOOCV out-of-fold prediction records to estimate 95% confidence intervals for AUC, TPR, FPR, and TSS, as well as for differences in AUC and TSS between the SMLM and MOLM.

The results show that adding measured pre-event snow depth improved overall discrimination for both avalanche types. For dry-snow avalanches, AUC increased from 0.72 (95% CI: 0.56–0.86) for the MOLM to 0.83 (0.71–0.94) for the SMLM; for wet-snow avalanches, it increased from 0.85 (0.76–0.92) to 0.94 (0.88–0.99). The paired AUC differences were 0.11 (95% CI: 0.01–0.24) and 0.09 (0.02–0.17), respectively, with both confidence intervals excluding zero. The threshold-dependent results further distinguish the two avalanche types. For wet-snow avalanches, TSS increased from 0.58 (95% CI: 0.39–0.77) to 0.79 (0.64–0.91), with a paired difference of 0.21 (0.01–0.38). For dry-snow avalanches, TSS increased from 0.51 (0.27–0.71) to 0.56 (0.33–0.77), but the paired difference of 0.05 (−0.08 to 0.19) had a confidence interval that included zero. These findings indicate that the contribution of pre-event snow depth is evident in overall discrimination for both avalanche types, whereas the improvement in threshold-dependent classification is more clearly supported for wet-snow avalanches.

The confusion-matrix counts in Table 2 clarify the practical meaning of these differences. For wet-snow avalanches, adding SD reduced false positives from 15 to 9 and missed events from 5 to 2. For dry-snow avalanches, false positives decreased from 16 to 14, while missed events remained at one. Thus, incorporating SD improved both event detection and false-alarm control for wet-snow avalanches, whereas its main contribution at the selected dry-snow classification threshold was to reduce false alarms without changing the number of missed events. These results provide local empirical evidence that directly observed pre-event snow depth contributes information beyond meteorological process intensity and duration.

In the revised manuscript, Sect. 4.4 acknowledges the limited two-season record and clarifies that applicability to other locations and snow seasons remains to be established. Within this scope, the model comparison provides local observational evidence that measured pre-event snow depth adds discriminatory information beyond snowfall or snowmelt intensity and duration, offering a basis for further developing locally adapted avalanche assessment and early-warning criteria.

Changes in the revised manuscript:

Sect. 2.5 describes the LOOCV procedure for assessing model discrimination and the paired bootstrap procedure for quantifying uncertainty in model performance and differences between the models. More details are shown in Sect. 2.5 (lines 181–184 and 205–207) of the revised manuscript:

> "The discriminative performance of the model was evaluated using leave-one-out cross-validation (LOOCV). In the LOOCV framework, one observation is withheld as the test sample in each iteration, whereas the remaining observations are used for model training. This procedure is repeated until out-of-fold classification scores and their corresponding ground-truth labels are obtained for all observations (Hastie et al., 2009)." (Sect. 2.5, lines 181–184)

> "To quantify uncertainty in model performance, 1000 paired bootstrap resamples of the LOOCV out-of-fold prediction records were used to calculate percentile 95% confidence intervals for AUC, TPR, FPR, and TSS, as well as for the differences in AUC and TSS between the SMLM and MOLM." (Sect. 2.5, lines 205–207)

Sect. 3.3 and Table 2 report the overall discrimination, uncertainty in threshold-dependent classification performance, and confusion-matrix counts. Together, these results quantify the additional discriminatory information provided by pre-event snow depth and show how its contribution differs between dry- and wet-snow avalanches. More details are shown in Sect. 3.3 (lines 306–312 and 318–324) and Table 2 (lines 326–329) of the revised manuscript:

> "To assess the discriminative capacity of the SMLM, we compared its overall classification performance with that of the baseline MOLM using the LOOCV-derived predictions. The ROC curve analysis indicates that the SMLM consistently achieves better discrimination than the MOLM for both dry- and wet-snow avalanche scenarios (Fig. 10). Specifically, for dry-snow avalanches, the AUC increased from 0.72 (95% CI: 0.56–0.86) for the MOLM to 0.83 (95% CI: 0.71–0.94) for the SMLM. For wet-snow avalanches, the AUC increased from 0.85 (95% CI: 0.76–0.92) for the MOLM to 0.94 (95% CI: 0.88–0.99) for the SMLM. Paired bootstrap comparisons yielded AUC differences (SMLM − MOLM) of 0.11 (95% CI: 0.01–0.24) and 0.09 (95% CI: 0.02–0.17) for dry- and wet-snow avalanches, respectively." (Sect. 3.3, lines 306–312)

> "The SMLM demonstrates a more effective trade-off between false alarms and missed events at the selected classification thresholds (Table 2). For wet-snow avalanches, the MOLM yields an FPR of 0.21 and a precision of 0.56. In contrast, the SMLM reduces the FPR to 0.13 while increasing precision to 0.71 and TPR from 0.79 to 0.92. For dry-snow avalanches, the FPR decreased from 0.41 to 0.36, while precision increased from 0.43 to 0.46. The TPR remained unchanged at 0.92. TSS increased from 0.58 to 0.79 for wet-snow avalanches and from 0.51 to 0.56 for dry-snow avalanches. The paired bootstrap TSS differences

(SMLM − MOLM) were 0.21 (95% CI: 0.01–0.38) for wet-snow avalanches and 0.05 (95% CI: −0.08–0.19) for dry-snow avalanches." (Sect. 3.3, lines 318–324)

Table 2. Comparison of classification performance metrics between the MOLM and SMLM. (Table 2, lines 326–329)

	Wet MOLM	Dry MOLM	Wet SMLM	Dry SMLM
TPR (95% CI)	0.79 (0.64–0.95)	0.92 (0.75–1.00)	0.92 (0.77–1.00)	0.92 (0.75–1.00)
FPR (95% CI)	0.21 (0.12–0.31)	0.41 (0.25–0.57)	0.13 (0.06–0.21)	0.36 (0.20–0.51)
TSS (95% CI)	0.58 (0.39–0.77)	0.51 (0.27–0.71)	0.79 (0.64–0.91)	0.56 (0.33–0.77)
TNR	0.79	0.59	0.88	0.64
FNR	0.21	0.08	0.08	0.08
Precision	0.56	0.43	0.71	0.46
F1	0.66	0.59	0.80	0.62
TN/FP/FN/TP	57/15/5/19	23/16/1/12	63/9/2/22	25/14/1/12

Sect. 4.4 acknowledges the scope of the available observations and clarifies that model applicability beyond the evaluated dataset remains to be established. It also identifies validation across additional snow seasons as a direction for future research. More details are shown in Sect. 4.4 (lines 392–396 and 399–400) of the revised manuscript:

"The improved discrimination achieved by the SMLM highlights the value of combining pre-event snow depth with meteorological forcing, while the scope of the available observations places some limits on its broader application. The two-season record and the monitoring stations in Galongla Valley provide valuable local evidence, but do not encompass the full variability of snowpack and meteorological conditions across southeastern Tibet. Moreover, because model performance was assessed using LOOCV, its applicability to other locations and snow seasons remains to be established." (Sect. 4.4, lines 392–396)

"Future research should focus on extending the observational record and examining model performance across a wider range of locations and snow seasons." (Sect. 4.4, lines 399–400)

4. Interpretation and operational scope of snow depth

I support replacing "background snow depth" with "pre-event snow depth" or "prior snow depth." Snow depth is an observable state variable, but it does not represent snowpack conditions in the broader sense because it contains no information on stratigraphy, weak layers, liquid-water content, or stability.

Its association with recorded avalanches may also reflect avalanche size, terrain roughness, the ability of an avalanche to reach the monitored corridor, season, elevation, or detection conditions. These alternative interpretations should be discussed, and the mechanical interpretation in Sect. 4.1 should be moderated.

The study is retrospective and internally validated. The manuscript states that predictors for positive samples are calculated up to the avalanche occurrence time. Please clarify how the model would be applied prospectively to an ongoing snowfall or snowmelt process, including how the observation window would be defined before the outcome is known. Until prospective validation is available, terms such as "candidate forecasting model" or "retrospective classification model" may be more appropriate than claims of demonstrated operational forecasting performance.

Response:

Thank you for this thoughtful comment. We agree that "pre-event snow depth" is more precise than "background snow depth", as it refers to the observed snow depth at the onset of the identified snowfall or snowmelt process, rather than implying a comprehensive description of snowpack conditions. We have adopted this terminology consistently throughout the revised manuscript.

Regarding the interpretation of snow depth, we have moderated the discussion in Sect. 4.1 to distinguish the observed statistical association from a specific physical mechanism. The revised discussion retains the positive association between pre-event snow depth and the model-derived relative likelihood of recorded avalanche occurrence, while placing this finding in the context of antecedent snow cover, terrain roughness, and previous studies of avalanches affecting transportation corridors. We recognize that recorded avalanche occurrence can also be influenced by event characteristics, local terrain and seasonal conditions, and whether an avalanche reaches and is detected along a monitored corridor. These factors provide important context for interpreting the observed association, although their individual contributions cannot be isolated using the present dataset. Accordingly, we interpret the snow-depth effect as an empirical association under the conditions represented

by our observations, without attributing it to a particular release mechanism. This qualification preserves the observed contribution of pre-event snow depth while keeping the physical interpretation within the scope of the available evidence.

Regarding prospective application, the present study is retrospective because the predictors for positive samples were calculated up to the recorded avalanche occurrence time, as described in Sect. 2.4. In a prospective setting, evaluation times and a forecast horizon would need to be defined in advance. At each evaluation time, the current qualifying snowfall or snowmelt process would be identified using predefined criteria. Its intensity and duration would be calculated only from observations available up to that time, while pre-event snow depth would be taken from the onset of the process. Avalanche occurrence would then be assessed over the predefined subsequent period.

With respect to operational scope, we agree that the current results establish internal discriminatory ability rather than validated operational forecasting performance. We have retained the study's broader forecasting motivation while framing the SMLM as an observation-based approach to distinguishing avalanche-event and non-event conditions. Sect. 4.4 acknowledges the limited spatial and seasonal coverage of the observations and the need to evaluate applicability across additional locations and snow seasons. The principal contribution is that measured pre-event snow depth provides additional discriminatory information beyond snowfall or snowmelt intensity and duration, with a clearer improvement in threshold-dependent classification for wet-snow avalanches. This provides local empirical support for further developing and prospectively testing locally adapted avalanche assessment and early-warning methods.

Changes in the revised manuscript:

Sect. 2.3 consistently uses “pre-event snow depth” and defines SD using the snow depth observed at the onset of the identified snowfall or snowmelt process. More details are shown in Sect. 2.3 (lines 116–118 and 132–135) of the revised manuscript:

> “Using hourly records, snowfall duration (SFD) was calculated as the duration of this event segment, and snowfall intensity (SFI) as the corresponding cumulative snowfall divided by SFD, representing the mean rate of snow loading. In parallel, the snow depth recorded at the onset of the event was defined as pre-event snow depth (SD).” (Sect. 2.3, lines 116–118)

> “Snowmelt duration (SMD) was calculated from hourly records as the duration of this event segment, whereas snowmelt intensity (SMI) was calculated as the corresponding cumulative snow-depth loss divided by SMD. Similarly, the snow depth recorded at the onset of the event was defined as pre-event snow depth (SD).” (Sect. 2.3, lines 132–135)

Sect. 4.1 interprets the positive association of pre-event snow depth with recorded avalanche occurrence in the context of antecedent snow cover, terrain roughness, and previous studies, without attributing the observed association to a single release mechanism. More details are shown in Sect. 4.1 (lines 333–340) of the revised manuscript:

> “Avalanche activity is influenced by the combined effects of meteorological forcing and antecedent snow cover under different terrain conditions. The model results show that SD, representing pre-event snow depth, is positively associated with the model-derived relative likelihood of avalanche occurrence for both dry- and wet-snow avalanches. This finding is consistent with previous studies indicating that sufficient pre-existing snow cover can reduce the influence of terrain roughness on avalanche formation and affect the amount of new snow required for avalanche release (Schweizer et al., 2003a). Previous investigations of large avalanches affecting transportation corridors have also identified antecedent snow depth as a relevant factor in avalanche forecasting (Schweizer et al., 2009). These findings highlight the value of considering pre-event snow depth together with meteorological forcing when assessing recorded avalanche occurrence.” (Sect. 4.1, lines 333–340)

Sect. 4.2 presents the contribution of measured pre-event snow depth as additional discriminatory information and positions the SMLM as an observation-based approach for further local method development. More details are shown in Sect. 4.2 (lines 374–378) of the revised manuscript:

> “In contrast to approaches that rely on detailed snowpack simulations, our results show that directly observed pre-event snow depth can provide additional discriminatory information when combined with forcing intensity and duration. By using a small set of readily available variables, the SMLM offers an observation-based approach to distinguishing avalanche-event and non-event conditions in data-limited mountain regions, providing a basis for the further development of locally adapted avalanche assessment methods.” (Sect. 4.2, lines 374–378)

Sect. 4.3 interprets model-score patterns as information for developing locally adapted warning criteria. More details are shown in Sect. 4.3 (lines 380–387) of the revised manuscript:

> "The SMLM provides continuous classification scores based on pre-event snow depth and the intensity and duration of snowfall or snowmelt processes. These scores allow the model response to different combinations of predictors to be examined beyond a single binary classification threshold. For illustrative purposes, the model-derived scores were divided into five equal-width intervals (Fig. 11). The resulting three-dimensional patterns illustrate how model scores vary with meteorological forcing and pre-event snow depth, and how different combinations of predictors can produce similar scores. Higher scores indicate predictor combinations more strongly associated with recorded avalanche events within the sampled data. These patterns provide an interpretable quantitative basis for identifying conditions associated with avalanche occurrence and informing the development of locally adapted avalanche early-warning criteria." (Sect. 4.3, lines 380–387)

Sect. 4.4 acknowledges the limits of the two-season record and internal LOOCV evaluation, and identifies broader validation as a necessary step towards future application. More details are shown in Sect. 4.4 (lines 392–396 and 399–404) of the revised manuscript:

> "The improved discrimination achieved by the SMLM highlights the value of combining pre-event snow depth with meteorological forcing, while the scope of the available observations places some limits on its broader application. The two-season record and the monitoring stations in Galongla Valley provide valuable local evidence, but do not encompass the full variability of snowpack and meteorological conditions across southeastern Tibet. Moreover, because model performance was assessed using LOOCV, its applicability to other locations and snow seasons remains to be established." (Sect. 4.4, lines 392–396)

> "Future research should focus on extending the observational record and examining model performance across a wider range of locations and snow seasons. Incorporating snowpack stability indices or mechanistic constraints within the existing statistical structure could strengthen the physical interpretation of the identified relationships. Furthermore, spatially distributed meteorological and topographical information could help account for local environmental differences and support the adaptation of this observation-based approach to other mountain regions. Together, these developments would provide a basis for more widely applicable, locally adapted avalanche assessment methods." (Sect. 4.4, lines 399–404)

Minor comments

1. Please provide concise summary statistics and distributions of the predictors for events and non-events, together with their units and correlations.

Response:

Thank you for this suggestion. Figure R1 presents the distributions of predictor variables for avalanche-event and non-event samples. It shows the distributions of SFI, SFD, and SD for dry-snow samples and SMI, SMD, and SD for wet-snow samples, with the corresponding units indicated on the axes. Table R2 provides the sample sizes and medians [Q1–Q3] for all predictors. Across the dry-snow dataset, Pearson's correlation coefficients were −0.03 for SFI–SFD, 0.39 for SFI–SD, and 0.06 for SFD–SD. Across the wet-snow dataset, the corresponding values were −0.31 for SMI–SMD, 0.48 for SMI–SD, and 0.003 for SMD–SD.

For both avalanche types, event samples had higher median forcing intensity and pre-event snow depth than non-event samples, but the distributions overlapped; these descriptive statistics therefore do not identify a single predictor that completely separates the two groups.

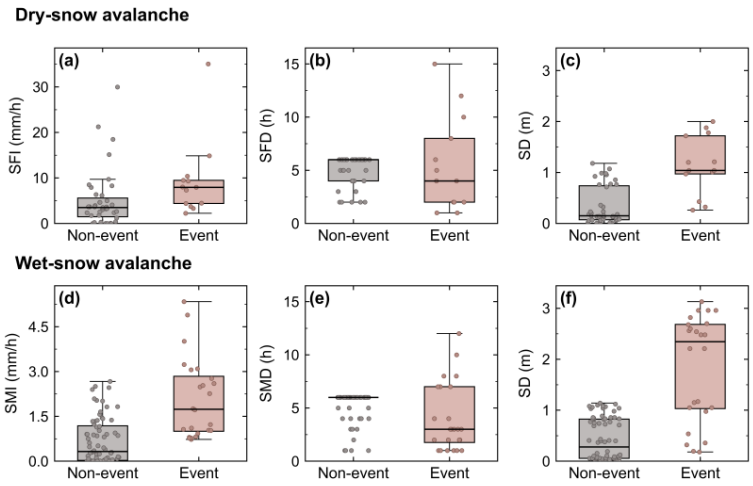


Figure R1. Distributions of predictor variables for avalanche events and non-events. Panels (a–c) show SFI, SFD, and SD for dry-snow samples, and panels (d–f) show SMI, SMD, and SD for wet-snow samples. Boxes indicate the interquartile range, horizontal lines indicate the median, whiskers extend to the most extreme observations within 1.5 times the interquartile range of the lower and upper quartiles, and points represent individual observations.

Table R2. Summary statistics of predictor variables for avalanche events and non-events.

Avalanche type	Sample	n	Rate (SFI/SMI; mm h^{-1})	Duration (SFD/SMD; h)	SD (m)
Dry snow	Event	13	7.94 [4.39–9.46]	4.00 [2.00–8.00]	1.04 [0.97–1.72]
	Non-event	39	3.49 [1.46–5.58]	6.00 [4.00–6.00]	0.15 [0.07–0.74]
Wet snow	Event	24	1.74 [1.00–2.84]	3.00 [1.75–7.00]	2.35 [1.03–2.69]
	Non-event	72	0.32 [0.03–1.18]	6.00 [6.00–6.00]	0.28 [0.05–0.82]

Note: Values are medians [Q1–Q3], where Q1 and Q3 denote the 25th and 75th percentiles, respectively.

2. Some events appear to have zero melt intensity. Please clarify how a zero melt intensity can occur in samples defined as snowmelt events.

Response:

Thank you for pointing this out. We confirmed that SMI = 0 occurs only in non-event samples; all wet-snow avalanche-event samples have SMI > 0. Candidate snowmelt-related processes were identified from hourly snow-depth data derived from the original 1 min observations, including very small decreases in snow depth. Negative samples were then restricted to intervals of up to 6 h, with SMI calculated over the retained intervals. Consequently, some retained intervals had snow-depth decreases too small to be represented at the reported precision, resulting in SMI = 0. We have clarified this distinction in Sect. 2.4 by referring to candidate snowfall or snowmelt-related process blocks rather than "event blocks" in the sample-selection description.

Changes in the revised manuscript:

The terminology in Sect. 2.4 was clarified to distinguish candidate meteorological process blocks from avalanche-event samples. More details are shown in Sect. 2.4 (lines 140–145) of the revised manuscript:

"For each recorded dry- or wet-snow avalanche, candidate snowfall or snowmelt-related process blocks were identified within a 3-d retrospective window at the monitoring station geographically closest to the avalanche location using the same process-identification criteria described above. The most recent qualifying process block preceding the recorded avalanche time was treated as the positive sample, with all variables calculated only up to the avalanche occurrence time. Earlier process blocks within the same retrospective window were considered candidate negative samples." (Sect. 2.4, lines 140–144)

"Candidate negative sample intervals were limited to a maximum duration of 6 h, even when the underlying processes lasted longer." (Sect. 2.4, lines 144–145)

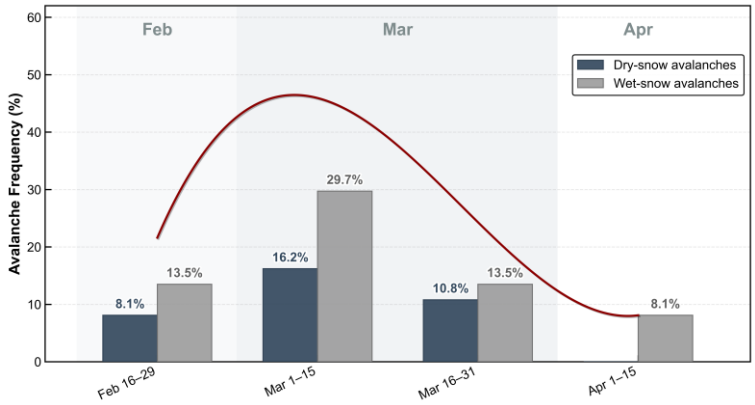
3. Please provide additional information in the caption of Fig. 6 and explain the meaning of the red curve.

Response:

Thank you for requesting this clarification. The red curve in the original Fig. 6 (Fig. 7 in the revised manuscript) is a cubic-spline interpolation of the combined dry- and wet-snow avalanche frequencies calculated for each half-month period. The bars show the observed frequencies of dry- and wet-snow avalanches separately, whereas the red curve provides a visual representation of the seasonal variation in their combined frequency. We have expanded the figure caption to clarify the calculation and meaning of the red curve. The interpolated curve is a visual guide to the observed seasonal pattern, not an independent prediction of avalanche frequency. The revised Fig. 7 and its complete caption are reproduced below.

Changes in the revised manuscript:

The caption of Fig. 7 was expanded to specify what the bars represent and to clarify that the red curve interpolates the combined frequencies of dry- and wet-snow avalanches. More details are shown in Fig. 7 (lines 229–230) of the revised manuscript:



"Figure 7. Seasonal distribution of avalanches. The bars show the frequencies of dry- and wet-snow avalanches in each half-month period, and the red curve represents a cubic-spline interpolation of their combined frequencies." (Fig. 7, lines 229–230)

4. Statements that snow depth has the "largest coefficient magnitude" are not directly meaningful when predictors have different units and scales. Standardised effects or odds ratios for meaningful increments would be more informative.

Response:

Thank you for pointing out this important issue regarding the interpretation of regression coefficients. We agree that raw coefficient magnitudes should not be used directly to compare predictors measured on different scales. We have therefore removed the statements referring to the SD coefficient as having the "largest magnitude" and no longer rank predictors by their raw coefficients. Although the predictors are transformed using ln(1 + x), this transformation does not standardise their scales. Rather than replacing the removed coefficient-magnitude comparison with a separate ranking based on standardised effects or odds ratios, we now interpret the coefficients in terms of the direction and uncertainty of their conditional associations and assess the incremental discriminatory value of adding SD through the comparison between SMLM and MOLM. This model comparison showed that adding SD increased LOOCV AUC by 0.11 for dry snow (95% CI: 0.01–0.24) and 0.09 for wet snow (95% CI: 0.02–0.17).

Changes in the revised manuscript:

Sects. 3.2.1 and 3.2.2 no longer rank SD by coefficient magnitude and instead describe its positive conditional association based on the confidence intervals. More details are shown in Sect. 3.2.1 (lines 241–243) and Sect. 3.2.2 (lines 280–282) of the revised manuscript:

"Its confidence interval lies entirely above zero, indicating that greater pre-event snow depth is associated with a higher model-derived classification score when the other variables are held constant." (Sect. 3.2.1, lines 241–243)

"Its confidence interval lies entirely above zero, indicating that greater pre-event snow depth is associated with a higher model-derived classification score when SMI and SMD are held constant." (Sect. 3.2.2, lines 280–282)

5. In the abstract, "pre-event snowpack conditions provides" should read "pre-event snowpack conditions provide." More generally, "pre-event snow depth" should replace "pre-event snowpack conditions" wherever SD is the only snow-related predictor.

Response:

Thank you. We have corrected the Abstract to refer to the measured variable as "pre-event snow depth," with the singular verb "provides," and standardized the terminology throughout the manuscript wherever SD is the only snow-related predictor. Broader terms such as "snowpack conditions" are retained only when referring to the general physical context, rather than to SD as the measured predictor.

Changes in the revised manuscript:

In the Abstract, we use "pre-event snow depth" to identify the measured predictor and use the singular verb "provides" accordingly. More details are shown in the Abstract (lines 18–20) of the revised manuscript:

> "These results indicate that pre-event snow depth provides additional discriminatory information beyond meteorological forcing alone, particularly for wet-snow avalanches, and supports the development of locally adapted avalanche assessment methods." (Abstract, lines 18–20)

In the Introduction, we explicitly define the snow-depth variable used in the models. More details are shown in the Introduction (line 66) of the revised manuscript:

> "Pre-event snow depth is defined as the snow depth at the onset of the preceding snowfall or snowmelt process." (Introduction, line 66)

Recommendation

The study is based on a valuable observational effort in an underrepresented mountain region. I would encourage the authors to revise the manuscript, with particular attention to the construction of positive and negative samples, the interpretation of probabilities under the sampling design, the interpretation of the forcing-duration coefficients, the limitations associated with the small dataset and short observation period, and the operational interpretation of snow depth. Addressing these points would substantially strengthen the scientific interpretation and potential practical relevance of the work.

Response:

Thank you for this overall assessment and the constructive recommendations. We have addressed the specific points in our responses above and revised the manuscript accordingly.