

Major comments:

Selection of key parameters for the study is introduced without further detail or sufficient explanation/citation. Some examples: i) Buffer of 2000 m defining the Bilanz-gebiete; ii) The fixed interval of 750 m for the PRAs; iii) Runoff coefficients for the accumulation and ablation zones; iv) Threshold value of $P_{df} > 50\%$. Please briefly explain or cite accordingly.

We propose to update the manuscript giving explicit justification for the parameter values as mentioned. These are:

- i) The 2000 m buffer around the Bilanzgebiete is, in principle, not required because the Bilanzgebiete represent self-contained catchments with their outlets located in gently sloping valley terrain. Nevertheless, we included a 2000 m buffer as a conservative measure to ensure that unlikely runoff beyond the delineated catchment boundaries is captured. The choice of this distance follows a similar assumption applied to (hydrologically non-contained) study areas by Van Zadelhoff et al. (2022). We will clarify this rationale in the revised manuscript.
- ii) In accordance with our response to Reviewer 1, the 750 m interval was selected empirically as a pragmatic compromise between computational burden and the need to capture spatial heterogeneity in potential release area properties along individual channels. We will explicitly state this rationale in the revised manuscript.
- iii) We assume that the glacier ablation zone responds hydrologically in a manner similar to non-glaciated areas, i.e. direct runoff, and therefore assign it the default runoff coefficient of 0.1. In contrast, the accumulation zone is treated as a hydrological buffer, where precipitation is assumed not to contribute directly to runoff (within the short temporal scale of debris flows). We therefore assign the accumulation zone a runoff coefficient of 0.0. We will make these assumptions and their implications more explicit in the revised manuscript.
- iv) The debris-flow probability is obtained by applying the trained random forest model to all potential release areas. The resulting probability ranges from 0 to 100%, where values approaching 0% indicate that an event is more likely to be classified as a flood and values approaching 100% indicate that it is more likely to be classified as a debris flow, according to the event categories used for model training in StorMe. We selected $P_{df} > 50\%$ as the threshold because, above this value, the model assigns a higher probability to the debris-flow class than to the alternative flood class. We will state this rationale explicitly in the revised manuscript.

Consistency between the parametrization of μ and ξ in the Voellmy model: The friction coefficient μ is spatially distributed from fan data, but ξ is kept constant at the national scale. Since RAMMS depends on both parameters jointly, please explain why a constant value and the reason behind this specific value.

In our implementation of the Voellmy model, we deliberately parameterize μ spatially while keeping ξ constant across the national domain. This choice limits computational complexity, facilitates interpretation of spatial variations in μ , and reduces the risk of equifinality that would arise from introducing an additional spatially varying parameter. Furthermore, this approach is supported by Bolliger et al. (2024), who found that ξ has a comparatively limited influence on debris-flow runout relative to μ . We will revise the manuscript to explain this modelling choice and the rationale for the constant ξ value more explicitly.

Please clarify in the discussion section why the output is framed as a hazard indication rather than a hazard map. I follow the intent to not introduce legal constraints directly, but this should be explicitly stated.

It indeed hangs on legal constraints. To be a hazard map, expert-knowledge and local validation is required. We propose to indicate this more clearly in the discussion.

L259-L265: Deriving μ spatially from data sourced is a good idea. However, two clarifications are needed: i) Which criterion was applied with “after visual inspection of results”?, and ii) Why the 5th percentile and not the 10th or the 25th?

i) The visual inspection consisted of two complementary comparisons. Firstly, we compared the resulting μ values with values reported in previous local RAMMS applications. For example, values of approximately 0.05

have been applied at Illgraben (Canton Valais) (Berger et al., 2011; Bolliger et al., 2024), whereas a value of 0.20 was used at Spreitgraben (Canton Bern) (Frank et al., 2015). However, direct comparison with these values is limited because previous studies differ in model parameterization, spatial resolution, and application context. Secondly, we qualitatively compared simulated RAMMS runout at several well-documented sites with the spatial extent of events represented in long-term inventories. Long inventories provide an approximate indication of the expected extent of high to medium frequency events (e.g., 30- to 100-year return period).

We acknowledge that this procedure relies on qualitative visual assessment rather than a formal quantitative calibration and therefore introduces uncertainty into the selection of μ . We will describe the evaluation criterion and this limitation more explicitly in the revised manuscript.

ii) We evaluated several candidate percentiles (50th, 30th, 10th, and 5th) using the procedure described above. Of these, the 5th percentile provided the most suitable representation when considering both previously reported RAMMS parameter values and the qualitative runout comparison at well-documented sites. We will expand the methods section to describe this testing procedure and the rationale for selecting the 5th percentile.

Use of the 10 km moving window: I follow the intent, but the choice of value should be stated as well as the authors should acknowledge that this smoothing trades local variability of μ for larger spatial assessments.

We fully agree and we will adapt the manuscript to explain our choice for the 10-km moving window and its consequences explicitly. This is indeed a compromise between local variability and preventing outlier influence.

L275: The decision criteria used to define the erosion categories are not given, and it is not clear how the erosion parametrization was implemented in RAMMS nor the obtained values presented in Tables A4/A5.

The RAMMS erosion module includes five parameters (RAMMS, n.d.):

1. Erosion density [kg/m^3]: Density of the erodible material.
2. Erosion rate [m/s]: The rate at which the debris flow entrains material from the bed.
3. Potential erosion depth [kPa]: Potential erosion depth as a function of applied shear stress.
4. Critical shear stress [kPa]: Shear stress threshold above which erosion can commence.
5. Maximum erosion depth [m]: Maximum thickness of erodible material.

Given the national scale of this study, a spatially detailed parameterization of all five erosion parameters is not feasible, and several simplifications are therefore required. First, we assume a spatially uniform density of erodible material (parameter 1). Second, parameters 2–5 have their value assigned according to four erosion classes: no erosion, low, medium, and high. The classification scheme is presented in Table 1 and is based on geological characteristics, particularly the Altschutt/Jungschutt concept, in combination with topographic classes.

The translation from these erosion classes to RAMMS parameter values is provided in Table A4 for parameters 2–4 and Table A5 for parameter 5. The default RAMMS erosion parameterization is based primarily on observations from the Illgraben catchment. To our knowledge, no established framework currently exists for representing spatially heterogeneous erosion conditions through the RAMMS erosion parameters at a regional let alone national scale. We therefore developed a simplified classification scheme for this study.

We use the RAMMS default parameterization to represent the medium erosion class and define the low and high classes relative to this reference. Based on exploratory model testing, the parameter values for the low erosion class are set to half of the default values, while those for the high erosion class are set to twice the default values. Areas classified as no erosion are assigned a potential erosion depth of zero, thereby preventing entrainment.

We acknowledge that this classification represents a substantial simplification of spatially variable erosion processes and that the scaling factors used to distinguish the erosion classes are empirically defined rather than independently calibrated. We propose to revise the methods section to describe these assumptions and parameter choices more explicitly and will discuss the associated uncertainties and limitations in the discussion.

Figure 6: A national-scale visualization of the flow-height is understandably difficult. However, the authors should motivate the specific examples shown, i.e., if they were chosen to represent contrasting geological, geomorphological or other environmental settings.

The three areas were indeed chosen to represent distinct environments which we state in table 3. We will update the text to explicitly motivate this choice before presenting any results.

Minor comments:

L276: define Altschutt/Jungschutt for the debris cover in the text, using the definitions in Table 1.

We will update the text applying the Altschutt/Jungschutt definition from the table explicitly in the text.

Section 6.2: If the Illgraben catchment is used as a representative, well-known site, state this before Figures 4 and 5.

It is indeed for recognizability in debris flow research that we choose this site as example. We will update the text as suggested to state this choice explicitly.

Section 6 partially interprets the results. Consider renaming it “Results and interpretation”.

We agree and will update the section name.

L340-L341: this would strengthen the outcome of this study: “Since then these approaches have been applied in large-scale modeling of rockfall (Bründl et al., 2025) and debris flow (this paper)” reviewed in “[...]. The results of this study show that a national application to debris-flow hazard can also be achieved and provide general indications for further detailed studies”.

We appreciate the suggestion to stronger state the relevance and outcome of our study and would like to incorporate this in the manuscript.

Technical corrections:

References are sometimes ordered chronologically and sometimes alphabetically; adopt one criterion and keep it consistent.

Thank you for bringing this to our attention. We will update and order all references in the text chronologically.

Equations should be referred to by number in the text, e.g., (1).

We will update the manuscript and refer to all equations in the text by number.

Ensure all variables in the equations are defined in the text before or after each equation.

We will check for this issue and update the text around the relevant equations.

L260 typo: “which is in turn is approximated”.

We will correct this.

Add subfigure labels (e.g., Figure 5a-c) and refer to them in the text - Figures 6 and 8 would benefit.

We will update the relevant figures with subfigure labels.

L285: Cite the reference work used to justify omitting the flow height values.

Flow heights < 0.01 were omitted from the analysis to exclude negligible modelled flow depths and numerical spreading at the margins of the simulated runout. The threshold was chosen in accordance with other applications of RAMMS (e.g., Hauser, 2011).

Apply abbreviations consistently after the first definition, eg. PRA in L347.

We will update L347 and check the manuscript for similar abbreviation issues.

Figure 7: Align labels for the three areas (Emmental, Tessin and Wallis) with Figure 6 (Emmental, Ticino and Valais).

That was overlooked, we will update the labels in the figure.

References:

- Berger, C., McArdell, B. W., & Schlunegger, F. (2011). Direct measurement of channel erosion by debris flows, Illgraben, Switzerland. *Journal of Geophysical Research: Earth Surface*, 116(F1), F01002. <https://doi.org/10.1029/2010JF001722> AGU Journals
- Bolliger, D., Schlunegger, F., and McArdell, B. W.: Comparison of debris flow observations, including fine-sediment grain size and composition and runout model results, at Illgraben, Swiss Alps, *Natural Hazards and Earth System Sciences*, 24, 1035–1049, 455 <https://doi.org/10.5194/nhess-24-1035-2024>, 2024
- Bründl, M., Ortner, G., Bühler, Y., Bebi, P., Christen, M., Kissling, E., Michel, A., Spieler, M., Völk, M.-S., & Zaugg, A. (2025). TASK 3.1 – Multi-hazard risk assessment framework. In A. Bast, M. Bründl, & M. Phillips (Eds.), *WSL research programme Climate Change Impacts on Alpine Mass Movements – CCAMM: Project report* (pp. 87–93). WSL Berichte, 181. Swiss Federal Institute for Forest, Snow and Landscape Research WSL. <https://doi.org/10.55419/wsl:41905>
- Frank, F., McArdell, B. W., Huggel, C., & Vieli, A. (2015). The importance of entrainment and bulking on debris flow runout modeling: Examples from the Swiss Alps. *Natural Hazards and Earth System Sciences*, 15, 2569–2583. <https://doi.org/10.5194/nhess-15-2569-2015>
- Hauser, D. (2011). *Interaktion Murgang–Wald: Rekonstruktion von Ereignissen mit Hilfe von RAMMS* [Master's thesis, ETH Zürich]. <https://doi.org/10.3929/ethz-a-006455792>
- RAMMS. (n.d.). Erosion. RAMMS—Rapid Mass Movement Simulation. <https://ramms.ch/ramms-debrisflow/erosion/>
- van Zadelhoff, F. B., Albaba, A., Cohen, D., Phillips, C., Schaefli, B., Dorren, L., & Schwarz, M. (2022). Introducing SlideforMAP: A probabilistic finite slope approach for modelling shallow-landslide probability in forested situations. *Natural Hazards and Earth System Sciences*, 22, 2611–2635. <https://doi.org/10.5194/nhess-22-2611-2022>