



Assessing the predictability of rockfall simulations constrained to simple objective input parameters

François Noël^{1,2}, Synnøve Flugekvam Nordang³

¹Risk Analysis Group, Institute of Earth Sciences, University of Lausanne, 1015 Lausanne, Switzerland

5 ²Landslides and Avalanches Section, The Norwegian Water Resources and Energy Directorate, 7030 Trondheim, Norway

³Mitigations Section, The Norwegian Water Resources and Energy Directorate, 7030 Trondheim, Norway

Correspondence to: François Noël (fran@nve.no)

10 **Abstract.** Rockfall simulations are used to predict runout distances in case potentially unstable rock compartments would eventually fail. Transit simulated values such as the bounce heights and involved energies are also useful for hazard and risk assessments and for mitigation design tasks. However, it has been shown that the predictions from simulation results can vary significantly from user to user and from site to site. This highlights the need for simulation models with quantified accuracy and precision, low parametric subjectivity, and with good performance at predicting the transit values. The
15 objective of this work is to present a validation methodology for rockfall simulation models and to objectively evaluate the predictive performance of stnParabel freeware simulation model when used with the Rolling friction rebound model. For this purpose, numerous mapped observations from a combination of back analyses of rockfall experiments and real events involving different remote sensing techniques were gathered. They cover twelve sites of various characteristics and geometries. The extensive collected observations include several hundred mapped deposited rock fragments of known
20 dimensions and respective source locations. Each individual rock's dimensions and masses were repetitively simulated without any other parameter adjustments in order to minimize the subjectivity of the simulation approach. In complement to the systematic objective process-based simulations, the runouts were also predicted for all sites with simulated trajectories from two additional process-based models for comparisons. Moreover, runout extents were also obtained geometrically with a commercial software and with a common geometrical approach for comparison. The results showed that the runout
25 prediction accuracy from our process-based simulated trajectories is generally stable from site to site. Moreover, the runout precision of the simulations with stnParabel is improved by 2× to 3× compared to those of all other methods tested. And this is achieved with limited errors on the predicted transit values such as the bounce heights and translational velocities.



1 Introduction

Combined with field observations, rockfall simulations are commonly used to predict areas that could be reached by rockfall events for rockfall susceptibility and hazard assessments. Additionally, the simulations are used to predict quantities such as the rock fragment's velocities, energies, and bounce heights for rockfall hazard, risk, and mitigation assessments. However, it is difficult to correctly predict such quantities at the same time as estimating proper runout distances (Labieuse, 2004; Dorren et al., 2013; Lambert and Bourrier, 2013). It has been shown through extensive benchmarks that there is a high variability with the accuracy and precision of the predictions from user to user (Berger et al., 2004; Berger and Dorren, 2006; Berger et al., 2011; Garcia, 2019; C2ROP, 2020). Since this is also the case even if the same software is involved, a part of this problem can be attributed to the subjectivity associated with the sensitive parametrization of some models that must be calibrated on a per-site basis. This has been confirmed for some models in Noël et al. (2023b) by qualitatively evaluating their sensitivity by initially varying their parameters within reasonable values and seeing the effect on their early predictions on one site. Previous studies also confirm the need for most simulation models to calibrate the simulations on a per-site basis (Jones et al., 2000; Labieuse, 2004; Crosta and Agliardi, 2004; Berger and Dorren, 2006; Labieuse and Heidenreich, 2009; Berger et al., 2011; Volkwein et al., 2011; Dorren et al., 2013; Valagussa et al., 2015; Bourrier et al., 2021; Noël et al., 2021). This led Noël et al. (2023b) to formulate associated drawbacks as follows: "there is a significant limitation caused by the inconvenient requirement for calibration of the simulations on a per-site basis from on-site observations, back calculations, and back analyses". And to ease the unfortunate related workflow, back calculations and guiding hazard zoning tools were provided (Noël et al., 2023b). However, in some situations, it might not be possible to observe past rockfall deposits on-site for fine-tuning the simulation parameters due to farming or other anthropogenic activities. Having predictable, accurate, and precise simulation models would be particularly welcome in such situations.

Another contributor to the sensitive parametrization of some simulation models derived from Pfeiffer and Bowen (1989) could be attributed to the potential confusion related to the subjective dissipating (or damping) coefficient parameters R_N and R_T . Indeed, because they are sometimes called "coefficient of restitution", it is likely to confuse them with the apparent kinematic coefficient of restitution (COR) obtained empirically from rockfall experiments and reconstructed trajectories (Fig. 1). For example, the empirical COR from (e.g., Wu 1985, Azzoni et al., 1992; Chau et al., 2002; Beladjine 2007; Bourrier, 2008; Labieuse and Heidenreich, 2009; Paronuzzi, 2009; Mathon et al., 2010; Buzzi et al., 2012; Asteriou et al., 2012; Spadari et al., 2013; Wyllie, 2014a; Noël et al., 2018; Wang et al., 2018; Ji et al., 2019; Ye et al., 2019; Caviezel et al., 2021; Noël et al., 2022; Noël et al., 2023a) should not be directly interchanged as R_N or R_T . Even if parallels can be made between these, they are not the same as explained in (Noël et al., 2021; Noël et al., 2022, and Noël et al., 2023a). It is thus difficult to transpose the correlations made on the COR from the literature work to the desired R_N and R_T parameters. Additionally, the effect of these parameters may differ from software to software, even when using the same equations (Noël et al., 2021).

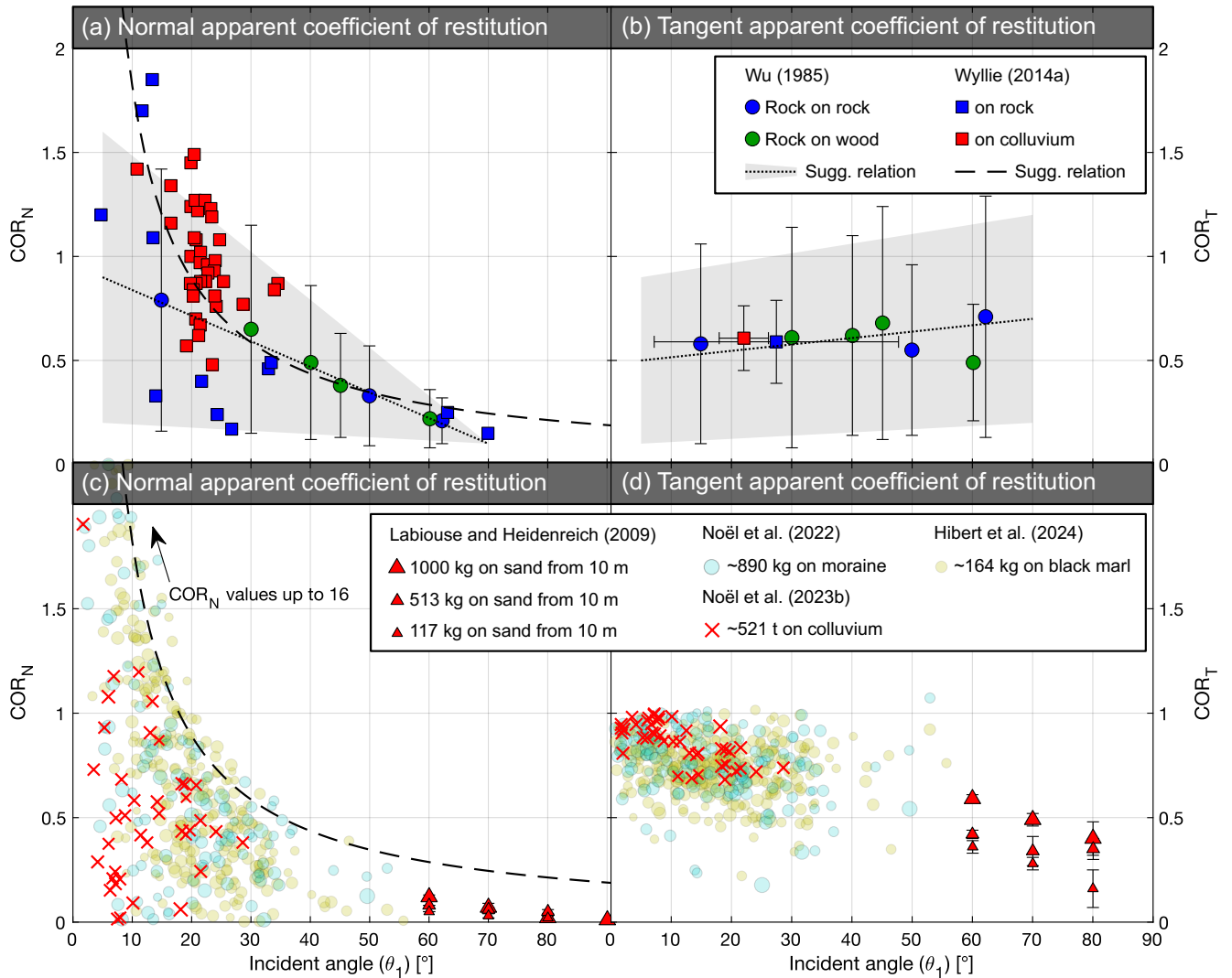


Figure 1: Compilation of observed velocity changes at impact. The changes are expressed as ratios of the returned velocities over the incident ones, i.e. expressed as apparent kinematic coefficient of restitution (COR). In a) and c), the changes are for the normal component of the velocities, and in b) and d), for the tangent component of the velocities. Note how the values are far from being constant and how the incident angle plays a major role in comparison to the impacted materials. Moreover, rebound models that calculate the rebounds solely based on their components could never reproduce the observed COR_N trend if they do not apply geometric deviations and use damping functions (e.g. based on fixed R_N and R_T parameters per terrain unit with values that cannot exceed one). The data points with error bars correspond to the mean values with their standard deviations for Wu (1985) and Wyllie (2014a). Modified from Noël et al. (2023a) in CC BY 4.0.

Moreover, Pfeiffer and Higgins (1990) and Jones et al. (2000) stated that the terrain geometry and the perceived roughness are the most important inputs for proper predictions of rockfall behaviours with their revised version of the Pfeiffer and Bowen (1989) model. Indeed, the much less significant dissipative effect of the subjective R_N and R_T parameters was obscured by the so much greater geometrical effect of the two previous inputs in their sensitivity analyses. Ondercin et al.



(2014) also noted that simulations are not sensitive to the material parameters given that they are performed with a detailed terrain model for properly accounting for the geometric aspect. With better predictions from the relatively objective geometrical parameters, one would assume that the models derived from Pfeiffer and Bowen (1989) would have focused on those geometrical parameters. However, some derived models seem to have put most emphasis on the subjective R_N and R_T dissipating parameters during the last three decades. In addition to degrading the predictive capacity of models, neglecting the geometrical roughness also leads to significant limitations when having to consider rockfall embankments and ditches (Lambert et al., 2013; Lambert and Bourrier, 2013). Others allow for randomly varying the slope angle artificially. But for those artificial slope angle variations, it is often omitted to mention that the user should adjust the added angle so that it matches what would be perceived by the rock size. To do so, the maximum perceived angle (θ_{max}) that should be added to the local terrain slope can be estimated in relation to the rock diameter (d) as follows (Pfeiffer and Bowen, 1989):

$$\theta_{max} = \tan^{-1} \frac{S_{\perp}}{0.5d} \quad (1)$$

where $S_{\perp}$ = perpendicular surface roughness [m]. The inaccuracy related to the omission of half of the most important inputs or to the imprecision on how to evaluate the artificial perceived surface roughness might have contributed to the variable results obtained at the previously mentioned benchmarks.

Additionally, setting fixed θ_{max} , R_N , and R_T parameters for numerous terrain units introduces a generous level of subjective freedom. This convenient flexibility comes at the expense of the fine-tuning simplicity and objectivity, however. Given that Wyllie (2014b) observed that rock, talus, and colluvium materials do not have significantly different apparent COR_N and COR_T , the use of numerous subjective terrain material units with fixed parameter values is difficult to justify (Fig. 1). For example, the 13 terrain units of the Rifle test site correspond to 39 terrain parameters that must be set and properly adjusted (Pfeiffer and Bowen, 1989; Pfeiffer and Higgins, 1990; Jones, 2000; Rocscience Inc. 2022; Noël et al., 2021). For that site, it has been shown that simplifying the terrain to only one unit did not produce significant change in the results (Noël et al., 2021). This came with the advantage of significantly reducing the number of terrain parameters to only three instead of 39. Similar trends were observed by Bourrier et al. (2009). Slightly better predictions were obtained from using a unique tangential dissipative parameter (R_m , related to R_T) per terrain unit than their three tangential dissipative parameters (Rg_{70} , Rg_{20} , and Rg_{10} , also related to R_T). Such simpler parametrization can potentially allow for a more efficient fine-tuning of the simulations due to the time saved when updating fewer parameters (Bourrier and Acary, 2022). And with finite time resources, the less time is spent with the simulation software to iteratively fine-tune the parameters and the more time remains to collect valuable observations from field work, to estimate the non-trivial temporal frequency aspects, and to reflect on the realism of the obtained predictions.

In the meantime, the sensitive, subjective, and complex per-site parametrization of some simulation models potentially leads to time-consuming assessments. The highly variable and sometimes inaccurate predictions may lead to a lack of homogeneity between delivered hazard maps, even for the same areas. Also, the uncertainty and related lack of confidence can potentially lead to overly conservative and constraining limits for the defined regulated levels of risk acceptance. With



these various issues stipulated, one realizes that an objective, accurate, and precise simulation model would be preferable. There is therefore a need for evaluating the predictive performance of rockfall simulations software (Dorren et al., 2013; Sarro et al., 2025) and assessing the sensitivity of their parametrization. This has been qualitatively evaluated for some simulation models in Noël et al. (2023b) by confronting their predictions against the mapped rockfall events from the Mel de la Niva Mountain described in Noël et al. (2023a). Also, recent developments related to RAMMS::ROCKFALL showed promising predictive capabilities related to the velocities, energies, and bounce heights when confronted against five reconstructed trajectories at the Chant Sura test site (Lu et al., 2019). Some rebound models have also been compared on a “per-impact” basis to reconstructed data in Noël et al. (2023a). Among them, inspired from Wyllie (2014b), our proposed Rolling friction model showed that its predictions are at least comparable to those of the other models and are in relatively good agreement with the “per-impact” observations. Moreover, this performance has been objectively obtained without any subjective adjustments, i.e., keeping its effective friction angle, cohesion, and dissipative rolling resistance coefficient unchanged.

The Rolling friction rebound model was then applied in this objective way on a “per-site” basis by Noël and Nordang (2025) to reproduce the rockfall events from the Mel de la Niva, similarly to how other simulation models were previously qualitatively compared (Noël et al., 2023b). To do so, the multi-model process-based simulation freeware stnParabel (Noël, 2020) built on our impact detection algorithm (Noël et al., 2021) was used with detailed 3D point cloud terrain models (DTM) for a proper geometrical terrain perception. Qualitatively, relatively good predictions were obtained (Noël and Nordang, 2025). The extent of these qualitative “per-site” validations was, however, limited due to being performed only at one site. Noël and Nordang (2025) thus pursued the validations to eleven additional sites. The Rolling friction rebound model was again tested with stnParabel (v. August 2021), as well as with a slightly adjusted impact detection algorithm to work with gridded DTMs, referred to as stnParabel raster. For additional context, rapid automatic simulations (r.a.sim.) from the Rockyfor3D (v5.2.15) process-based simulation software (Dorren, 2015) were also performed to the twelve sites. To complement the process-based simulated trajectories from stnParabel, stnParabel raster, and Rockyfor3D, maximal rockfall runout extents were also predicted geometrically with the Flow-R software (v2.1.0) (Horton et al., 2013), similar to its open-source analogue Flow-Py (D’Amboise et al., 2021; 2022), and with the Cone $\alpha - \beta$ approach (Derron et al., 2016; Keylock and Domaas, 1999). Validation results for the twelve sites were only preliminarily analyzed in the form of the quantified predictive capabilities of the tested models to reach the precisely mapped observations (Fig. 2). Supporting the Swiss Confederation recommendations about the use of rockfall trajectory simulations (Loup and Dorren, 2022), the process-based predictions performed the best at that metric, i.e. the predictions with the Rolling friction rebound model via stnParabel, stnParabel raster, and with Rockyfor3D. They were followed closely by the geometrically predicted runout extents from the Cone $\alpha - \beta$ approach, while the predictive performance of Flow-R showed to be relatively limited for that metric.

The metric of verifying that a modelling approach can predict runout extents from its capacity to cover precisely mapped rockfall deposits is largely sufficient for susceptibility assessments. However, it does not provide a comprehensive

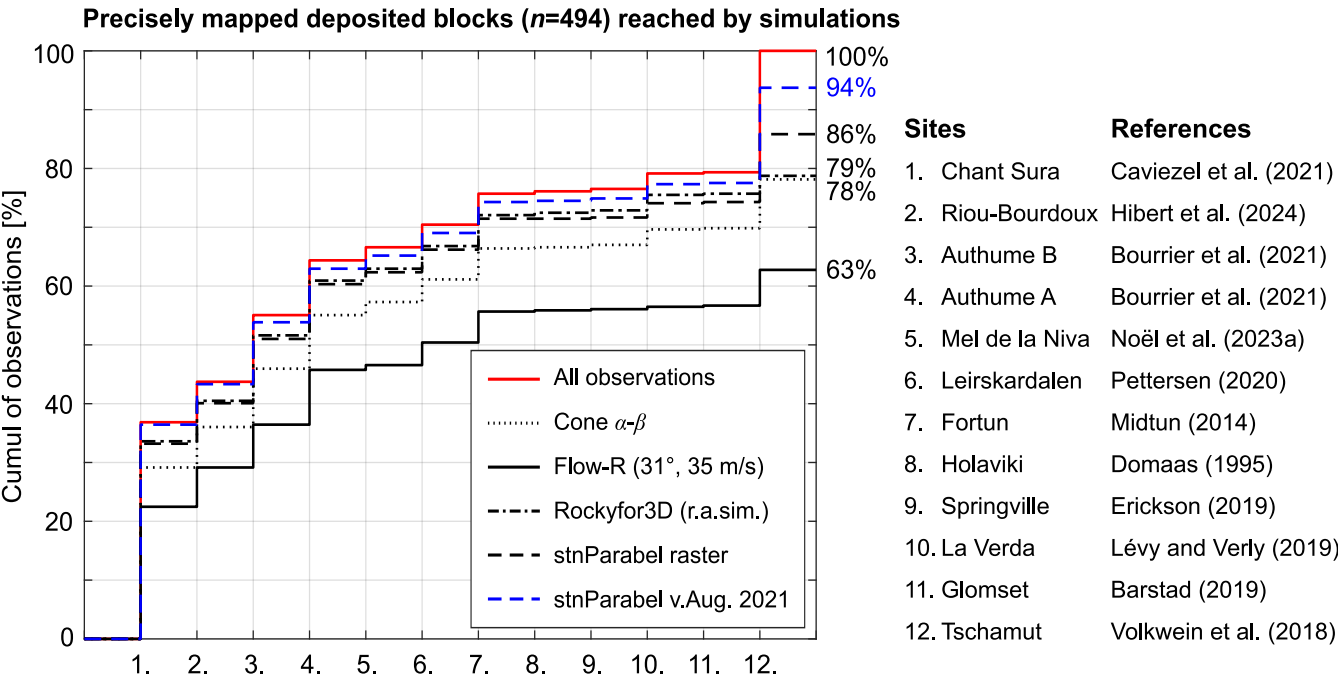


Figure 2: Preliminary validations by Noël and Nordang (2025) of the model’s capacities to reach observed rockfall deposits from rockfall experiments and real events mapped for several sites. Perfect predictions and predictions with exaggerating long-runout distances would cover most observations. Their performance at covering observed deposits would thus lie close to the red cumulative curve of all observations. Figure reproduced from Noël and Nordang (2025) in CC BY 4.0.

- 130 understanding of a method’s required predictive capabilities for more complex assessments, such as rockfall hazard and risk assessments or for designing mitigation measures. For these assessments, simulated rockfall statistics, including frequencies, bounce heights, and energies, must be spatially distributed appropriately. Therefore, the objectives of this paper are to enhance these qualitative and quantitative validations in an attempt to address the following related general questions:
1. How to evaluate the precision and accuracy of simulated predictions from models in terms of:
- 135
- Distribution of runout distances?
 - Distribution of lateral runout dispersion?
 - Bounce heights and translational velocities?
2. How to cope with the observational biases associated with limited mapped observations for some sites?
3. And above all, how precise and accurate are the predictions from our proposed objective simulation approach, which is
- 140 based on the Rolling friction rebound model with stnParabel?



2 Approach

As mentioned in Noël et al. (2023b), the extent of the objective predictability of a simulation model should be investigated on more than one site. For that purpose, the Rolling friction model was repetitively used with stnParabel to reproduce observed rockfalls at eleven additional sites (Noël and Nordang, 2025). Since these simulation results are analyzed further in this paper, the approach used to produce them is here summarized with the site characteristics. Further related details are given in Appendix A. The simulation results are also presented visually reusing the quantitative guiding hazard land use zonation concepts of Noël et al. (2023b) and Noël and Nordang (2025) based on Jaboyedoff et al. (2005), Abbruzzese and Labiouse (2020), and Hantz et al. (2021). For conciseness, methodological details regarding the complementary process-based simulations with stnParabel raster and Rockyfor3D are given in Appendix B. Those of the geometrical Flow-R and Cone $\alpha - \beta$ approaches are also given in Appendix B.

The sites cover a wide variety of slope geometries, sizes, and terrain materials to challenge the predictability of stnParabel to a large extent (Fig. 3). The Mel de la Niva site from Noël et al. (2023a, 2023b) is also shown next to the eleven other sites for context. To maximize the objectivity, only the rock fragment characteristics were adjusted for each site to match those of the observed rock fragments while everything else was kept to the default settings of stnParabel. The wide variety of sites' geometries, volumes, sizes, and materials involved can be separated into two categories: 1) rockfall test sites and 2) real rockfall event sites (Fig. 3 and Appendix A). For the first category, the number of well-documented rockfalls is high, the observation precision is high, and information related to the involved velocities and bounce heights is sometimes known. Indeed, the observations come from peer-reviewed, well-designed, and controlled rockfall experiments. The height differences from the sources to the deposited zones and the related involved rock volumes and masses are limited for the practicality of the rockfall experiments.

The second category contrasts with its real rockfall events, with generally greater height differences and larger involved volumes and masses. The quality of the observations is also variable depending on the available information used for mapping the rockfall events. Like the flexible trajectory reconstruction strategies employed to cope with non-optimal footage from witnessed rockfall events not initially captured with that purpose in mind (Noël et al., 2022), different remote mapping strategies were used from the available documentation of the rockfall events. As highlighted by Asteriou et al. (2025), structure from motion photogrammetry (SfM) is a powerful technique for such assessments when the available information allows. This was thus used for many sites to obtain orthophotos and 3D models of the source, transit, and deposit areas (e.g. Fig. 4). Terrain models were built iteratively, starting with the best pictures, and gradually adding those that could confuse the SfM algorithm if given initially. Low precision was set for the camera position metadata, when available, and ground control points were manually added during the iterative process from recognizable features on the pictures (e.g. roof corners and chimney, characteristic blocks, etc.) and with 3D coordinates from corresponding detailed DTM from airborne lidar data (ALS). The obtained 3D model positions and scales were further refined by using the ICP algorithm in CloudCompare from the reference ALS DTM (Girardeau-Montaut, 2006). On their end, the obtained orthophotos were further rectified in QGIS

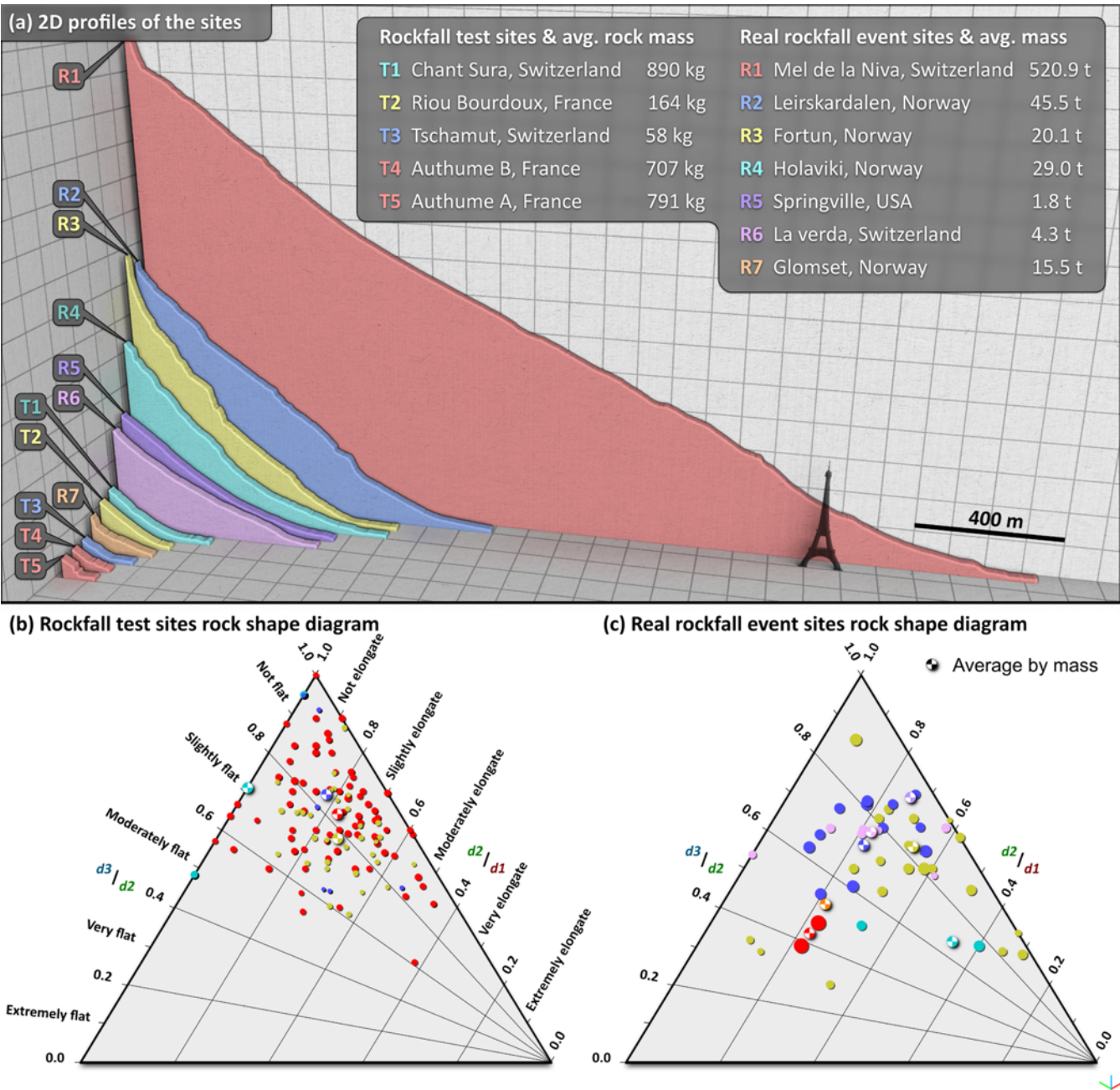


Figure 3: Comparison of the different terrain profiles of the covered sites and of their rock fragments' shapes. The terrain profile and rock fragments' shapes of the Mel de la Niva site covered in (Noël et al., 2023a; Noël et al., 2023b; Noël and Nordang, 2025) are also included for comparison.

to refine their correspondence with the existing orthophotos of the sites. A special attention has been put on the identification of the source areas (e.g. Fig. 4e and f). Indeed, it is imperative for proper validations that the initial conditions for the

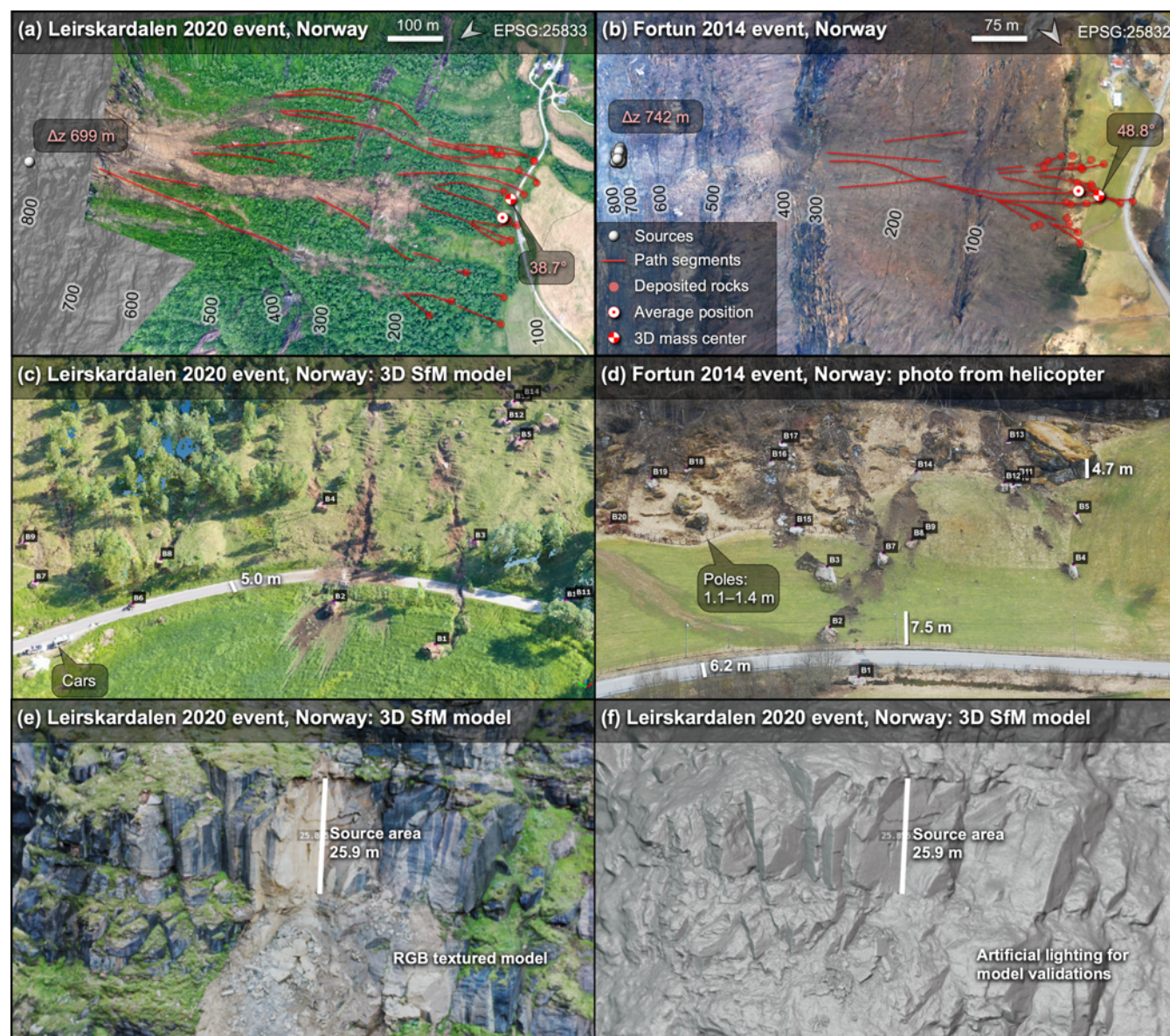


Figure 4: Examples of mapped source, rockfall path segments, and deposited rock fragments from orthophotos and SfM photogrammetric models. Across the paper, those types of observations are always shown in red to be consistent with the symbology used in Noël et al. (2023a; 2023b) and Noël and Nordang (2025). As shown in (f), digital artificial lighting and shaders were used to visually verify the quality of the 3D SfM models in CloudCompare and Blender open-source software. Drone photos Leirskardalen: Sølve Pettersen (Nordland County Council). Helicopter photos Fortun: The Norwegian Public Roads Administration and Yngve Midtun (The Norwegian Water Resources and Energy Directorate, NVE).

simulation match the initial elevation and related potential energy of the real events.

Regarding the simulations, the 3D point cloud DTMs were prepared in CloudCompare from different inputs, including airborne and terrestrial LiDAR surveys and SfM photogrammetry from UAV pictures (Table 1). The point spacing has been



Table 1: Simulation characteristics with stnParabel for each site. The point spacing has been adjusted for the simulations by sampling on a meshed version of the terrain models.

Sites	DTM-acquisition type and original format	Point spacing [m]	Nr. of points [$\cdot 10^6$]	Cell-size* (stnParabel raster) [m]	Nr. of rock scenarios; (unique size & shapes)	Total nr. of simulations
Chant Sura	Photogrammetry from drone, gridded	0.07	17.93	0.2	182; (8)	10,010
Riou Bourdoux	Photogrammetry from drone and terrestrial LiDAR, dense points	0.06	10.88	0.2	31; (31)	9920
Tschamut	Unspecified, gridded	0.03	12.47	0.1	102; (6)	9996
Authume A & B	Photogrammetry from drone and terrestrial pictures, gridded	0.06	8.22	0.2	A: 46; (46) B: 56; (46)	A: 10,350 B: 9,800
Mel de la Niva	Airborn LiDAR, gridded	2.30	1.81	2.0	1; (1)	10,000
Leirskardalen	Airborn LiDAR, point cloud	0.45	12.54	1.0	12; (12)	9996
Fortun	Airborn LiDAR, point cloud	0.35	14.90	1.0	19; (19)	9994
Holaviki	Airborn LiDAR, point cloud	0.33	10.84	0.9	2; (2)	10,000
Springville	Airborn LiDAR, point cloud	0.28	10.95	0.7	1; (1)	10,000
La Verda	Airborn LiDAR, point cloud	0.23	7.01	0.5	7; (7)	10,010
Glomset	Airborn LiDAR, point cloud	0.23	3.17	0.3	1; (1)	10,000

*The Cone $\alpha - \beta$ geometric runout extents were also projected on these raster DTM grids. They were resampled by average to 2 m for all Rockyfor3D simulations, and to 10 m for all Flow-R geometric predictions.

defined based on the smallest rock fragment scenario of each site to ensure that the simulations don't pass through the
 180 DTMs. Combining all scenarios, an arbitrary 10,000 simulations were targeted per site. To challenge the simulation model
 and evaluate its predictability, no other parameterization tuning was allowed. The simulations were thus all run from a
 simple set of objective inputs for an objective comparison and assessment of the predictability of the model as recommended
 in (Noël et al., 2023b). Therefore, the simulations were done without any added artificial roughness or subjective
 adjustments of terrain material properties, i.e., keeping the effective friction angle, cohesion, and dissipative rolling
 185 resistance coefficient unchanged as default settings.



To better communicate how the simulated trajectories would translate into corresponding hazard zones, the quantitative guiding hazard zoning concepts of Noël et al. (2023b) and Noël and Nordang (2025) were used. The local yearly hazard probability of being reached at least once (P_{Hazard}) depends on the on-site frequencies of rockfall fragments (λ) that are expected to propagate to the deposition zones and the size of the hypothetical exposed objects. Common for residences, a 30 m object width was used for the examples. Instead of using a fixed λ of 1/25 (per 10,000 m² source cliff area when used for diffuse sources), as it was found to be a good average starting point for Norway (Noël and Nordang, 2025), a range of λ was used for the “single source area” examples. Indeed, since frequency of failure and fragmentation usually depend on local onsite conditions, different quantified guiding zones were produced from the simulations for four arbitrary λ ranging from 1/100, 1/30, 1/10, to 1/1 expected averaged yearly propagating rock fragments per site. For a site with 10,000 simulated trajectories, a single trajectory carries a 1/10,000 portion of λ . When combining the carried frequency portion of each reaching trajectories at a given location, the trajectories with low intensities from their local low translational velocities and rock fragments mass were conservatively not removed for simplicity. A P_{Hazard} value of 1/300 was used as the threshold for drawing the hazard boundaries from “intolerable” to “tolerable” hazard. This value is similar to the 300 years return period hazard threshold of Switzerland (OFEV, 2016; OFDT et al., 2005; Raetzo et al., 2002). For Norway’s building code TEK17 (Norwegian Building Authority, 2017) and other similar countries tolerating less residual hazard than Switzerland for their land-use zonations, the same drawn boundaries correspond to those for P_{Hazard} of 1/1000 when the arbitrary expected λ values are lowered by about 3×.

2.1 Runout distribution validating approach

The validation approach for the predictive performance of the simulations to predict proper runout distribution in distance and lateral dispersion is based on a combination of qualitative and quantitative evaluations. Since quantitative rankings do not give the full picture, they should be complemented with qualitative performance evaluation. For the latter, the qualitative visual validation concepts from Noël et al. (2023b) were used. Maps showing the simulations in overlay to the observations were thus produced for each site. To help visually compare the simulated propagation paths, random samples of ten trajectories were picked for each site and displayed overlayed by the mapped observed paths. This way, it is possible to verify if the simulated paths unrealistically oscillate laterally in comparison to the observations, and if they share similar propagation paths. For the simulated deposited rock fragment positions, they are shown with high transparency. This way, areas with high density of deposited rocks are highlighted from the additive colour of the overlaying rocks, in a pseudo “heat map” way. These highlighted deposited regions can then be visually compared with the overlaying mapped deposited rock fragment positions. In complement, the 3D average position and centre of mass of the deposits, observed and simulated, are shown on the maps for a semi-quantitative visual validation.

Values for the average deposited rock fragment positions and centre of mass are also used quantitatively, expressed in terms of corresponding reach angles (angle in line-of-sight from the source starting position to the end deposited position). This allowed evaluating the errors on the predicted average runout distances despite the large variety of site geometries and



scales. Indeed, the relatively restrained observational error of about $\pm 0.2^\circ$ at worse remains relatively constant from site to site. For example, a ± 0.1 m imprecision inversely applied perpendicularly to both ends of the source-deposit line in a vertical plane rotates the reach angle by about $\pm 0.2^\circ$ for the ~ 60 m source-deposit average distance of the smallest site (Authume A site, T5 in Fig. 3.a). The same near maximal induced angular error of $\pm 0.2^\circ$ is obtained from a ± 3 m imprecision applied in the same way to the ~ 1475 m source-deposit average distance of the largest site (Mel de la Niva site, R1 in Fig. 3.a). And this observational error of $\pm 0.2^\circ$ is relatively pessimistic, since it is mostly induced from a vertical imprecision component in the previous examples, while the slight imprecision on the observations for the real sites lay mostly in the planimetric space given the use of airborne LiDAR data (ALS) for the vertical positioning. In the previous examples, if the reach angle lines are steepened by moving both ends (source-deposit) closer to each other horizontally by ± 0.1 m and ± 3 m respectively, the estimated observational angular errors are of $\pm 0.1^\circ$, better than the near worse estimates of $\pm 0.2^\circ$.

For the errors on the maximal predicted runout distances, observational biases must be matched by the simulations. Indeed, the rock fragments from a rockfall event or experiment followed some possible propagating paths leading to the mapped observed runout distances. However, this does not mean that all possible paths were explored, nor that all possible maximal runout distances were observed. With around 10,000 simulated paths, the likelihood of stochastically approaching all possible paths is higher, thus creating an imbalance where simulations should show longer maximal runout distances in comparison to the observations. To match biases, random samples of comparable size to a site's observations were iteratively used to evaluate their maximal runout distances. This was repeated for as long as there were sufficient remaining simulation results to be compared to the observations. The maximal runout distances of the samples were then averaged for each site in order to obtain the simulation maximal predicted runout distances, in reach angle, to compare to the observations.

For hazard, risk, and mitigation assessments, simulations with minimal error for both the centre of mass and maximal runout distances are desired for proper quantified rockfall statistics between these points. To complement these two metrics, the simulated deposits were also compared to the observations in terms of cumulative curves of reach angles. For the lateral dispersion, cumulative curves of deposits were used, this time with the lateral dispersion angle of each rock fragment from a reference trend line of the source average position to the centre of mass of the mapped deposit.

2.2 Velocities and bounce heights validating approach

Concerning the predicted translational velocities and bounce heights, it was possible to compare them to the reconstructed trajectories for the Chant Sura and Riou Bourdoux test sites (Caviezel et al., 2019; Noël et al., 2022; Hibert et al., 2024). The reconstructed trajectory segments cover most of the sites, so instead of being limited to a few 1D evaluation screens, it was possible to extend the comparison to cover most of the sites through numerous screens distributed concentrically around the sources. This innovative transition to a 2D rockfall statistical analysis from 3D results as in Crosta and Agliardi (2004) has never been compared before with 3D reconstructed observations to the knowledge of the authors. This was made possible thanks to the freely accessible dataset from the WSL Institute for Snow and Avalanche Research SLF in collaboration with Geobruigg (Caviezel et al., 2020) and Hibert et al. (2024).



The sites were subdivided into 200 evaluation screens/bins of widths of 1.4 m and 1.1 m, respectively. Box plot statistics were extracted as well as the minimums, averages, maximums, and standard deviations for every evaluation screen. For each screen, comparable sample sizes were iteratively used based on the local number of observations to reduce comparison biases. This was repeated for as long as there were sufficient remaining simulation results to be compared to the observations. The resulting 2nd, 25th, 50th, 75th, and 98th percentiles from the numerous simulation samples were respectively averaged for each screen. Also, to locally ensure a proper statistical representation of the data, only the local maximal translational velocity and bounce height values of each unique trajectory segment were used so that each segment was counted as one “item” per screen. Otherwise, considering all vertices points of a trajectory segment would have introduced a bias over-representing slower reconstructed trajectories. Indeed, if the points forming a trajectory segment are equally spaced in time (e.g., from a constant time step), a slower trajectory segment has a higher density of points than a faster trajectory segment of equal length. Without accounting for this bias, a 5 m s⁻¹ trajectory segment would have had the same weight as four 20 m s⁻¹ segments, for example.

3 Simulation results and analyses

As for the results of the Mel de la Niva site presented in Noël and Nordang (2025), the simulation results for the eleven additional sites are presented in the next sub-sections in a way that facilitates the qualitative comparison following the validating reasoning steps given in (Noël et al., 2023b) summarized as follows:

- First, the simulations were performed with objective parameters kept within the anticipated realistic range: e.g. default settings of stnParabel keeping the Rolling friction model as described in Noël et al. (2023a), rapid automatic simulations of Rockyfor3D, default angles and coefficients of the cone $\alpha - \beta$ method from Derron et al. (2016), and parameters of Oppikofer et al. (2024) for Flow-R. This way, one can objectively verify if the simulations fail/succeed at reproducing the observed rockfall deposits with the following question: are the runout distances, bounce heights, and velocities exaggerated or underestimated? Since parameters were not iteratively adjusted, the sensitivity of the models is not explored here, however.
- For proper hazard probabilities, the temporal rockfall frequencies must be realistically distributed spatially. From a randomly selected small sample set of simulated trajectories, one can visually verify if the lateral dispersion and overall distribution seem realistic with the following questions: (1) are the simulated paths unrealistically oscillating laterally and undertaking unnatural paths, (2) are they reproducing the distribution of the observed paths for proper distributed expected frequencies, (3) are the bounce heights and velocities similar to those of the observations when compared with a similar sample size?
- From a larger set of simulated trajectories, similar to what may be used in a delivered project, ensuring that the runout extent and distribution are close to and proportional to the observations. Then, the following questions can be explored: (1)



Are the envelopes of the simulations including most of the observed rockfall deposits? (2) Are the simulated paths and deposited rock fragments overly channelized, leading to skewed spatial rockfall statistics? (3) Are the simulated rockfall path widths representative of the real potential reach of the propagating rock fragments?

For conciseness, only the results for the Rolling friction with stnParabel are given and analyzed here. Those of the other models are given as supplementary material, and their qualitative analysis is left to the discretion of the reader.

3.1 Rockfall test sites

Starting with the rockfall test sites, similar, relatively well-matching results as for the Mel de la Niva site (Noël and Nordang, 2025) are obtained (Fig. 5 and Fig. 6), but on the other spectrum of volumes and sizes involved. The deposited rocks from the simulations are close to those observed. Indeed, the simulated average reach angles by mass shown in Fig. 5 have a mean error (ME) of 0.6° as accuracy, with a mean absolute error (MAE) of 1.3° as precision (ME of 0.5° , MAE of 1.2° for the averages by rocks). This relatively high predictability is achieved despite the wide range of average reach angles varying by about 16° (27.5° to 42.1° for the 3D mass centres of the observed values). Moreover, not only the runout extents are obtained with the process-based trajectory simulations, but also useful complementary information such as the runout paths, lateral dispersion, velocities, energies, bounce heights, etc.

The cumulative distributions of deposited rocks match particularly well with the observations for the two first sites (Fig. 6a and b). This is of course to be expected for the Riou Bourdoux site since its rockfall impacts from reconstructed trajectories have been used for the development of the Rolling friction rebound model in combination with the reconstructed impacts from the Mel de la Niva site. To quantitatively complement the visual comparisons of cumulative distributions, two-sample Kolmogorov-Smirnov (KS) tests were performed with the distributions by rock fragment numbers. To reduce biases, samples of corresponding sizes were compared (Antoch et al., 2010). This was repeated for as long as there were sufficient remaining simulation results to be compared to the observations. If the KS test does not reject the null hypothesis (H_0), it can be assumed that the simulated dataset reproduces the same continuous distributions as those observed. The situation is the opposite if the result is rejected (H_1). The ratios of tests that do not reject the null hypothesis are given in Fig. 6.

Focusing on the discrepancies, the simulated runout distances for Tschamut are more clustered (Fig. 5c and Fig. 6c), and a proportion of the rocks stops earlier compared to the observations. The model may have a slightly too high rolling resistance for the combination of small and non-elongated rocks on freshly mown grass, especially for the artificial “ball-shaped” rock (EOTA₁₁₁). The average reach angle by mass of the simulations still reaches a relatively low value of 28.7° for that site, significantly lower than for the other test sites.

For Authume, the simulated rocks travel slightly longer distances on average than the observations (Fig. 5d and Fig. 6d). Despite this, the average reach angle by mass of the simulated rocks is steeper than the observations for profile A. This is because numerous simulated rocks come to rest on the lower level of the quarry right after passing over a ramp. This increases their height difference with their source to a larger extent than their horizontal runout distances, which steepens their reach angles as a result. Note how the cumulative distributions of reach angles significantly vary from site to site and

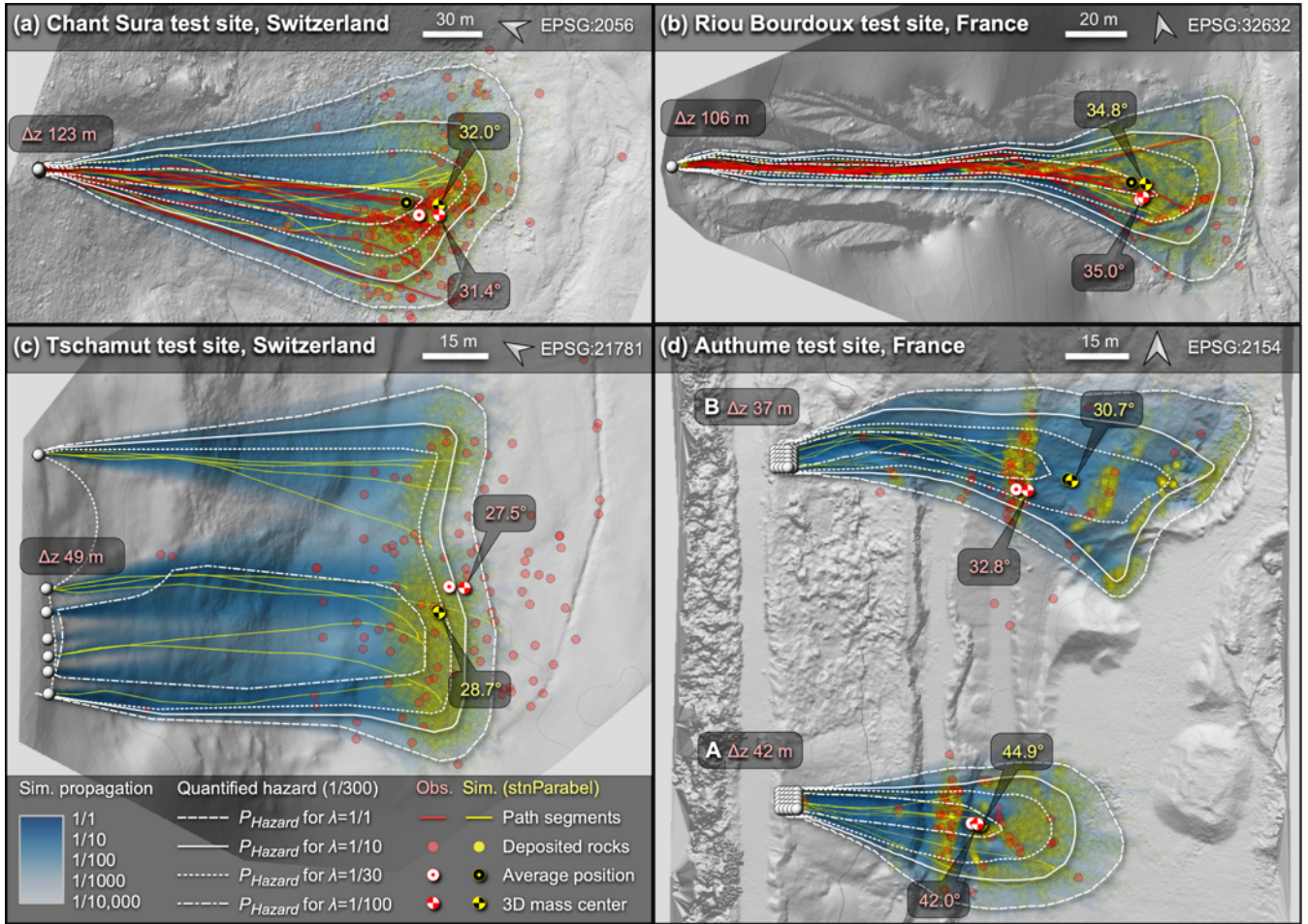


Figure 5: Comparison for the rockfall test sites of the mapped deposited rocks vs. the simulations from stnParabel. The observations are shown in red while the simulations are shown in yellow and blue. The deposited rocks are shown with semi-transparent colours so that areas with a high density of overlapping rocks are highlighted with stronger corresponding colours. The average deposited locations per rock fragment and per mass are overlaid with characteristic symbols. A small sample set of trajectories randomly selected for each site is shown with subtle yellow lines to help evaluate the realism of the simulated paths undertaken by the rocks. The simulated propagation ratio of the spatially distributed frequencies evaluated from rockfall paths of width corresponding to their respective rock's d_1 diameter are shown with semi-transparent blue colours to contrast with the overlaying red and yellow colours. White contours show the 1/300 quantitative guiding hazard zonation based on different expected frequency of rock fragments. Low intensities were not filtered out for drawing the guiding contours.

from profile A to B for the Authume site (Fig. 6). Impressively, the simulation model manages to objectively reproduce those differences to a relatively high degree prior to any per-site fine-tuning of parameters.

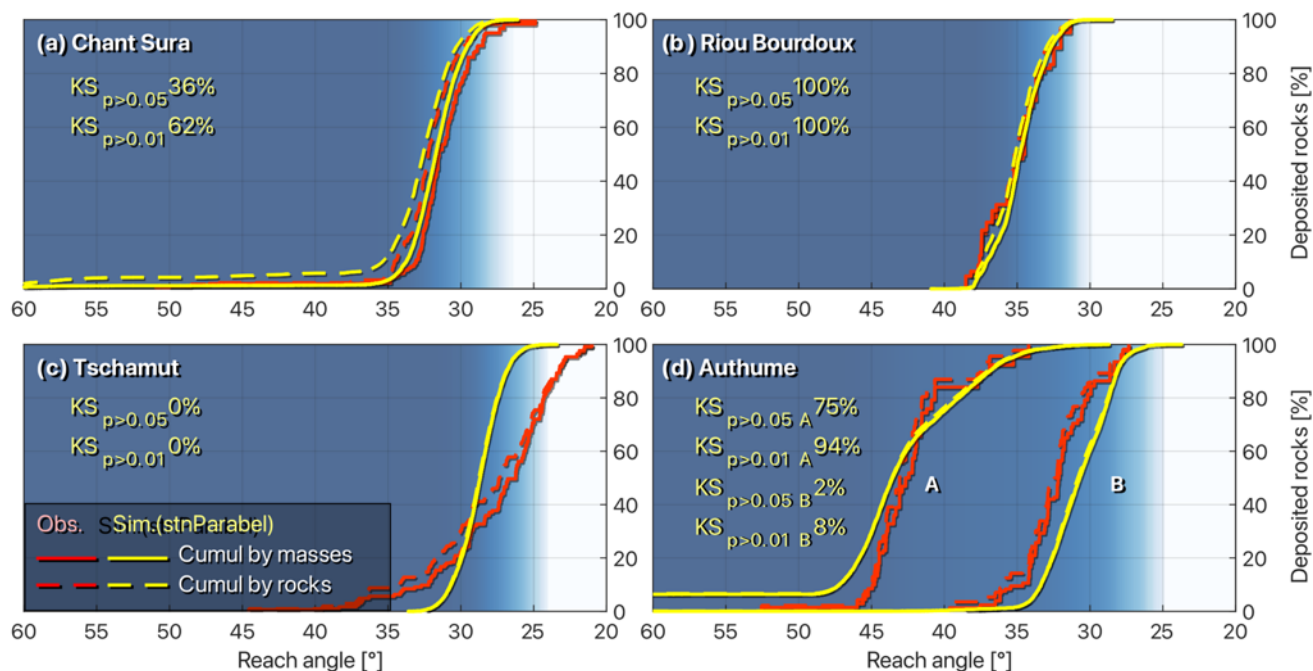


Figure 6: Comparisons of the cumulative distribution of the observed and simulated deposited rock fragments in terms of reach angles (energy line angles from the source to the deposited location). The observations are shown in red and the simulations from stnParabel are shown in yellow and blue. Also, for homogeneity across the paper and continuity with (Noël et al., 2022; Noël et al., 2023a; Noël et al., 2023b; Noël and Nordang, 2025), the runout propagation goes from left to right. The reach angle axes are thus inverted since the reach angles usually decrease as the runout distances increase.

3.1.1 Test sites lateral dispersions

Regarding the lateral dispersion obtained for the test sites, the simulations manage to reproduce relatively well the observed spreading (Fig. 5 and Fig. 7). Apart from the Riou Bourdoux site where both the observations and simulations are naturally channelized by the ravine morphology of the site, none of the results for the other sites appear overly channelized. The variable general dispersion distributions seem also respected for all sites, with less spreading at the Riou Bourdoux site (near 90% of the deposited rocks within $\pm 5^\circ$) compared to the Chant Sura and Authume A sites (around 90% of the deposited rocks within $\pm 15^\circ$). For Authume B, the dispersion seems slightly skewed by about 5° toward the north compared to the observations (Fig. 5d and Fig. 7d).

3.1.2 Velocities and bounce heights

The obtained values are compared to the ones of the reconstructed trajectories in Fig. 8 and summarized in Table 2. The values given in Table 2 for the Chant Sura test site are subdivided into two regions shown in Fig. 8a corresponding to those of the inspiring work from Lu et al. (2019) related to RAMMS::ROCKFALL. For those two regions, our averaged median

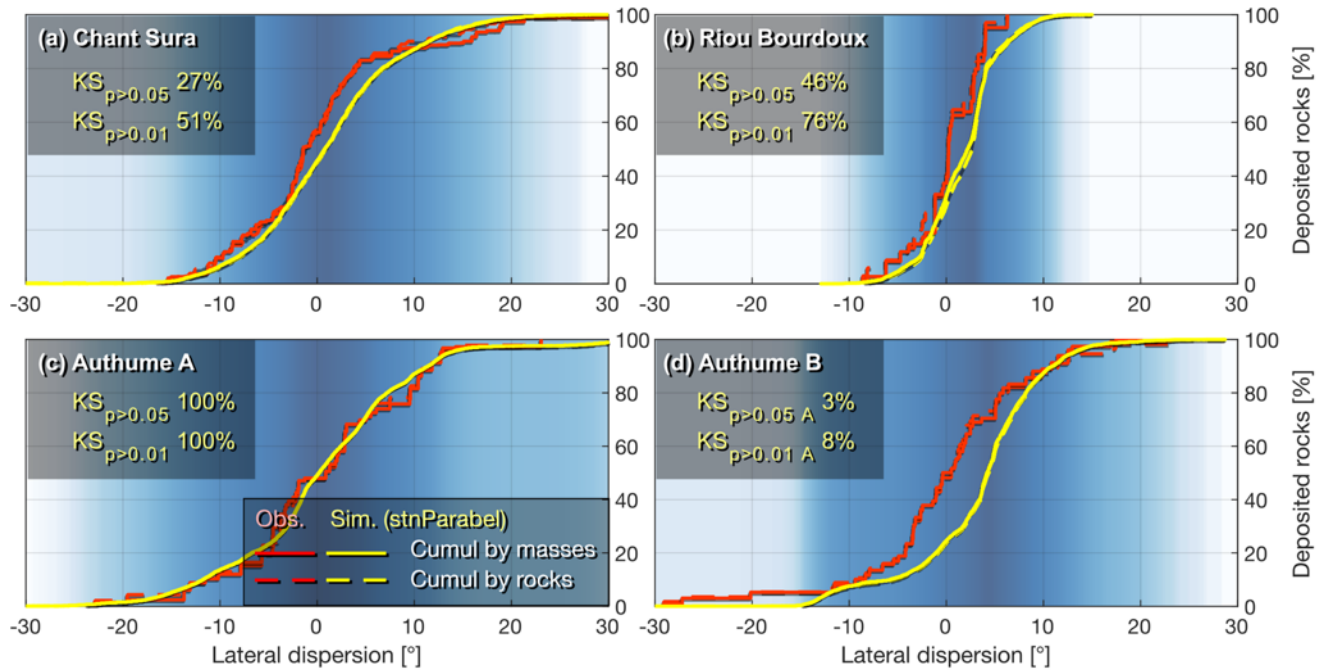


Figure 7: Comparison of the lateral dispersion of the deposited rocks for the test sites, excluding Tschamut since several release points were used. The observations are shown in red and the simulations from stnParabel are shown in yellow and blue. The cumulative curves are built by scanning in a counterclockwise direction around each average source point and by setting the zero-degree dispersion at the observed mass centre.

330 (P_{50}) velocities and bounce heights from the numerous screens match very well with the average simulated values from Lu et al. (2019). Compared to the observations that vary significantly through the screens inside each region, the standard deviations of our simulated values properly reflect those variations from screen to screen. As for the simulated velocities of Lu et al. (2019), our averaged values are also slightly underestimated for region 1. This results in a relative error (RE) of -20% for the average median of our simulations, i.e., a difference of 2.7 m s^{-1} . A part of this relative error can be attributed to the fact that the slower parts where the observed rocks are not free-falling were not reconstructed, mostly near the source and in the deposited area. This introduces a slight observational bias that can raise the values of the observations in comparison to the simulations. This should also be true for the observed bounce heights. The RE of 17% on our slightly overestimated median bounce height for region 1 is similar, in terms of absolute deviation, to the RE of -17% on the slightly underestimated mean bounce height of Lu et al. (2019). To put things into perspective, those minor differences of 20–30 cm are less than the diameter of the smallest rock tested for this site and are thus within the precision limits of the reconstructed data. The same narrow range of error was obtained for the other percentiles and for the maximum bounce heights. The performance of stnParabel with the Rolling friction model at accurately predicting the bounce heights in that region is thus remarkable, especially given the limited fine-tuning of the parameters kept at their default values.

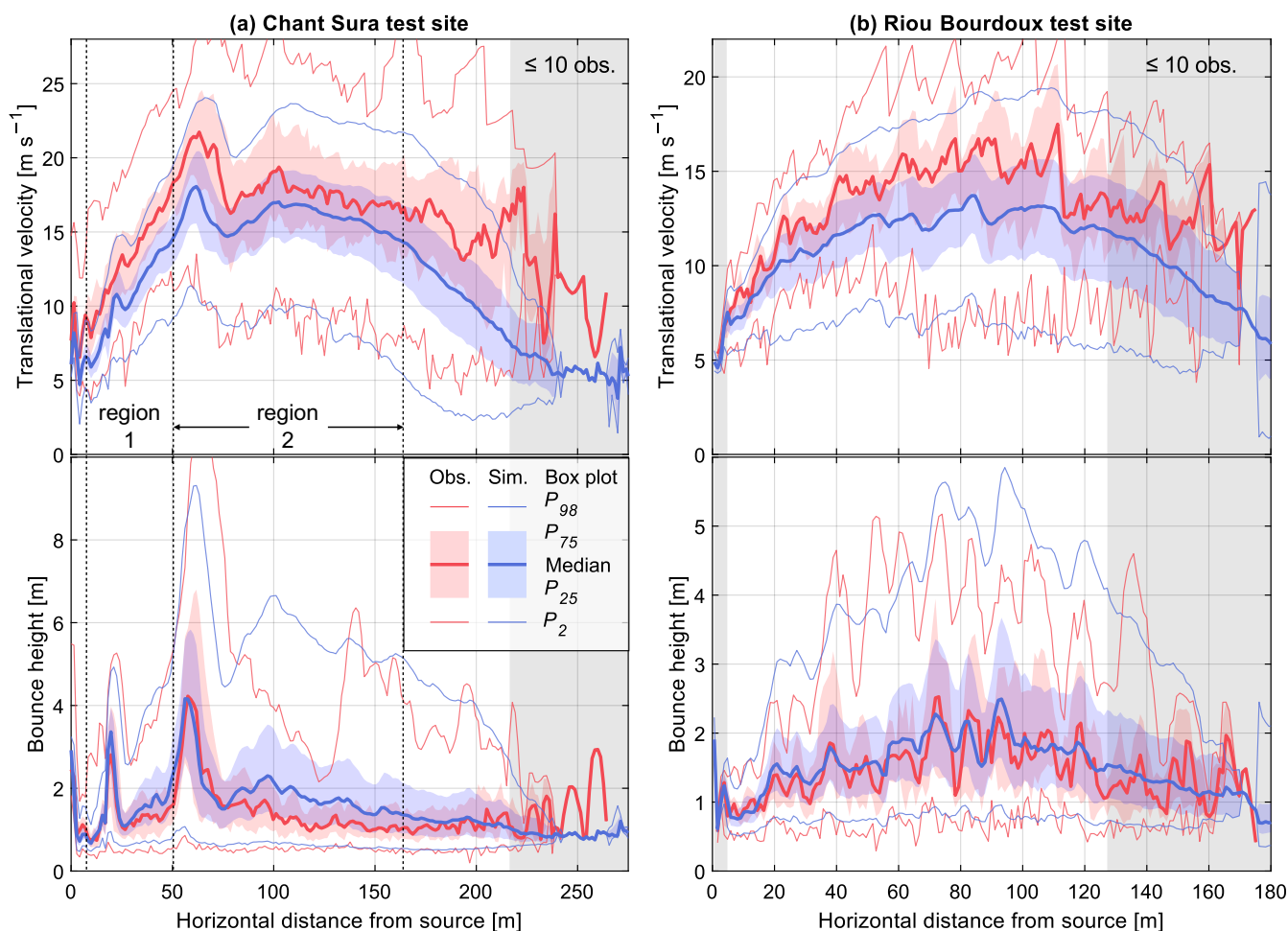


Figure 8: Comparison of the translational velocities and bounce heights of the "observed" reconstructed trajectories and simulated ones from stnParabel with box plot statistics.

Concerning region 2, which covers most of the Chant Sura test site (Fig. 8a), the simulated translational velocities are again



Table 2: Summarized box plot statistics related to the translational velocities and bounce heights for the regions covered.

Chant Sura test site, region 1 (10–50 m)												
	Velocities [m s ⁻¹]						Bounce heights [m]					
	<i>P</i> ₂	<i>P</i> ₂₅	<i>P</i> ₅₀	<i>P</i> ₇₅	<i>P</i> ₉₈	Max.	<i>P</i> ₂	<i>P</i> ₂₅	<i>P</i> ₅₀	<i>P</i> ₇₅	<i>P</i> ₉₈	Max.
Observed	8.5	11.9	13.7	15.5	20.3	21.3	0.5	0.9	1.4	2.0	3.8	4.4
Simulated	7.0	9.6	11.0	12.4	15.0	15.6	0.7	1.2	1.6	2.2	3.6	4.1
S.D. obs.	2.5	2.7	2.6	2.4	2.8	2.7	0.1	0.2	0.5	0.7	0.9	1.0
S.D. sim.	1.6	2.2	2.4	2.6	3.0	3.1	0.1	0.4	0.6	0.7	1.1	1.2
RE [%]	−15	−19	−20	−21	−27	−28	45	27	17	10	−6	−7
S.D. RE [%]	13	4	5	6	8	9	18	16	13	17	19	23
Chant Sura test site, region 2 (50–165 m)												
	Velocities [m s ⁻¹]						Bounce heights [m]					
	<i>P</i> ₂	<i>P</i> ₂₅	<i>P</i> ₅₀	<i>P</i> ₇₅	<i>P</i> ₉₈	Max.	<i>P</i> ₂	<i>P</i> ₂₅	<i>P</i> ₅₀	<i>P</i> ₇₅	<i>P</i> ₉₈	Max.
Observed	9.3	15.0	18.0	21.0	26.4	27.7	0.5	1.0	1.6	2.4	5.3	6.3
Simulated	8.8	13.6	16.0	18.2	22.3	23.2	0.7	1.2	1.9	3.0	5.9	6.8
S.D. obs.	1.4	1.5	1.4	1.5	1.7	1.5	0.1	0.4	0.8	1.4	2.3	2.6
S.D. sim.	1.4	0.9	0.9	1.0	1.0	1.0	0.1	0.4	0.6	0.9	1.1	1.1
RE [%]	−4	−9	−11	−13	−15	−16	20	24	32	40	30	28
S.D. RE [%]	16	6	5	5	5	4	15	16	25	33	50	60
Riou Bourdoux test site (all ≥ 10 m)												
	Velocities [m s ⁻¹]						Bounce heights [m]					
	<i>P</i> ₂	<i>P</i> ₂₅	<i>P</i> ₅₀	<i>P</i> ₇₅	<i>P</i> ₉₈	Max.	<i>P</i> ₂	<i>P</i> ₂₅	<i>P</i> ₅₀	<i>P</i> ₇₅	<i>P</i> ₉₈	Max.
Observed	7.9	11.7	14.0	15.8	18.4	18.4	0.7	1.1	1.6	2.3	3.5	3.5
Simulated	6.8	9.9	11.9	13.8	16.9	16.9	0.8	1.2	1.7	2.4	4.1	4.1
S.D. obs.	1.5	1.7	2.0	2.5	2.9	2.9	0.2	0.2	0.4	0.7	1.0	1.0
S.D. sim.	0.7	1.0	1.3	1.7	2.5	2.5	0.1	0.2	0.3	0.5	1.0	1.0
RE [%]	−11	−15	−14	−12	−8	−8	19	10	9	11	20	20
S.D. RE [%]	19	7	5	6	7	7	30	15	16	20	24	24

345 slightly underestimated. This is also the case for some of the averaged velocities of Lu et al. (2019) but depends on their



parameters used (see Lu et al. 2019). Our RE on the velocities is better than for region 1, but slightly larger than some of those from the mean velocities of Lu et al. (2019). Still, with a value of -11% for our median, and remaining below $\pm 17\%$ for all the other percentiles and maximum velocity values, the RE are exceptionally good for simulations constrained to simple objective inputs. As for Lu et al. (2019), the simulated bounce heights for region 2 are slightly overestimated.

350 Looking at Fig. 8, it is possible to note that the simulated bounce heights diverge mostly after the cliff present from 50–70 m. Because of those differences, the RE are larger but the ME remain reasonable with differences within 30 cm for the average of the lower percentiles including the median, and within 60 cm for the higher percentiles and the maximum value at worst. To put those ME values in perspective, they are similar to the radii of the larger rocks used for that site.

For the Riou Bourdoux test site (Fig. 8b), the simulated velocities from stnParabel are slightly underestimated, with again the

355 potential observational bias partly to blame. Note that the maximal value and the 98th percentile are the same due to the smaller size of the samples (Table 2). The RE is of a reasonable -14% for the average median across the site (excluding the first 10 m and the values with ten observations or less). It remains close to that value from screen to screen, as shown by the low S.D. on the RE. The RE does not exceed -15% for the other averaged percentiles or the maximal velocity value. The ME for those averaged values is within -2.1 m s^{-1} at worst. Overall, the predictive performance of the objective simulation

360 model related to the velocities is very good given the relatively small RE obtained. Concerning the bounce heights, all the RE are within 20% for the averaged percentiles and the average maximum value. The average median has the lowest RE at 9%. With positive relative errors, the predicted bounce heights are slightly overestimated. Their ME is of 60 cm at worst on the averaged statistical values from the evaluation screens, which is a relatively good performance for the objective predictability of the simulations.

365 To put those promising results from stnParabel into perspective, let us compare them to the interesting benchmark from Berger and Dorren (2006). Twelve participants were tasked with predicting the runout distances, velocities, and bounce heights at two evaluation screens for a site where over 100 rocks were experimentally rolled down. With the freedom of subjectively fine-tuning their simulations iteratively by expert judgement for their preferred simulation software, 25.0% of the participants managed to predict either the maximal velocities or the bounce heights for at least one of the two evaluation

370 screens within a RE of $\pm 20\%$. Only one participant (8.3%) managed to predict both the maximal velocities and bounce heights for the two evaluation screens within the RE margins.

For comparison, we iteratively (10,000 times) randomly selected a set of screen pairs from our hundreds of evaluation screens for the previously covered regions at the Chant Sura and Riou Bourdoux test sites. The same conditions were verified for each pair: 1) is there at least one of the maximal translational velocity or maximal bounce height predicted within

375 a RE of $\pm 20\%$, and 2) are all quantities predicted within the RE margins for the evaluation screen pair?

For the region 1 of the Chant Sura test site, the objective stnParabel predictions succeed for 96.6% of the evaluation screen pairs at fulfilling the first condition. This is a significant improvement compared to the 25.0% prediction performance of the combined 12 participants from Berger and Dorren (2006). Different sites are, however, involved here, which limits how far



one could draw conclusions from the comparison. Regarding the second condition, stnParabel's predictions succeed for
380 10.2% of the evaluation screen pairs of the region 1, which is similar to the 8.3% from Berger and Dorren (2006). For the
region 2 evaluation screens, the predictions from stnParabel fulfil the conditions 1 and 2 for 98.7% and 27.1% of the
randomly selected evaluation screen pairs, respectively. Again, this contrasts with the prediction performance (25.0% and
8.3%) of the simulations involving subjective fine-tuning from expert judgment in Berger and Dorren (2006).

Finally, for the Riou Bourdoux test site, the predictions fulfil the conditions for 99.1% and 21.3% of the screen pairs,
385 respectively. Therefore, stnParabel's predicted maximal translational velocities and bounce heights are within the margins
defined by Berger and Dorren (2006) for a large portion of the hundreds of evaluation screens. If the sites are comparable to
the one used in Berger and Dorren (2006), then this constitutes a significant improvement in terms of predictability.
Moreover, this performance is combined with low errors on the runout predictions. Such combined predictions of rockfall
kinematics and runout distance is rarely achieved together within one simulation (Labieuse, 2004; Berger and Dorren, 2006;
390 Lambert and Bourrier, 2013; Dorren et al., 2013). And furthermore, as envisaged by Pfeiffer and Higgins (1990), this
achievement is obtained from simple objective inputs that greatly simplify the practical use of the simulation model. This
thus has the potential of considerably reducing the time resources spent for subjective iterative fine-tuning from expert
judgment.

Of course, a model remains a model; there are always simplifications compared to the complex true nature of the
395 phenomena. Therefore, as with any simulation model, one should not trust the simulation results blindly without on-field
validations. Despite the progress, improvements could be made regarding the slightly underestimated translational velocities
and the marginally overestimated median bounce heights. Given that the velocities after impact are relatively well predicted
on a "per-impact" basis from imposed incident angles (Noël et al., 2023a), it is possible that the impact detection algorithm
used (Noël et al., 2021) introduces some roughness from the point cloud surface. If so, this may explain the marginal
400 deviations of the predictions, combined with the slight observational bias. On the model side, the impact detection algorithm
could potentially be improved to be more robust relatively to the roughness introduced by the point spacing. This may raise
the simulated velocities while lowering the bounce heights, thus potentially reducing the RE on those points if properly
balanced. Additionally, the Rolling friction model could be fitted to the now larger datasets available. And this could be
repeated on segmented data per terrain material involved if it is judged to bring significant improvements without
405 introducing excessive subjectivity. On the observational side, the data from the previous rockfall experiments should be
completed to also include the trajectory segments that are not in the free-falling phase. Also, the materials involved should
be noted for every rock-ground interaction to better document the remaining subtle effect of the materials.

3.2 Real rockfall event sites

Continuing the objectivity and predictability investigation with the sites previously affected by real rockfall events, as for the
410 rockfall test sites, similar well-matching runout results are obtained by stnParabel (Fig. 9 and Fig. 10). The deposited rocks

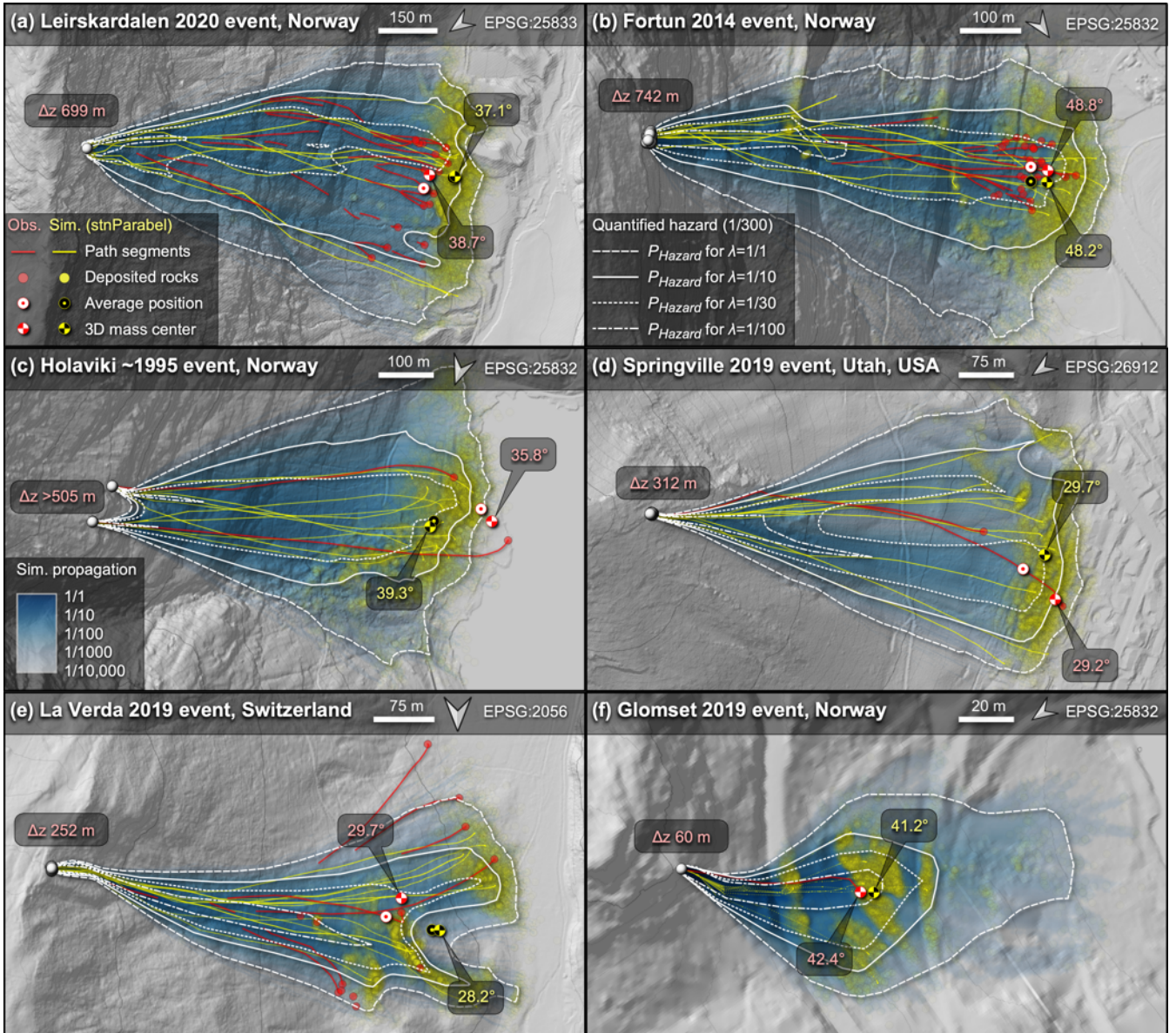


Figure 9: Comparison for the real rockfall event sites of the mapped deposited rocks vs. the simulations from stnParabel. The observations are shown in red while the simulations are shown in yellow and blue. The deposited rocks are shown with semi-transparent colours so that areas with a high density of overlapping rocks are highlighted with stronger corresponding colours. The average deposited locations per rock fragment and per mass are overlaid with characteristic symbols. A small sample set of trajectories randomly selected for each site is shown with subtle yellow lines to help evaluate the realism of the simulated paths undertaken by the rocks. The simulated propagation ratio of the spatially distributed frequencies evaluated from rockfall paths of width corresponding to their respective rock's d_1 diameter are shown with semi-transparent blue colours to contrast with the overlaying red and yellow colours. White contours show the 1/300 quantitative guiding hazard zonation based on different expected frequency of rock fragments. Low intensities were not filtered out for drawing the guiding contours.

from the process-based simulations are again close to those observed. This time, the accurately simulated average reach

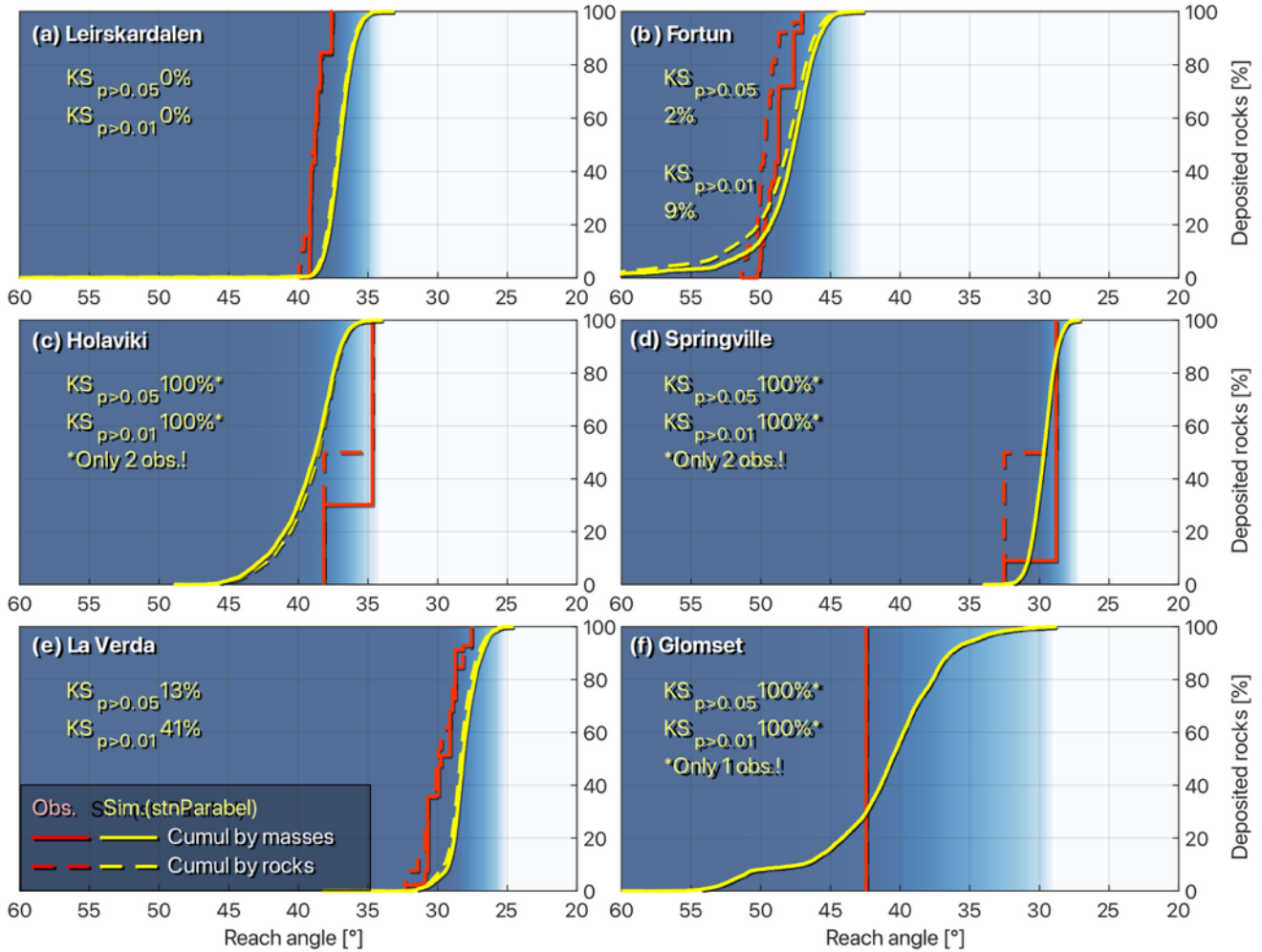


Figure 10: Comparisons of the cumulative distribution of the observed and simulated deposited rock fragments in terms of reach angles (energy line angles from the source to the deposited location). The observations are shown in red, and the simulations from stnParabel are shown in yellow and blue. Also, for homogeneity across the paper, the runout propagation goes from left to right. The reach angle axes are thus inverted since the reach angles usually decrease as the runout distances

angles by mass shown in Fig. 9 have a mean error (ME) of -0.1° from those of the observations, with a slightly lower precision given by the MAE of 1.5° however (ME of -0.6° , MAE of 1.5° for the averages per rock fragment). The larger variability may partly be attributed to the less representative smaller samples of the observed rock fragments. Combined with the previous values from the test sites, the mean error of the stnParabel process-based simulations becomes 0.2° with a MAE of 1.4° (ME of -0.1° with a MAE of 1.4° for the averages per rock fragment). And this is now achieved despite the wider range of average reach angles varying by about 21° , from as low as 27.5° at the Tschamut test site to as high as 48.8° at the Fortun site for the 3D mass centres of the observed values.

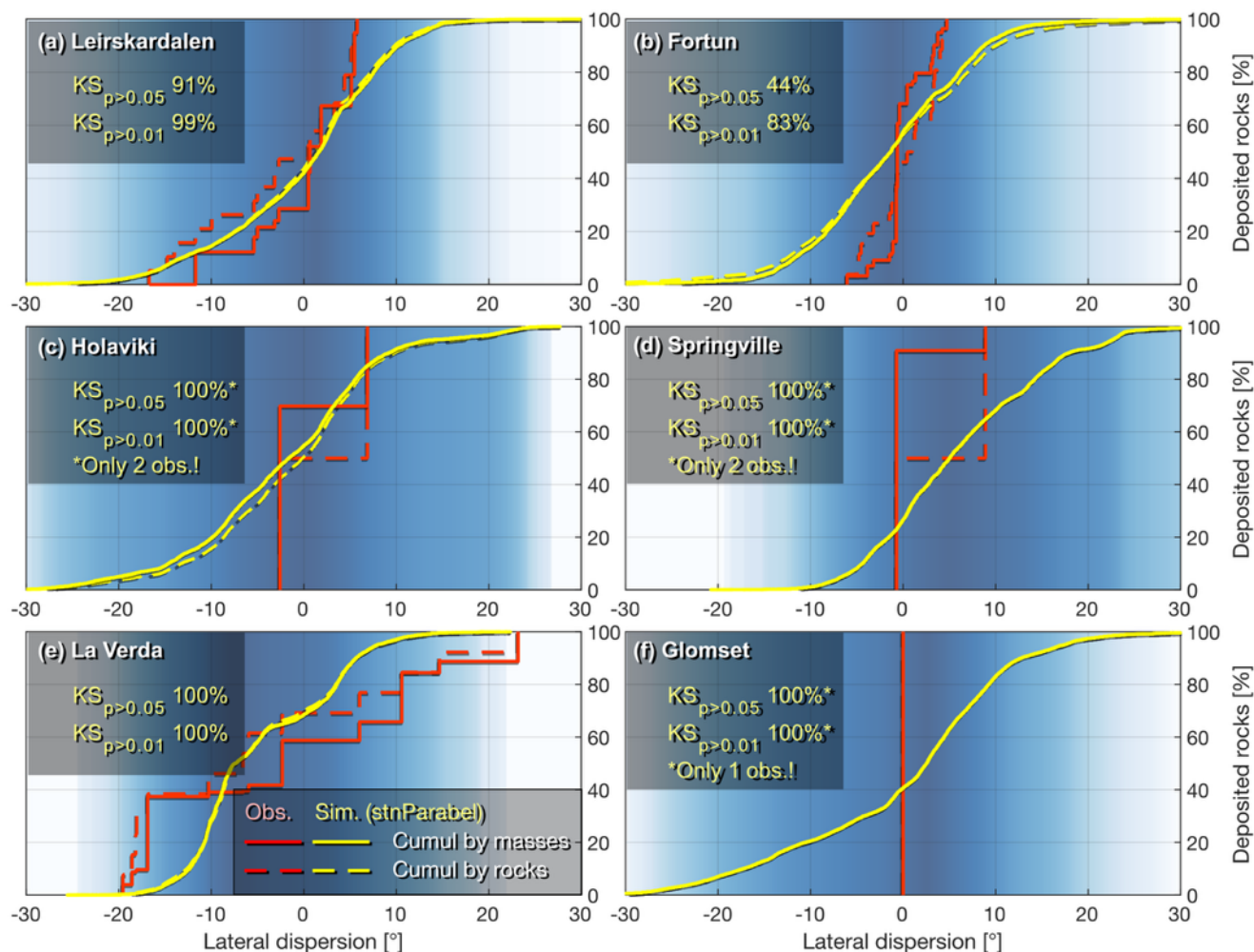


Figure 11: Comparison of the lateral dispersion of the deposited rock fragments of the real rockfall event sites. The observations are shown in red, and the simulations from stnParabel are shown in yellow and blue. The cumulative curves are built by scanning in a counterclockwise direction around each average source point and by setting the zero-degree dispersion at the observed mass centre.

As for the previous Mel de la Niva site, the randomly selected small sample set of simulated trajectories can be compared visually with the partially mapped path segments (yellow vs. red paths in Fig. 9). Note how most simulated trajectories are parallel to and follow the observed rockfall paths. Again, the overall distribution of the simulated trajectories follows the observed mapped paths. The simulations are not over-channelised and do not unnaturally oscillate from left to right. Also, when sufficient observed paths are present, the samples of simulated trajectories cover similar lateral spreading, which suggests that the probabilistic aspect of the simulations is close to reality. The qualitative verification steps given in Noël et al. (2023b) are thus fulfilled. Focusing on the discrepancies, none of the simulated trajectories from the small sample set of the Springville site reproduces the kink observed at 150 m from the source on the mapped path. This may be partly attributed



to the aggressive filtering of the vegetation applied by the provider of the ALS DTM. Indeed, this removed a considerable amount of surface roughness that would have induced higher lateral deviations for that site. As described by Noël et al. (2021), this can be compensated for with further fine-tuning simulation attempts by adding artificial surface roughness to the simulations. However, this fine-tuning of the simulation model parameters was omitted here to avoid introducing subjectivity to the comparisons.

Considering the limited samples of observed deposited rock fragments, the cumulative distributions of simulated deposited rocks match well with the highly variable observations, especially if cumulated by mass (Fig. 10). For the three real event sites with samples of more than ten observations, the simulation model tends to be slightly too conservative. For the others, the trends should not be considered too seriously due to the limited samples. Nevertheless, the simulated rocks seem to come to rest too early for one of the two observed rock fragments of the Holaviki site (Fig. 9c and Fig. 10c). Has this rock with the longest runout significantly changed size and mass during its course so that the simulations from its final dimensions could not reach as far? Has the rock followed a highly improbable trajectory so that none of the 5000 simulated attempts could reproduce it perfectly (the closest simulated rock fragments came to rest at 30 m of the observed rock after propagating beyond 730 m horizontally from the source)? Was the source located further up, which could have allowed the rock to gain more velocity than in the simulations? It is difficult to know for sure, but this rock fragment certainly challenges the simulation model. It should, however, be noted that only the position of the two fragments with longer runouts that stopped near buildings were reported in Domaas (1995). Nonetheless, it was mentioned that numerous other fragments stopped earlier (Domaas, 1995). The steeper reach angles of the other fragments would pull the average reach angle back toward the one of the simulations. The cumulative curve of the observed rocks would also be pushed up, closer to the one of the simulations.

3.2.1 Real events lateral dispersions

Regarding the lateral dispersion obtained for the real rockfall event sites, the simulations manage again to reproduce relatively well the observed spreading (Figures 9 and 11). However, the results here are less representative given the smaller number of observations compared to the test sites. The dispersion for the Leirskardalen site (Fig. 9a and Fig. 11a) matches relatively well the observations, with a greater dispersion toward North-West ($\sim 15^\circ$ for 45% of the rock fragments) than toward South-East ($\sim 10^\circ$ for another 45% of the simulated rock fragments). The South-East dispersion part seems even more constrained on the observations, so the simulations exaggerated slightly the lateral spreading. This seems more pronounced at the Fortun site (Fig. 9b and Fig. 11b), with most observations surprisingly laying within a narrow $\pm 5^\circ$ in comparison to near $\pm 15^\circ$ for 90% of the simulated rock fragments. Could it be because the final shapes of the rock fragments were simulated from the source, while in reality these shapes may have been created through late fragmentation during the propagation? It is hard to comment on the dispersion from the Holaviki, Springville and Glomset sites given their limited number of observations. For the La Verda site (Fig. 9e and Fig. 11e), despite perfect KS scores, the simulations seem to



slightly underestimate the lateral dispersions in comparison to the observations ($\pm 15\%$ and $\pm 20\%$ for around 90% of the
460 deposited rock fragments respectively).

4 Runout analysis in light of existing methods

So far, the results were mostly shown with maps and cumulative distributions (Fig. 8, Fig. 9, Fig. 10, and Fig. 11). They
were mostly analyzed in terms of runouts expressed as average reach angles at the centre of mass of the deposits. For
perspective, these results, including those for the Mel de la Niva site, are here summarized in Fig. 12a next to those of the
465 other process-based approach, respectively stnParabel raster in Fig. 12d and Rockyfor3D in Fig. 12g. And for a simpler
overview, they are also shown as pie charts per simulation approach and in histogram form per precision classes in Fig. 13.
Compared to the other process-based approach, i.e. stnParabel raster and Rockyfor3D, stnParabel got the best accuracy at
predicting the centre of mass of the observed runouts with a null mean error (ME = 0.0°). The others are not far behind with
mean errors of -0.2° and -0.4° for stnParabel raster and Rockyfor3D, respectively. With negative ME, stnParabel raster and
470 Rockyfor3D overshoot slightly the observed centre of mass on average, which is safe. Compared to Rockyfor3D, the
precision of the predictions from stnParabel and stnParabel raster is doubled or tripled ($\sim 1/2$ MAE and $\sim 1/3$ RMSE of
Rockyfor3D's predictions). However, when looking at the predicted points along their 1:1 perfect fit line or on their
precision-accuracy targets, it can be noted that most predictions from Rockyfor3D exhibit similar decent precision as
stnParabel and stnParabel raster. It actually has slightly more sites with predictions in the best precision class (AE from $0-1^\circ$)
475 than the other methods, as shown on the pie charts and histograms (Fig. 13). The low obtained precision score is due to two
strong outliers on Rockyfor3D's predictions, where it got challenged for the Glomset and Tschamut sites. Otherwise, its
precision is comparable to the one of stnParabel and stnParabel raster predictions.

For quantified guiding hazard zones, not only predicting the average runout distances matters. Indeed, the whole range of the
cumulative curves is of interest for quantified guiding hazard zones. As shown in Fig. 5 and Fig. 9, different guiding zones
480 quantified according to different hypothetical expected frequencies are drawn at different local cumulated reach values. If the
simulations predict well the average expected runouts and reproduce their distributions, then the longest expected runouts
should also be properly predicted. So, to complement the previous comparisons of average reach angle values, they are also
put aside with comparisons for the longest runouts expressed as reach angles in Fig. 12 and Fig. 13. This time, the scores for
the predictions of the geometric methods are also compiled. Moreover, this complementary comparison can favour the
485 geometric methods since they are developed to rather predict conservative propagation distances for susceptibility
assessments, or as early coarse estimations in the other types of assessments (Loup and Dorren, 2022; Noël et al., 2023b).

Without excluding the sites with the most limited number of observations, the accuracy of the process-based simulations
with the Rolling friction model (stnParabel) appears to be similar to the perceived accuracy of the other approaches for
predicting the longest runout distances by reach angles (ME of 0.3° for stnParabel vs. 0.7° , 0.6° , 0.3° , and -0.5° in Fig. 12).
490 Like the perceived accuracy of the Cone $\alpha - \beta$ approach, stnParabel appears to have the best accuracy for predicting the

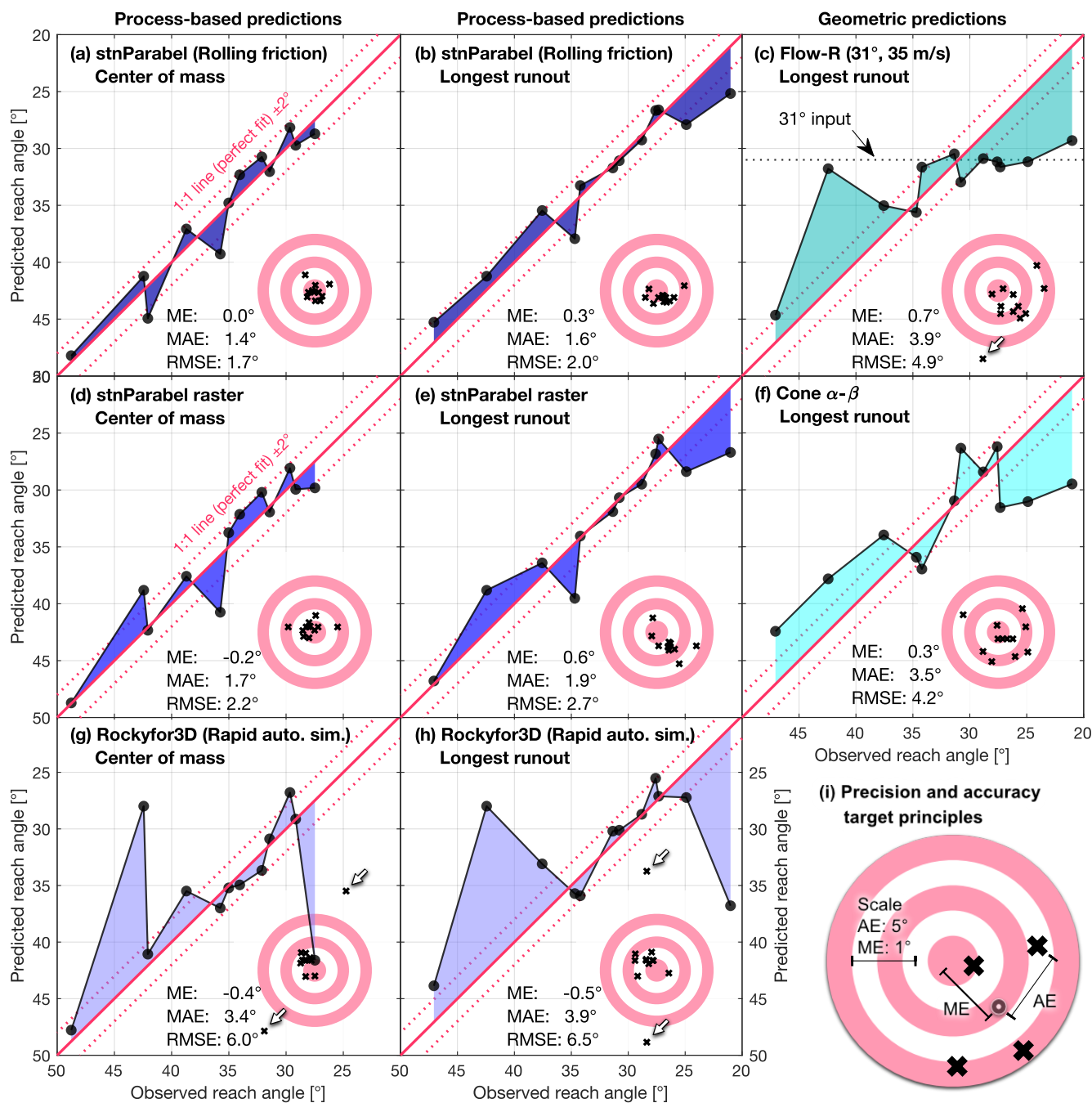
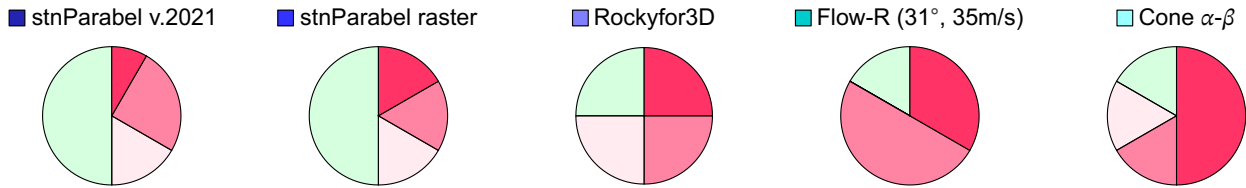


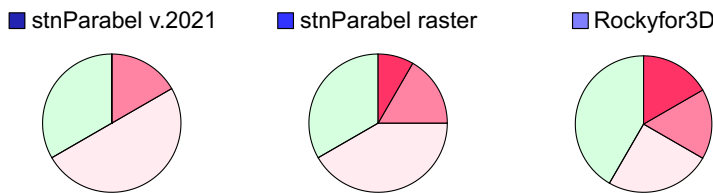
Figure 12: Comparison overview of the average and longest runout distances expressed as reach angles (angle with horizontal and the source-deposited position line of sight). Note that the axes are inverted for homogeneity across the paper. Predicted values above the 1:1 line (negative ME) overestimate the runout distances due to their gentler reach angle, which is safe but more constraining.

longest runout distances, with a mean error of about half of stnParabel raster or Rockyfor3D. The negative ME for

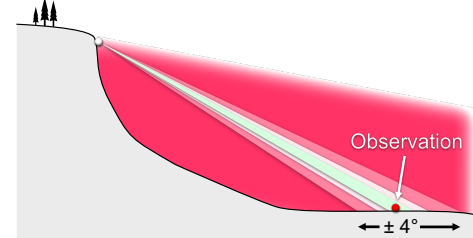
(a) Absolute error (AE) at predicting the longest mapped runouts



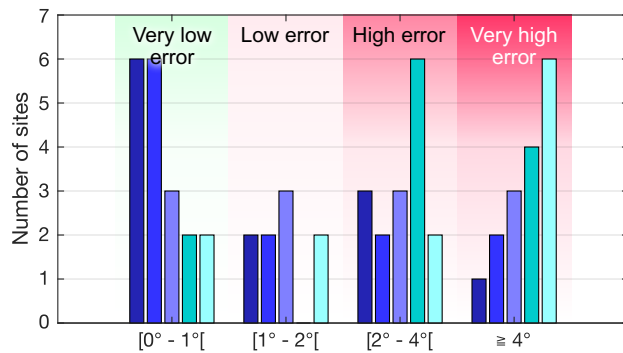
(b) Absolute error (AE) on the average mapped runouts (by mass)



(c) Conceptual representation of the AE



(d) Absolute error (AE) on the longest runouts



(e) Absolute error (AE) on the average runout (by mass)

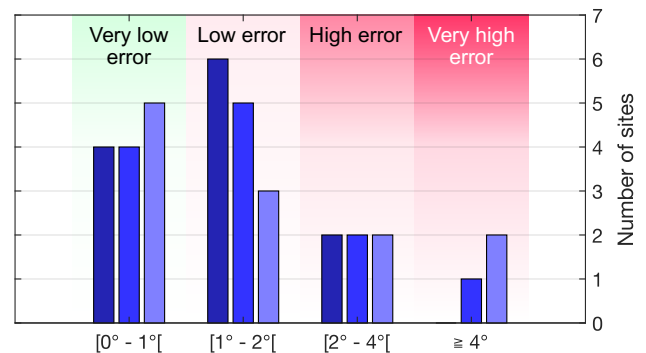


Figure 13: Summary overview of the quantitative predictive performance of the models at predicting the maximal and average observed runout distances of the twelve sites. The performances categorized per model are shown in the form of pie charts for the longest runouts in (a) and for the average runout distances in (b). The same performances, but categorized per performance classes, are shown in the form of histograms in (d) and (e). However, the predictive performance can't be simplified to solely these two metrics. One should thus complement their model choice following the qualitative validation steps with the numerous maps and distributions of deposits of the result section (and given in supplementary material for the other models).

Rockyfor3D means that it slightly exaggerates the maximal runout distances on average, which is safe. It is surprising that stnParabel has a double accuracy in comparison to stnParabel raster (ME of 0.3 vs. 0.6) despite using the same Rolling friction rebound model. This effect may be attributed to the slightly rougher terrain from the use of 3D point cloud DTMs instead of gridded DTMs (rasters). It would be interesting to confirm this hypothesis (and improve stnParabel raster) by repeating the simulations with stnParabel raster while adding subtle random terrain roughness before each impact to emulate the effect of the 3D point clouds. Like for the validations for the capacity at covering the mapped observed rock fragments of Noël and Nordang (2025) mentioned in the introduction (Fig. 2), Flow-R's apparent performance at predicting the longest



runout seems again behind in terms of accuracy (ME of 0.7), but this time only by a narrow margin. However, this gap may increase with its open-source counterpart Flow-Py, given a D8 grid bias that artificially shortens runouts by about 2° in odd grid directions (Noël et al., 2023b). Also, as shown in the following paragraphs, the perceived accuracy may considerably vary when some sites are excluded depending on the precision of the predictions.

Regarding the precision of the methods at predicting the longest runout distances, the Rolling friction rebound model with stnParabel and stnParabel raster performed relatively well. Indeed, their mean absolute error (MAE) below 2° is approximately half of those of the other methods, resulting in double the precision. This conveniently implies that practitioners would not need to refine their zonation limits significantly beyond those predicted by stnParabel and stnParabel raster. The same is also partly true for Rockyfor3D, although its MAE may suggest the opposite. Indeed, with a precision that seems as limited as for Flow-R (MAE of 3.9°), one can see on Fig. 12h that this score is mostly penalized by two outliers from Rockyfor3D, for the Glomset and Tschamut sites. Apart from these, most of Rockyfor3D's predictions are similar to those of stnParabel: the process-based predictions follow the trend of the 1:1 perfect fit line, and are clustered close to the target centre. The geometrical predictions from the Cone $\alpha - \beta$ method (Fig. 12f) exhibit a similar trend, albeit not precisely. Conversely, those obtained from Flow-R (Fig. 12c) tend to align more closely with the 31° input rather than accurately predicting the observed reach angles.

If the under-sampling carried out to obtain samples of comparable sizes had been neglected, the likelihood of simulating longer runouts than the observations would have been high because of having much more process-based simulation attempts than for the observations. For transparency, such biased comparison gives a mean error (ME) on the longest distances by reach angles of -2.3° for stnParabel with a MAE of 3.1° when only keeping the sites with at least 10 observations. As presumed, this comparison gives the biased impression that the simulations overestimate the maximum runout distances. But this can also simply be attributed to the insufficient number of observations. Indeed, even if greater than 10 rock fragments per site, this is not necessarily sufficient to conclude that the maximum distances expected for each site have been observed, as explained by Dorren et al. (2013). Comparing the 99th percentile of the longest simulated runouts by reach angle (1st percentile of the reach angles), the bias is already greatly reduced with a ME of -0.2° and a MAE of 1.9°.

When properly comparing corresponding sample sizes for the nine sites with more than ten observations, the mean error of stnParabel simulations is reduced to 0.1° (MAE of 1.6°) for the longest runout distances, which corresponds to near-perfect accuracy and relatively good precision. For the same precision, stnParabel raster predictions have a slightly lower accuracy with a ME of 0.6° in that case. Given the relatively precise predictions of stnParabel and stnParabel raster for all sites (MAE of 1.6°), excluding a few sites barely changes their ME. In contrast, the corresponding ME for Rockyfor3D's predictions is shifted to 0.9° (MAE of 3.5°). This is a significant shift from -0.5° to 0.9° when neglecting the sites with fewer observations. This is because the outlier prediction of Rockyfor3D for the Glomset site is this time excluded and thus does not counterbalance the other outlier result of the Tschamut site. The ME of the geometrical predictions are this time shifted to 0.8° and 1.8° (MAE of 4.0° and 3.7°) for the Cone $\alpha - \beta$ and Flow-R methods respectively. This again illustrates how their relative imprecision affects the perceived accuracy of their predictions whether some sites are excluded or not. For



transparency, it should be reminded that the Rolling friction model has been developed from the impacts of the reconstructed trajectories of the Mel de la Niva and Riou Bourdoux sites (Fig. 1b). But when excluding these sites and those with fewer than ten observations, the good accuracy and precision of this model remain relatively stable, with ME of -0.1° and 0.2° (MAE of 1.9° and 2.1°) for stnParabel and stnParabel raster respectively (and ME of 0.3° and 0.7° ; MAE of 1.9° and 2.3° for all sites but the Mel de la Niva and Riou Bourdoux sites).

5 Implications in rockfall assessments

In most cases, although at different degrees, the positive mean error inaccuracies of the method tested show that they tend to underestimate the maximum runout distances more than they overestimate them. This can be problematic if the least precise methods are used for susceptibility purposes without being adjusted on a per-site basis. A workaround could consist of extending the resulting inaccurate zonation limits by the equivalent of the ME of the chosen method, plus a buffer of one or two MAE as safety margins, especially if no accurate and precise method is available. To avoid excessively constraining susceptibility zonation, one should prefer precise and accurate approaches, such as those based on process-based rockfall simulated trajectories for the most cases (e.g. Noël and Nordang, 2025). Also, with similar validation methodology, future versions and alternative simulation models could also have their precision and accuracy objectively quantified, and their general behaviour qualitatively assessed. This way, practitioners could roughly know by how much their zonation limits should be refined from field observations with judgement.

The accuracy becomes even more important as the complexity of the assessments increases, for indicative hazard, hazard, and risk assessments, for example, and for the design of mitigation measures. In these cases, rockfall statistics such as the local number of passing trajectories are necessary for evaluating the potential local frequencies at which an exposed object may be reached. Velocities and bounce heights are also required for fulfilling intensity criteria or designing mitigation measures. As shown, the extent of the detailed guiding hazard zones from the process-based simulations with the Rolling friction rebound model naturally variates considerably from site to site. The process-based simulations utilizing the roughness available from the detailed terrain models managed to reproduce the observed variability, as shown by their relative improved accuracy and precision for predicting both the mean runout distances (by mass or by rocks) and the maximum runout distances. With comparable cumulative curve trends, this also applies to a certain extent to any values picked in between the mean and maximum runout distances or across the lateral dispersion. This way, the frequencies can be properly carried by the simulated trajectories for statistically valid predicted reach values, related hazard and risk values, and for estimating mitigation measure lengths. And, the process-based simulations not only predict runout distances, but also their spatial distributions by simulating individual propagating rock's trajectories in 3D and potentially the rock-tree interactions across forested areas.

Moreover, at the same time as relatively precisely predicting the runout distances, it has been shown that the velocities and bounce heights are properly predicted across most of the transit areas of the test sites with reconstructed trajectories, and to



565 the reconstructed trajectories from the Mel de la Niva in Noël and Nordang (2025). The novel quantitative validation
approach across many evaluation screens over the whole sites allows a spatial validation of the predicted velocities and
bounce heights. Moreover, placing the compared simulations to the same level as the observations by comparing similar
sample sizes helps cope with some observational biases. Again, this highlights the important value of the reconstructed
trajectories and the collaborative rockfall experiments performed by the WSL with GeoBrugg (Caviezel et al., 2019, 2021),
570 Hibert et al. (2024), and Bourrier et al. (2021). Our presented validation methodology can be applied with other simulation
models and future versions, as long as they output non-gridded 3D simulated trajectories. With the validated results from
stnParabel with the Rolling friction rebound model, one can confidently choose characteristic energies and heights for
mitigation designs from simulations of a chosen reference block (assuming that the validated behaviours of the Chant Sura
and Riou Bourdoux sites can be transposed to other sites). And the quantified precision and accuracy can provide arguments
575 for justifying the associated safety margins used for passing from the characteristic values to the design / service values.
With this, and also by properly accounting for the detailed geometry of the terrains and impacted structures thanks to the
impact detection algorithm (Noël et al., 2021), several modelling approach limitations highlighted by Lambert and Bourrier
(2013) are resolved with stnParabel's approach with the Rolling friction rebound model (Noël et al., 2023a).

As shown with stnParabel's simulations, the Rolling friction rebound model combined with the impact detection algorithm
580 that works on point clouds produced a simulation model that manages to objectively reproduce the observed rockfall
experiments and events to a high degree prior to any calibration or fine-tuning of the parameters. Thus, independently of its
initial experience, a practitioner would objectively obtain the same guiding results close to the expected reality for this wide
range of sites and events. Combined with field observations, the results could be used to set boundaries/margins inside which
fine-tuned simulations and finalized manually drawn hazard zonation are expected. Such a use case can be particularly useful
585 when little is known about the past rockfall events for a site and their possible runout distances. Additionally, those
reasonable boundaries can also be useful for independent third-party quality control. Furthermore, the variability of the
results produced from user to user would be greatly reduced with the increased predictability of the simulation model. With
predictable runout evaluations from simple parameters, one could eventually redirect some resources toward the important
collection of on-field observations, the non-trivial evaluation of the frequency of rock failure events, and the expected
590 corresponding frequency of rock fragments.

6 Conclusions

Numerous back analyses of previous rockfall experiments and real events were objectively and quantitatively compared to
rockfall simulations from the stnParabel freeware (Noël, 2020) using the Rolling friction rebound model proposed by Noël et
al. (2023a) combined with the impact detection algorithm that works with point clouds proposed by Noël et al. (2021). The
595 performance of the rockfall simulations was compared to those of other process-based models and simple geometric
methods. Doing so, comparison methods were given and used to qualitatively and quantitatively assess the predictive



performance of different approaches while coping with different biases. The results showed that the stable accuracy of the runout predictions from our process-based simulated trajectories is generally superior since it remains relatively stable from site to site. Moreover, for the wide variety of sites tested, the precision of stnParabel's simulations is improved by $2\times$ to $3\times$ compared to all other tested methods prior to any calibration or fine-tuning of the simulations from stnParabel using the Rolling friction rebound model. Also, combined with the relatively good performance at predicting realistic runout distances and lateral dispersions, the predictions on the bounce heights and translational velocities (and energies by extension) are greatly improved compared to past simulation benchmarks.

As previously foreseen (Noël et al., 2023a) and suggested by the results from Noël and Nordang (2025) (Fig. 2), the Rolling friction rebound model does not seem to require extensive per-site calibration. The impact detection algorithm that properly perceives the initial terrain geometry and roughness from detailed terrain models is probably a large contributor to this objective predictability. The objective simple parametrization of the model thus improves the predictability and objectivity of the simulations, which can help homogenize zonation boundaries while providing realistic bounce height and velocity predictions, increase the applicability of process-based simulated trajectories over large regions, and potentially help practitioners save precious time and resources. More emphasis could thus be allocated to the important collection of on-field observations, on the non-trivial on-field evaluation of the frequency of rock failure events and its expected corresponding frequency of rock fragments.

Despite the considerable improvements in terms of objectivity and predictability of the simulations, a model remains a simplified attempt at reproducing the complex nature of the phenomena. As with any model, one must thus remain critical of the simulated results. The relatively good precision within $\pm 1-2^\circ$ of the average and maximum runout distances could still be improved further. The same is true for the mean error of a few metres per second on the velocities and of a few decimetres for the bounce heights. But overall, although simplified, the Rolling friction rebound model's consideration of sliding and rotation mechanisms combined with its redirection of the returned velocities emulating the combined effect of scarring and rock shape mechanisms provided good predictions. This suggests that newer model development should also explore alternative avenues than those solely based on the traditional normal and tangential components with damping functions. Indeed, given the generally complex and highly dynamic rock-ground interactions at impact, limitations of overly simplified traditional approaches that neglect the instantaneous contact surface and use damping coefficients R_N - R_T are likely to require per-site adjustments and cautious result interpretations (Labieuse and Heidenreich, 2009). Independently of using newer or traditional approaches, the reader is encouraged to objectively confront models to various real events and to constructively share their findings. This would thus help the rockfall community to be aware and understand the level of precision and accuracy of the tools used. We also thank the reader for reporting any residual glitch if any is found with stnParabel.

Further work should focus on completing the compared "observational" datasets of reconstructed rockfall trajectories and impacts to also include the slower rock-ground interaction segments. More details could also be attributed relatively to the



630 material involved to better document their remaining subtle effect. On the simulation model side, the impact detection algorithm could be enhanced to reduce the potential remaining roughness artificially introduced when the algorithm fails to fit a surface through the point cloud. This may help increase slightly the simulated velocities and reduce the bounce heights. Refining the Rolling friction rebound model from the finalized reconstructed impact data from the Mel de la Niva 2015 events and the Chant Sura test site could also help improve the model further.

635 **Appendix A: Mapped observations and stnParabel simulation details**

A.1 Rockfall test sites

The SLF Chant Sura test site in Switzerland described in Caviezel et al. (2019), Caviezel et al. (2021), and Noël et al. (2022) is the first rockfall test site of the comparison (T1 in Fig. 3a). Its related data from observations (Caviezel et al., 2020) is used as a reference. The DTM from structure from motion photogrammetry (SfM) from Caviezel et al. (2020) is used for the
 640 simulations. The site characteristics other than the terrain geometry are summarized in Table A1 for all the rockfall test sites, while those for the real rockfall event sites are summarized in Table A2. The corresponding characteristics for the simulations are summarized in Table 1. Two artificial rock shapes (equant EOTA₁₁₁ and disk EOTA₂₂₁) of different sizes were used for the 182 documented controlled rockfalls from a total of eight different rock sizes and shapes. We simulated them proportionally with 55 simulated trajectories for every one of the 182 rockfalls, for a total of 10,010 simulated
 645 trajectories (Table 3). The corresponding rock size and shape from the eight different ones was chosen for every simulation for a proper comparison with the observations. The rocks are characterized by the two turquoise points in the rock shape diagram shown in Fig. 3b. Note that their d_3/d_2 ratios are not exactly 0.5 and 1.0 once the rock dimensions are determined from a bounding box aligned on the principal axes of inertia using the rock block geometry method described by Noël et al. (2022). The rocks for the Riou Bourdoux (T2), Tschamut (T3), Mel de la Niva (R1), and La Verda (R6) sites were measured
 650 from 3D models in the same way and are shown next to each other in Fig. 6 of Noël et al. (2023a).

The Riou Bourdoux test site in France featured in Hibert et al. (2017), Noël et al. (2018), Noël et al. (2022), Noël et al. (2023a), and Hibert et al. (2024) is the second test site for the comparison (T2 in Fig. 3a). Its related data from observations (Hibert et al., 2024) is used as a reference, as well as the related DTM from terrestrial LiDAR surveys (TLS) and SfM photogrammetry for the simulations. From the 31 initial rocks released (+1 artificial EOTA₁₁₁ instrumented metal shape here
 655 ignored), two rocks were instrumented with accelerometers and gyroscopes. A total of 34 end locations of the deposited rocks and larger fragments were measured. It should be noted that the Rolling friction model of stnParabel has been developed and empirically calibrated mostly from the reconstructed data from those experimental rockfalls (Fig. 1b), together with preliminary versions of the reconstructed data from the 2015 rockfall event from the Mel de la Niva (Noël et al., 2023a). As for the Chant Sura site, the 31 rockfalls were simulated proportionally, this time with 320 simulated
 660 trajectories per unique rock, for a total of 9920 simulated trajectories (Table 1).



Table A1: Rockfall test site characteristics.

Sites	Nr. of total obs. rock fragments*	Avg. rock mass; range [kg]	Observed terrain materials			References
			Source vicinity	Transit zone	Deposition zone	
Chant Sura ($<\pm 1$ m) ⁶	n_{tot} : 182 n_{mass} : 182 n_a : 182	Avg.: 890 Min.: 45 Max.: 2670	Outcrops of hard meta granitoid / dwarf shrubs	Alpine meadow interspersed with rocks	Blocky scree / rough scree field	Caviezel et al. (2019); Caviezel et al. (2021)
Riou Bourdoux ($<\pm 1$ m) ⁷	n_{tot} : 34 n_{mass} : 34 n_a : 34	Avg.: 164 Min.: 55 Max.: 427	Soft black marl and moraine	Soft black marl	Debris flow fan, limestone rocks	Hibert et al. (2024)
Tschamut ($<\pm 1$ m) ⁷	n_{tot} : 102 n_{mass} : 102 n_a : 102	Avg.: 58 Min.: 19 Max.: 79	Shallow soil covering bedrock, some exposed bedrock			Volkwein et al. (2018)
Authume B ($<\pm 1$ m) ⁷	n_{tot} : 56 n_{mass} : 56 n_a : 56	Avg.: 707 Min.: 170 Max.: 1869	Medium soft quarry waste, mixing sand, clay, and limestone fragments			Bourrier et al. (2021); Garcia et al. (2022)
Authume A ($<\pm 1$ m) ⁷	n_{tot} : 46 n_{mass} : 46 n_a : 46	Avg.: 791 Min.: 178 Max.: 1751	Numerous boulders	Dry newly deposited quarry waste, mixing sand, clay, and limestone fragments. Compact quarry waste		Bourrier et al. (2021); Garcia et al. (2022)

* n_{tot} : total number of mapped rock fragments; n_{mass} : number of mapped rock fragments from which the dimensions and masses are known, used for defining the deposit centre of mass and cumulative curves by mass, and for the simulations; n_a : number of rock fragments used for defining the deposit's average positions, and the cumulative curves by rocks.

Precision estimated from: ¹GNSS, ²national orthophoto, ³SfM orthophoto rectified from national orthophoto + elevation from national ALS data, ⁴national ALS data, ⁵SfM point cloud with GCP from national ALS data and then rectified by ICP with national ALS data, ⁶RTK, ⁷total station.

The third site (T3 in Fig. 3a), Tschamut in Switzerland, is described in Volkwein et al. (2018). Its observation data and DTM used as reference and for the simulations are available via Volkwein and Gerber (2018). On average, this site has the smallest and lightest rocks (Table A1). The lower section of the site in the deposition zone is very smooth and was freshly mowed before performing the rockfall experiments (Volkwein et al., 2018). The six different rocks were thrown repetitively a various number of times from different sources linearly spread at the top of the site (along a roadside) to totalize 102 controlled rockfalls. As before, they were simulated proportionally with 98 trajectories per controlled rockfall, for a total of 9996 simulated trajectories (Table 1).

The Authume test site in France, featured in Garcia (2019), Bourrier et al. (2021), and Garcia et al. (2022), is used as the last test site of the comparison and is separated into two profiles (A and B, respectively T5 and T4 in Fig. 3a). The observation data and SfM DTM from Bourrier et al. (2021) are used as a reference and for the simulations. Unlike for the other test sites,



the dimensions of the rocks were directly taken from the provided $d_1 - d_2 - d_3$ values in Bourrier et al. (2021) (Fig. 3b). For the profile A, 46 controlled rockfalls with different rocks were documented. For the profile B, 56 controlled rockfalls from 46 different rocks were documented. They were again simulated proportionally, with 100 trajectories per controlled rockfall, for a total of 10,200 simulated trajectories (4600 and 5600 for the profiles A and B respectively, Table 1).

675 A.2 Real rockfall event sites

Concerning the real rockfall event sites, the site R1 in Fig. 3a corresponds to the 2015 event from the Mel de la Niva Mountain described in detail in Noël et al. (2023a). The site has already been described and simulated in Noël and Nordang (2025). Nevertheless, its characteristics are summarized in Table A2 as for all the other sites from real events.

The two first real rockfall event sites covered in this paper correspond to the 2020 and 2014 rockfall events, respectively, 680 from Leirskardalen and Fortun in Norway (R2 and R3 in Fig. 3a, Fig. 4, and Table A2). The authors remotely mapped those events from pictures taken shortly after the events, respectively provided by the Nordland County Council (Pettersen, 2020) and the Norwegian Water Resources and Energy Directorate (NVE) (Midtun, 2014). In the first case, the pictures were acquired with a DJI Mavic 2 Pro drone and were geotagged from the drone GNSS and altimeter measurements. For the Fortun case, they were acquired by helicopter using a Nikon D40 with a 55 mm lens (82 mm full-frame equivalent). 685 Although the pictures were not taken with the intention of producing photogrammetric models and derived orthophotos, it was possible to use them for this purpose after carefully selecting the most suited ones and by guiding the reconstruction with the help of markers and masks in the software Metashape Professional (Agisoft LLC, 2018) (Fig. 4).

Three 3D models were produced for the Leirskardalen site (R2 in Fig. 3a), one for an overview of the event (Fig. 4a), one close-up for the deposition zone (Fig. 4c), and one for the release zone (Fig. 4e and f). With geotagged locations, the two 690 first models have respective XY planimetric errors of 2.0 and 9.7 m, and all three correspond relatively well with the DTM from airborne LiDAR (ALS) of the area from Høydedata (2019). Their position and scale were nevertheless refined using the ICP algorithm in the CloudCompare open-source software (Girardeau-Montaut, 2006) and the ALS DTM as a reference. The $d_1 - d_2 - d_3$ dimensions of 12 out of the 19 identified rock fragments could be measured on the close-up 3D model (dense cloud) (Fig. 4c). It was thus possible to estimate the mass for 12 rock fragments from a rock density of 2850 kg m³ 695 (Li, 2017; Myrvang, 2001) and the volume estimated from ellipsoids of corresponding dimensions. After the massive 490 and 552 t rock fragments from the Mel the la Niva case, the rocks from the Leirskardalen 2020 event are the heaviest by average (Table A2).

An orthophoto was produced from the overview SfM model of the Leirskardalen site. Its orthorectification was refined manually in the QGIS open-source software (QGIS Development Team, 2009) to match with the orthophoto of the area from 700 NiB (2022). The XY planimetric positions of the rock fragments and undertaken paths were mapped from those inputs in QGIS (Fig. 4a). The Z elevations of the rock fragments were estimated from the corresponding elevation of the ALS DTM in line with their XY positions. The ALS DTM was used for the simulations, with 833 trajectories for each one of the 12 rock fragments whose dimensions were measured, for a total of 9996 simulated trajectories (Table 1).



Table A2: Real rockfall event site characteristics.

Sites	Nr. of total obs. rock fragments*	Avg. rock mass; range [metric ton]	Observed terrain materials			References
			Source vicinity	Transit zone	Deposition zone	
Mel de la Niva** (±3 m) ^{1,5}	n_{tot} : 449 n_{mass} : 2 n_a : 449, 219 & 11	bl1: 489.7 bl2: 552.1	Bedrock of gray carbonate-silicate schist / thin scree slopes of colluvium	Scree slopes of colluvium / blocky meadows	Grassy alpine montane meadow, morainic gray fine sediments, some rockfall block fragments	Steck et al. (2001); Noël et al. (2023a)
Leirskardalen (±2 m) ^{3,5}	n_{tot} : 19 n_{mass} : 12 n_a : 19	Avg.: 45.5 Min.: 13.3 Max.: 86.2	Bedrock of mica gneiss	Bedrock with thin scree slopes of colluvium and moraine	Scree slopes of colluvium and moraine / cultivated fields	NGU (2022a); NGU (2022b); NiB (2022); Pettersen (2020)
Fortun (±3 m) ^{3,5}	n_{tot} : 26 n_{mass} : 21 n_a : 26	Avg.: 20.1 Min.: 0.3 Max.: 101.1	Bedrock of tonalite to granitic gneiss and phyllite	Bedrock with thin scree slopes of colluvium and moraine	Bedrock with thin scree slopes of colluvium and moraine / cultivated fields	NGU (2022a); NGU (2022b); NiB (2022); Midtun (2014)
Holaviki (±5 m) ^{2,4}	n_{tot} : 2 n_{mass} : 2 n_a : 2	bl1: 40.5 bl2: 17.6	Bedrock of phyllite and mica schist	Bedrock with thin scree slopes of colluvium and moraine	Scree slopes of colluvium / fluvial deposits with cultivated fields	NGU (2022a); NGU (2022b); NiB (2022); Domaas (1995)
Springville (±3 m) ⁵	n_{tot} : 2 n_{mass} : 2 n_a : 2	bl1: 3.2 bl2: 0.3	Bedrock of limestone	Weathered bedrock / lacustrine sand and gravel	Lacustrine sand and gravel	Baker (1973); Erickson (2019)
La Verda (±1 m) ^{1,2,3}	n_{tot} : 13 n_{mass} : 13 n_a : 13	Avg.: 4.3 Min.: 0.2 Max.: 15.5	Bedrock of sandstone and conglomerate	Shallow soil covering bedrock	Shallow soil covering bedrock	© swisstopo; Noël et al. (2021); Lévy and Verly (2019)
Glomset (±2 m) ^{2,4}	n_{tot} : 1 n_{mass} : 1 n_a : 1	bl1: 15.5	Bedrock of garnet mica schist	Thin scree slopes of colluvium and moraine	Blanket of glaciomarine deposits	NGU (2022a); NGU (2022b); NiB (2022); Barstad (2019)

* n_{tot} : total number of mapped rock fragments; n_{mass} : number of mapped rock fragments from which the dimensions and masses are known, used for defining the deposit centre of mass and cumulative curves by mass, and for the simulations; n_a : number of rock fragments used for defining the deposit's average positions, and the cumulative curves by rocks.

**Thirteen rockfall path segments, including the reconstructed 2015 trajectories of the bl1 and bl2 were mapped in Noël et al. (2023a), and 449 rock fragments were mapped by the BEG SA (219 certain). Out of the 13 paths, 11 deposited rock fragments sharing a common source origin were identified. The latter are used for the statistics given in this paper.

Precision estimated from: ¹GNSS, ²national orthophoto, ³SfM orthophoto rectified from national orthophoto + elevation from national ALS data, ⁴national ALS data, ⁵SfM point cloud with GCP from national ALS data and then rectified by ICP with national ALS data, ⁶RTK, ⁷total station.

For Fortun (R3 in Fig. 3a, Fig. 4b and d), the 3D model (dense cloud) was roughly scaled and geolocated in Agisoft



705 Metashape Professional by setting eight ground control points (GCPs) from characteristics 3D features recognizable over the area both in the pictures and on the raw ALS DTM point clouds from Høydedata (2012; 2018; 2020) (e.g., the base where the street light poles meet the ground, the corner of a house chimney, a top corner of a hunting blind, the characteristic sharp edge of a large boulder, etc.). With those, the rough XY planimetric error estimation is of 8.9 m. The 3D model lacks local surface roughness exhibited by a general smoothing of the height of small topographic objects. This is a sign that there is a lack of overlapping photos and that their ground resolutions vary greatly (average of 8.6 cm pix⁻¹) so that it is not possible to resolve the desired level of detail. Nevertheless, an orthophoto was produced from the SfM model following the same methodology as for the Leirskardalen site. The rockfall paths and rock fragments were located on the orthophoto (Fig. 4b). The XY planimetric dimensions of the rock fragments were measured in QGIS. They were completed for the third dimension with estimations from the available photos, including photos from media where people stand next to the rocks (the latter cannot be reproduced here for copyright reasons). Some photos also suit for stereographic visualization which helped validating the estimated dimensions (Fig. 3d). A total of 26 rock fragments were identified, of which 21 could be measured. Of the set of measured rocks, the largest 19 were simulated on the ALS DTM, with 526 trajectories per rock for a total of 9994 simulated trajectories (Table 1).

The rockfall event at the Holaviki site in Norway (R4 in Fig. 3a) was mapped by the Norwegian Geotechnical Institute (NGI) (Domaas, 1995). The previously mapped source locations, rockfall paths, and deposited locations of the two rock fragments from (Domaas, 1995) were digitalized in QGIS for this paper. The d_1 - d_2 - d_3 dimensions of the rock fragments were estimated from the descriptions and dimensions given in (Domaas, 1995) (Fig. 3c). The corresponding volumes of 15 and 6.5 m³ were taken directly from (Domaas, 1995) and used to estimate their mass (Table A2). The two source locations estimated by Domaas (1995) were used for the simulations with their respective rock characteristics. It is, however, possible that they originated from a higher common point. A total of 10,000 trajectories were simulated on the ALS DTM from Høydedata (2016), with 5000 trajectories for each rock-source pair (Table 1).

The 2019 rockfall event that happened at the Springville site in the United States of America (R5 in Fig. 3a) was mapped by the Utah Geological Survey (UGS) (Erickson, 2019). The UGS produced 3D SfM models of the event and the rock fragments (Erickson, 2019). They measured the volume and dimension of the main rock fragment of the event and located the impact marks, the paths undertaken by the rock fragments, and the deposited location of the rock fragments. This great detailed mapped information from the UGS was directly used in this paper (Table A2). We roughly estimated the mass of a smaller fragment that detached from the main falling rock but did not consider it for the simulations as its dimensions could not be precisely estimated. The public domain ALS DTM dataset of the area (State of Utah Acquired LiDAR Data - Wasatch Front, 2015) was used for the simulations. A total of 10,000 trajectories were simulated for the main rock fragment (Table 1).



Concerning the 2019 rockfall event at the La Verda site (R6 in Fig. 3a), the site was quickly visited in person a few months after the event. Due to limited equipment and time on-field, most rockfall paths were quickly walked up and down and roughly documented with pictures and video footage taken with a smartphone (Samsung S7 Edge). Despite their limited quality, the captured pictures and video footage helped identify on the 2020 SWISSIMAGE orthophoto by ©swisstopo the subtle corresponding features like the impact marks with their characteristic surrounding landmarks. The rockfall paths were digitalized based on the orthophotos in QGIS as well as the position of 13 visible deposited rock fragments. Additional information about the rockfall event was provided by the General Environment Directorate of the Vaud Canton of Switzerland (DGE) (communication with Lévy and Verly, 2019). It consists of complementary estimated positions of the source and of seven deposited rock fragments and was used for validation.

The characteristics of the seventh largest rock fragment were obtained by photogrammetry with Agisoft Metashape Professional. Their 3D SfM models were built from 20 to 50 pictures per rock (Fig. A1). The geotagged locations from the smartphone were ignored due to their imprecision, so the non-pro version of the software would have sufficed in that case. A bike pannier was used to roughly scale the seven models in CloudCompare. The scales were then refined to match the dimensions measured on the orthophoto. The models were cleaned manually to remove artefacts and to close the hidden base of the rocks as in (Noël et al., 2023a; Hibert et al., 2024). The volume and mass of the rock fragments were estimated from their 3D shapes with a volumetric density of 2700 kg m^{-3} (Table A2). Their dimensions were measured with a bounding box aligned with the principal axes of inertia of the rocks estimated with the MeshLab open-source software (Cignoni et al., 2008) (Fig. 3c). For the smaller rock fragments, their d_1 and d_2 were estimated from the planimetric dimensions measured on the orthophoto. Their d_3 dimensions were extrapolated from their d_2 dimensions using the average d_3/d_2 ratio of the 3D models of the larger rock fragments (Fig. 3c). Their masses were estimated from ellipsoids. The seventh largest rock fragments with characteristics obtained from 3D SfM models were simulated on the ALS DTM from the swissALTI3D product by ©swisstopo. A total of 10,010 trajectories were simulated, with 1430 trajectories per rock fragment (Table 1). The last real rockfall event site of Glomset in Norway (R7 in Fig. 3a) was roughly mapped from pictures of the 2019 event covered in the media and from the report of the event by Multiconsult Norge AS (Barstad, 2019) (Table A2). A total of 10,000 trajectories were simulated on the ALS DTM from Høydedata (2015) for the main rock fragment (Table 1).

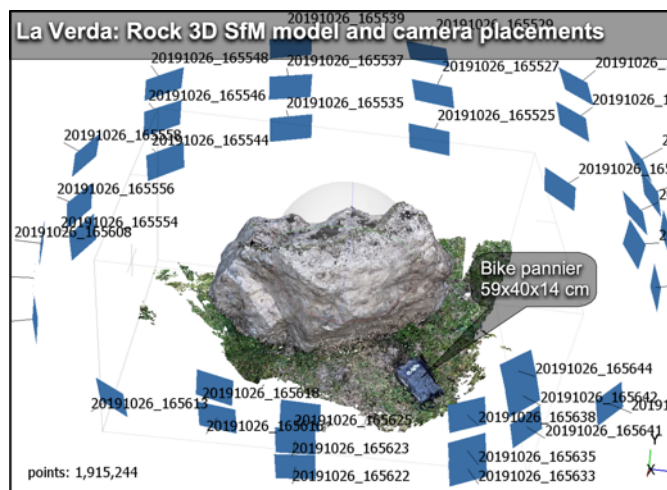


Figure A1: Example of a rock fragment 3D model from SfM photogrammetry shown with its cameras. This rock is shown among those from the other sites in (Figure 6 in Noël et al., 2023a).



Appendix B: Approaches with the other models

770 The methodology with stnParabel raster is basically the same as with stnParabel (Section 2 and Appendix A), apart from the gridded DTM inputs instead of the point clouds, mostly. The gridded DTMs were produced from the 3D point clouds DTMs used by stnParabel's simulations. The cell-sizes summarized in Table 1 were chosen to preserve the roughness details of the terrains in relation to the size of the simulated rock fragments. The same source and simulation settings files were read by the scripts, and the same default parameters were used. In case stopping criteria were met at the source of some rockfall test

775 sites, preventing any propagation from happening, the source points were raised by up to 2 m in elevation. This was, of course, not needed for the real rockfall event sites, given the presence of natural cliffs at the source locations.

Regarding Rockyfor3D v2.5.15 simulations, the raster inputs were produced like for stnParabel raster, but this time sampled to a 2 m cell-size by averaging. The source 3D points from the observations used with stnParabel were rasterized and provided to Rockyfor3D in the form of the *rockdensity.asc* input. For the test sites, an additional elevation of up to 2 m was

780 given in case stopping criteria were met too early. The rock dimensions were provided through the graphical interface for each simulated rock scenario (Table A1 and Table A2), and ellipsoid rock shapes were used. Damping terrain roughnesses Rg_{70} , Rg_{20} , and Rg_{10} , related to R_T , not the geometrical roughness θ_{max} , and soil types were objectively set based on the rapid automatic simulations slope thresholds. The number of simulations per rock scenario was set, as for the other process-based simulations from stnParabel, to target 10,000 simulations per site when all related scenarios are combined. For the

785 Glomset site, the obtained simulated runouts were extreme, with most reach angles between 19° and 27° . In a subjective attempt to improve the situation, a second run of simulations was performed, but this time by attributing soil types for water to prevent the trajectories from propagating further than the shoreline at the base of the Glomset terrains ($\sim 28^\circ$). The simulated deposited rock fragments were converted from raster to points (with a number of points per cell given by the count of deposited rock fragments). The points were then used to produce the cumulative curves and statistical analyses as for

790 stnParabel's results, with elevations sampled from the detailed DTMs used for stnParabel raster.

Regarding the geometrically obtained maximal runout extents, two geometric approaches were used to predict the expected reach angles. The first one consists of the alpha-beta ($\alpha - \beta$) method (Keylock and Domaas, 1999; Derron et al., 2016) whose extent is obtained by projecting the α angle with a cone (Jaboyedoff and Labiouse, 2011). That method used the 2D profiles of the sites (Fig. 3a) smoothed with an averaging window of a radius of 10 m to find the beta points where the slopes

795 at the base of the profiles equal 23° . The elevations of the original non-smoothed profiles in line with the beta point were then used to evaluate the beta angles (β). The predicted alpha reach angles (α) were then obtained with Eq. (A1) modified from (Keylock and Domaas, 1999; Derron et al., 2016) as follows:

$$\alpha = 0.77\beta + 3.9^\circ \quad (A1)$$

The second geometric method, with Flow-R v2.1.0, simply consists of using a fixed reach angle, but to limit the height of the

800 related propagating energy line above the slope profile to never exceed a maximum value (h_{max}). This limiting height can be



expressed in the form of a pseudo velocity cutoff (v_{pseudo}) with the following relation (Jaboyedoff and Labiouse, 2011; Dorren et al., 2013):

$$h_{max} = \frac{v_{pseudo}^2}{2g} \quad (A2)$$

With this method, also used by the open-source Flow-Py model, the predefined fixed reach angle is steepened if steep and high topographies are encountered, which shorten the predicted runouts as a result. The predicted reach angle thus cannot become gentler than its initial predefined value. Regarding the spatial extent, instead of using a cone, it is this time given by the hydrological soil moisture gridded flow runoff of Holmgren (1994). Like a queen on a chessboard, the soil moisture quantity status can be updated to its lateral and diagonal neighbouring cells. The distributed moisture quantity status is proportional to the height difference of the cells and controlled by an exponent parameter (x). The runout reachable extent is given by two conditions: 1) all reachable cells must have a predicted soil moisture remaining above a given susceptibility threshold, and 2), the height of the cells along the moisture gridded paths leading to all reachable cells must remain below their respective energy line heights. Since the gridded moisture paths are not in a straight line, shorter runout extents are likely to be obtained in comparison to the Cone $\alpha - \beta$ method for comparable reach angle values.

For the complementary runout extents obtained with this method, a 10 m cell-size grid was used with an initial reach angle value of 31° combined with a limiting height (h_{max}) of 62 m, which roughly corresponds to a pseudo velocity of 35 m s^{-1} based on Oppikofer et al. (2024). For the diffusive soil moisture status, the exponent parameter was set to one, and the minimal amount of soil moisture susceptibility threshold was kept to its default value. Then, for evaluating the obtained runouts in terms of reach angle, the edges of the reachable cells in the deposition zones were used instead of the cell centre to avoid reducing the runout extents. As a result, the corresponding reach angles were slightly lowered, which favoured Flow-R's results. The runout extent edges were converted to closely spaced points (1 m spacing), with reach angles obtained from sampling the elevation on the detailed DTMs used with stnParabel raster. For each site, the point with the lowest reach angle was used.

Code and data availability

The mapped observations including their sources, transit paths, and deposited rocks (including their d_1 , d_2 , and d_3 attributes) when available for the sites where we were involved will be openly shared via a Zenodo repository upon publication. Those from other works are available via their respective publications or upon request to their respective authors.

The alpha version of stnParabel (v.August 2021) used for the rockfall simulation examples is freely available via <https://stnparabel.org> or upon request to the first author.



830 **Supplement**

The supplement related to this article is available online.

Author contributions

FN and SFN conceptualized the research. FN and SFN oversaw the data curation. FN did the formal analysis. FN and SFN contributed to the investigation. FN and SFN developed the methodology. FN oversaw the project administration. FN and
835 SFN provided the resources. FN developed the software. FN and SFN supervised the experiment and the research. FN and SFN validated the approach. FN and SFN designed and produced the figures. FN wrote the original draft. FN and SFN reviewed and edited the original draft.

Competing interests

The contact author has declared that none of the authors has any competing interests.

840 **Disclaimer**

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Funding

The open access to this article was financed by University of Lausanne. Otherwise, this research received no additional
845 external funding and has been realized mostly on the author's spare time.

Acknowledgements

We would like to acknowledge Michel Jaboyedoff and Jacques Locat for their constructive feedback, which significantly enhanced the quality of this manuscript. Furthermore, we would like to thank all participants to the rockfall experiments, and Andrin Caviezel, Axel Volkwein, Bruna Garcia, Franck Bourrier, the Canton of Valais, the BEG AS, Sølve Pettersen, Priska
850 H. Hiller, Vebjørn Opdahl, Yngve Midtun, Ulrik Domaas, Ben Erickson, Sébastien Lévy, and Mélanie Verly for sharing their mapped observations and data. Finally, we acknowledge the anonymous reviewers for their time and constructive comments.



855 References

- Abbruzzese, J. M. and Labiouse, V.: New Cadanav Methodology for Rock Fall Hazard Zoning Based on 3D Trajectory Modelling, *Geosciences*, 10, 434, <https://doi.org/10.3390/geosciences10110434>, 2020.
- Agisoft LLC: Metashape Professional, 2018.
- Antoch, J., Hušková, M., and Sen, P. K.: Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis : A Festschrift in honor of Professor Jana Jureš, 182–193 pp., 2010.
- 860 Asteriou, P., Saroglou, H., and Tsiambaos, G.: Geotechnical and kinematic parameters affecting the coefficients of restitution for rock fall analysis, *International Journal of Rock Mechanics and Mining Sciences*, 54, 103–113, <https://doi.org/10.1016/j.ijrmms.2012.05.029>, 2012.
- Asteriou, P., Zekkos, D., and Manousakis, J.: Fully remote assessment of rockfall incidents based on crowdsourced imagery, *Bull Eng Geol Environ*, 84, 204, <https://doi.org/10.1007/s10064-025-04218-x>, 2025.
- 865 Azzoni, A., Rossi, P. P., Drigo, E., Giani, G. P., and Zaninetti, A.: In situ observation of rockfall analysis parameters, in: International symposium on landslides, Christchurch, Citation Key: azzoni1992situ, 307–314, 1992.
- Baker, A. A.: Geological map of the Springville quadrangle, Utah County, Utah, 1973.
- Barstad, O. H.: Akuttvurdering Rabben 3 Glomset (10214790-RIGberg-NOT-001), Multiconsult Norge AS, Oslo, Norway, 870 2019.
- Beladjine, D.: Étude expérimentale du processus de collision d'une bille sur un empilement granulaire : application au mécanisme de saltation dans le transport éolien, Université de Rennes 1, 2007.
- Berger, F. and Dorren, L. K. A.: Objective comparison of rockfall models using real size experimental data, in: Proceedings of the INTERPRAEVENT International Symposium, Mitigation of Debris Flows, Slope Failures and Landslides, Niigata, Japan, 25-29 September 2006, Citation Key: Berger2006, 245–252, 2006.
- 875 Berger, F., Corominas, J., Lopez Carreras, C., Brauner, M., Kienholz, H., and Bartelt, P.: RockFor project - Efficiency of the protective function of mountain forest against rockfall : elements for strategic planning tools dedicated to a forest sustainable mitigation against rockfall, Cemagref, Grenoble, France, 2004.
- Berger, F., Martin, R., Auber, B., and Mathy, A.: Etude comparative, en utilisant l'événement du 28 décembre 2008 à Saint Paul de Varces, du zonage de l'aléa chute de pierre avec différents outils de simulation trajectographique et différentes matrices d'aléa., Cemagref UREMGR, Grenoble, France, 2011.
- 880 Bourrier, F.: Modeling the impact of a boulder on a natural soil, application to rockfall trajectory simulations., Institut National Polytechnique de Grenoble - INPG, Grenoble, France, 246 pp., 2008.



- Bourrier, F. and Acary, V.: Predictive Capabilities of 2D and 3D Block Propagation Models Integrating Block Shape
885 Assessed from Field Experiments, *Rock Mechanics and Rock Engineering*, 55, 591–609, <https://doi.org/10.1007/s00603-021-02696-5>, 2022.
- Bourrier, F., Dorren, L., Nicot, F., Berger, F., and Darve, F.: Toward objective rockfall trajectory simulation using a stochastic impact model, *Geomorphology*, 110, 68–79, <https://doi.org/10.1016/j.geomorph.2009.03.017>, 2009.
- Bourrier, F., Toe, D., Garcia, B., Baroth, J., and Lambert, S.: Experimental investigations on complex block propagation for
890 the assessment of propagation models quality, *Landslides*, 18, 639–654, <https://doi.org/10.1007/s10346-020-01469-5>, 2021.
- Buzzi, O., Giacomini, A., and Spadari, M.: Laboratory Investigation on High Values of Restitution Coefficients, *Rock Mechanics and Rock Engineering*, 45, 35–43, <https://doi.org/10.1007/s00603-011-0183-0>, 2012.
- C2ROP, A. participants of A. 3. 1: Benchmark des approches d'analyse trajectographique par analyse comparative de simulations prédictives et d'essais de terrain, *Revue Française de Géotechnique*, 6, 12,
895 <https://doi.org/10.1051/geotech/2020015>, 2020.
- Caviezel, A., Demmel, S. E., Ringenbach, A., Bühler, Y., Lu, G., Christen, M., Dinneen, C. E., Eberhard, L. A., von Rickenbach, D., and Bartelt, P.: Reconstruction of four-dimensional rockfall trajectories using remote sensing and rock-based accelerometers and gyroscopes, *Earth Surface Dynamics*, 7, 199–210, <https://doi.org/10.5194/esurf-7-199-2019>, 2019.
- Caviezel, A., Demmel, S. E., Bühler, Y., Ringenbach, A., Christen, M., and Bartelt, P.: Induced Rockfall Dataset #2 (Chant
900 Sura Experimental Campaign), Flüelapass, Grisons, Switzerland, EnviDat, <https://doi.org/https://doi.org/10.16904/envidat.174>, 2020.
- Caviezel, A., Ringenbach, A., Demmel, S. E., Dinneen, C. E., Krebs, N., Bühler, Y., Christen, M., Meyrat, G., Stoffel, A., Hafner, E., Eberhard, L. A., Rickenbach, D. von, Simmler, K., Mayer, P., Niklaus, P. S., Birchler, T., Aebi, T., Cavigelli, L., Schaffner, M., Rickli, S., Schnetzler, C., Magno, M., Benini, L., and Bartelt, P.: The relevance of rock shape over mass—
905 implications for rockfall hazard assessments, *Nature Communications*, 12, 5546, <https://doi.org/10.1038/s41467-021-25794-y>, 2021.
- Chau, K. T., Wong, R. H. C., and Wu, J. J.: Coefficient of restitution and rotational motions of rockfall impacts, *International Journal of Rock Mechanics and Mining Sciences*, 39, 69–77, [https://doi.org/10.1016/S1365-1609\(02\)00016-3](https://doi.org/10.1016/S1365-1609(02)00016-3), 2002.
- 910 Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., and Ranzuglia, G.: MeshLab: An open-source mesh processing tool, in: 6th Eurographics Italian Chapter Conference 2008 - Proceedings, Citation Key: Cignoni2008, 129–136, 2008.
- Crosta, G. B. and Agliardi, F.: Parametric evaluation of 3D dispersion of rockfall trajectories, *Natural Hazards and Earth System Science*, 4, 583–598, <https://doi.org/10.5194/nhess-4-583-2004>, 2004.



- 915 D'Amboise, C. J. L., Teich, M., Hormes, A., Steger, S., and Berger, F.: Modeling Protective Forests for Gravitational Natural Hazards and How It Relates to Risk-Based Decision Support Tools, IntechOpen, <https://doi.org/10.5772/intechopen.99510>, 2021.
- D'Amboise, C. J. L., Neuhauser, M., Teich, M., Huber, A., Kofler, A., Perzl, F., Fromm, R., Kleemayr, K., and Fischer, J.-T.: Flow-Py v1.0: a customizable, open-source simulation tool to estimate runout and intensity of gravitational mass flows, 920 Geoscientific Model Development, 15, 2423–2439, <https://doi.org/10.5194/gmd-15-2423-2022>, 2022.
- Derron, M.-H., Stalsberg, K., and Sletten, K.: NGU Report 2016.003 Method for susceptibility mapping of rock falls in Norway, Technical report, Trondheim, Norway, 2016.
- Domaas, U.: Natural rockfalls - Descriptions and calculations, Norwegian Geotechnical Institute (NGI), Oslo, Norway, 1995.
- 925 Dorren, L. K. A.: Rockyfor3D (v5.2) revealed - Transparent description of the complete 3D rockfall model, Bern, Switzerland, 2015.
- Dorren, L. K. A., Domaas, U., Kronholm, K., and Labiouse, V.: Methods for Predicting Rockfall Trajectories and Run-out Zones, in: Rockfall Engineering, John Wiley & Sons, Inc., Hoboken, NJ, USA, 143–173, <https://doi.org/10.1002/9781118601532.ch5>, 2013.
- 930 Erickson, B.: Personal communication, 2019.
- Garcia, B.: Analyse des mécanismes d'interaction entre un bloc rocheux et un versant de propagation: application à l'ingénierie, Grenoble Alpes, <https://doi.org/HAL Id: tel-02445651>, version 1, 2019.
- Garcia, B., Villard, P., Richefeu, V., and Daudon, D.: Comparison of full-scale rockfall tests with 3D complex-shaped discrete element simulations, Engineering Geology, 310, 106855, <https://doi.org/10.1016/j.enggeo.2022.106855>, 2022.
- 935 Girardeau-Montaut, D.: Détection de changement sur des données géométriques tridimensionnelles, Télécom ParisTech, 207 pp., <https://doi.org/HAL Id: pastel-00001745>, version 1, 2006.
- Hantz, D., Corominas, J., Crosta, G. B., and Jaboyedoff, M.: Definitions and Concepts for Quantitative Rockfall Hazard and Risk Analysis, 11, 158, 2021.
- Hibert, C., Malet, J. P., Bourrier, F., Provost, F., Berger, F., Bornemann, P., and Mermin, E.: Single-block rockfall dynamics 940 inferred from seismic signal analysis, Earth Surface Dynamics, 5, 283–292, <https://doi.org/10.5194/esurf-5-283-2017>, 2017.
- Hibert, C., Noël, F., Toe, D., Talib, M., Desrues, M., Wyser, E., Brenguier, O., Bourrier, F., Toussaint, R., Malet, J.-P., and Jaboyedoff, M.: Machine learning prediction of the mass and the velocity of controlled single-block rockfalls from the seismic waves they generate, Earth Surface Dynamics, 12, 641–656, <https://doi.org/10.5194/esurf-12-641-2024>, 2024.
- Holmgren, P.: Multiple flow direction algorithms for runoff modelling in grid based elevation models: An empirical 945 evaluation, Hydrological Processes, 8, 327–334, <https://doi.org/10.1002/hyp.3360080405>, 1994.



- Horton, P., Jaboyedoff, M., Rudaz, B., and Zimmermann, M.: Flow-R, a model for susceptibility mapping of debris flows and other gravitational hazards at a regional scale, *Natural Hazards and Earth System Sciences*, 13, 869–885, <https://doi.org/10.5194/nhess-13-869-2013>, 2013.
- Luster Stryn 2012: <https://hoydedata.no/LaserInnsyn2/>, last access: 2 October 2022.
- 950 Haram Skodje Ørskog Vestnes 2pkt 2015: <https://hoydedata.no/LaserInnsyn2/>, last access: 7 June 2020.
- NDH Høyanger-Vik-Stølsheimen 2pkt 2016: <https://hoydedata.no/LaserInnsyn2/>, last access: 11 October 2022.
- NVE Flaum Skjolden i Luster 2018: <https://hoydedata.no/LaserInnsyn2/>, last access: 2 October 2022.
- NDH Hemnes øst-Rana sør 2pkt 2019: <https://hoydedata.no/LaserInnsyn2/>, last access: 20 September 2020.
- NDH Årdal 2pkt 2020: <https://hoydedata.no/LaserInnsyn2/>, last access: 2 October 2022.
- 955 Jaboyedoff, M. and Labiouse, V.: Technical Note: Preliminary estimation of rockfall runout zones, *Natural Hazards and Earth System Science*, 11, 819–828, <https://doi.org/10.5194/nhess-11-819-2011>, 2011.
- Jaboyedoff, M., Dudt, J. P., and Labiouse, V.: An attempt to refine rockfall hazard zoning based on the kinetic energy, frequency and fragmentation degree, *Natural Hazards and Earth System Sciences*, 5, 621–632, <https://doi.org/10.5194/nhess-5-621-2005>, 2005.
- 960 Ji, Z.-M., Chen, Z.-J., Niu, Q.-H., Wang, T.-J., Song, H., and Wang, T.-H.: Laboratory study on the influencing factors and their control for the coefficient of restitution during rockfall impacts, *Landslides*, 16, 1939–1963, <https://doi.org/10.1007/s10346-019-01183-x>, 2019.
- Jones, C. L., Higgins, J. D., and Andrew, R. D.: MI-66 Colorado Rockfall Simulation Program, version 4.0, , 127, 2000.
- Keylock, C. and Domaas, U.: Evaluation of Topographic Models of Rockfall Travel Distance for Use in Hazard
- 965 Applications, Arctic, Antarctic, and Alpine Research, 31, 312–320, <https://doi.org/10.1080/15230430.1999.12003314>, 1999.
- Labiouse, V.: Fragmental rockfall paths: comparison of simulations on Alpine sites and experimental investigation of boulder impacts, in: *Landslides: Evaluation and Stabilization/Glissement de Terrain: Evaluation et Stabilisation*, Set of 2 Volumes, CRC Press, Rio de Janeiro, Brazil, 457–466, <https://doi.org/10.1201/b16816-65>, 2004.
- Labiouse, V. and Heidenreich, B.: Half-scale experimental study of rockfall impacts on sandy slopes, *Natural Hazards and*
- 970 *Earth System Sciences*, 9, 1981–1993, <https://doi.org/10.5194/nhess-9-1981-2009>, 2009.
- Lambert, S. and Bourrier, F.: Design of rockfall protection embankments: A review, *Engineering Geology*, 154, 77–88, <https://doi.org/10.1016/j.enggeo.2012.12.012>, 2013.
- Lambert, S., Bourrier, F., and Toe, D.: Improving three-dimensional rockfall trajectory simulation codes for assessing the efficiency of protective embankments, *International Journal of Rock Mechanics and Mining Sciences*, 60, 26–36, <https://doi.org/10.1016/j.ijrmms.2012.12.029>, 2013.
- 975 Lévy, S. and Verly, M.: Personal communication, 2019.
- Li, C. C.: Rock Mechanics – Compendium in TGB 4210, Norwegian University of Science and Technology (NTNU), 2017.
- Loup, B. and Dorren, L. K. A.: Utilisation de modèles pour l'évaluation des dangers de chute de pierres - Fiche d'information, Office fédéral de l'environnement (OFEV), Bern, Switzerland, 2022.



- 980 Lu, G., Caviezel, A., Christen, M., Demmel, S. E., Ringenbach, A., Bühler, Y., Dinneen, C. E., Gerber, W., and Bartelt, P.: Modelling rockfall impact with scarring in compactable soils, *Landslides*, 16, 2353–2367, <https://doi.org/10.1007/s10346-019-01238-z>, 2019.
- Mathon, C., Kobayashi, H., Dewez, T., Sedan, O., Nachbaur, A., Berger, F., Desgarets, E., and Nowak, E.: Calibration of restitution coefficients, based on real size experiment in the weathered volcanic context of Tahiti, French Polynesia, in: Proceedings of the rock slope stability, Citation Key: mathon2010calibration, 2010.
- 985 Midtun, Y.: Tilbakemelding etter synfaring – Steinsprang i Fortun – Luster kommune (201402064-3), The Norwegian Water Resources and Energy Directorate (NVE), Oslo, Norway, 2014.
- Myrvang, A.: Bergmekanikk – Mechanical properties of rocks tested in the rock mechanics laboratory NTH / NTNU / SINTEF, Norwegian University of Science and Technology (NTNU), Trondheim, 2001.
- 990 NGU Bedrock – National bedrock database: https://geo.ngu.no/kart/berggrunn_mobil/, last access: 30 December 2022a.
- NGU Superficial deposits – National Database: https://geo.ngu.no/kart/losmasse_mobil/?lang=eng, last access: 30 December 2022b.
- Norge i Bilder – Ortofoto: <https://norgeibilder.no/>, last access: 30 December 2022.
- Noël, F.: stnParabel, 2020.
- 995 Noël, F. and Nordang, S. F.: Quantitative Rockfall Hazard Assessment of the Norwegian Road Network and Residences at an Indicative Level from Simulated Trajectories, *Remote Sensing*, 17, 819, <https://doi.org/10.3390/rs17050819>, 2025.
- Noël, F., Wyser, E., Jaboyedoff, M., Derron, M.-H., Cloutier, C., Turmel, D., and Locat, J.: Real-size rockfall experiment: How different rockfall simulation impact models perform when confronted with reality?, in: In Proceedings of the 7th Canadian Geohazards Conference 2018, Canmore, Alberta, Canada, 3–6 June 2018, Citation Key: Noel2018, 2018.
- 1000 Noël, F., Cloutier, C., Jaboyedoff, M., and Locat, J.: Impact-Detection Algorithm That Uses Point Clouds as Topographic Inputs for 3D Rockfall Simulations, *Geosciences*, 11, 188, <https://doi.org/10.3390/geosciences11050188>, 2021.
- Noël, F., Jaboyedoff, M., Caviezel, A., Hibert, C., Bourrier, F., and Malet, J.-P.: Rockfall trajectory reconstruction: a flexible method utilizing video footage and high-resolution terrain models, *Earth Surface Dynamics*, 10, 1141–1164, <https://doi.org/10.5194/esurf-10-1141-2022>, 2022.
- 1005 Noël, F., Nordang, S. F., Jaboyedoff, M., Travelletti, J., Matasci, B., Digout, M., Derron, M.-H., Caviezel, A., Hibert, C., Toe, D., Talib, M., Wyser, E., Bourrier, F., Toussaint, R., Malet, J.-P., and Locat, J.: Highly energetic rockfalls: back analysis of the 2015 event from the Mel de la Niva, Switzerland, *Landslides*, 20, 1561–1582, <https://doi.org/10.1007/s10346-023-02054-2>, 2023a.
- Noël, F., Nordang, S. F., Jaboyedoff, M., Digout, M., Guerin, A., Locat, J., and Matasci, B.: Comparing Flow-R, Rockyfor3D and RAMMS to Rockfalls from the Mel de la Niva Mountain: A Benchmarking Exercise, *Geosciences*, 13, 200, <https://doi.org/10.3390/geosciences13070200>, 2023b.
- 1010 Norwegian Building Authority: Veiledning om tekniske krav til byggverk - Byggteknisk forskrift (TEK17) med veiledning, 2017.



- OFDT, OFEG, and OFEFP: Recommandation: Aménagement du territoire et dangers naturels, Rapport final, Confédération
 1015 Suisse, 2005.
- OFEV: Protection contre les dangers dus aux mouvements de terrain. Aide à l'exécution concernant la gestion des dangers
 dus aux glissements de terrain, aux chutes de pierres et aux coulées de boue, Office fédéral de l'environnement, Bern,
 Switzerland, 2016.
- Ondercin, M., Kromer, R., and Hutchinson, D. J.: A Comparison of Rockfall Models Calibrated Using Rockfall Trajectories
 1020 inferred from LiDAR Change Detection and Inspection of Gigapixel Photographs, in: Geohazards 6, Kingston, 7, 2014.
- Oppikofer, T., Michoud, C., Horton, P., and Böhme, M.: Improved rockfall susceptibility map for Norway, in: ISL2024,
 2024.
- Paronuzzi, P.: Rockfall-induced block propagation on a soil slope, northern Italy, *Environmental Geology*, 58, 1451–1466,
<https://doi.org/10.1007/s00254-008-1648-7>, 2009.
- 1025 Pettersen, S.: Notat etter steinskred i Leirskardalen, Nordland County Council, 2020.
- Pfeiffer, T. J. and Bowen, T. D.: Computer Simulation of Rockfalls, *Environmental & Engineering Geoscience*, XXVI, 135–
 146, <https://doi.org/10.2113/gseegeosci.xxvi.1.135>, 1989.
- Pfeiffer, T. J. and Higgins, J. D.: Rockfall hazard analysis using the Colorado rockfall simulation program, *Transportation
 Research Record*, 1288, 117–126, 1990.
- 1030 QGIS Development Team: QGIS Geographic Information System, 2009.
- Raetzo, H., Lateltin, O., Bollinger, D., and Tripet, J. P.: Hazard assessment in Switzerland - Codes of practice for mass
 movements, *Bulletin of Engineering Geology and the Environment*, 61, 263–268, <https://doi.org/10.1007/s10064-002-0163-4>, 2002.
- Rocscience Inc.: RocFall, version 8, 2022.
- 1035 Sarro, R., Rossi, M., Reichenbach, P., and Mateos, R. M.: From rockfall source area identification to susceptibility zonation:
 a proposed workflow tested on El Hierro (Canary Islands, Spain), *Natural Hazards and Earth System Sciences*, 25, 1459–
 1479, <https://doi.org/10.5194/nhess-25-1459-2025>, 2025.
- Spadari, M., Kardani, M., De Carteret, R., Giacomini, A., Buzzi, O., Fityus, S., and Sloan, S. W.: Statistical evaluation of
 rockfall energy ranges for different geological settings of New South Wales, Australia, *Engineering Geology*, 158, 57–65,
 1040 <https://doi.org/10.1016/j.enggeo.2013.03.007>, 2013.
- Distributed by OpenTopography: <https://doi.org/10.5069/G9TH8JNQ>, last access: 26 February 2023.
- Steck, A., Epard, J.-L., Escher, A., Gouffon, Y., and Masson, H.: Carte tectonique des Alpes de Suisse occidentale, Office
 fédéral des eaux et de la géologie (OFEG), Bern, Switzerland, 2001.
- Valagussa, A., Crosta, G. B., Frattini, P., Zenoni, S., and Massey, C.: Rockfall Runout Simulation Fine-Tuning in
 1045 Christchurch, New Zealand, in: *Engineering Geology for Society and Territory - Volume 2*, vol. 2, edited by: Lollino, G.,



- Giordan, D., Crosta, G. B., Corominas, J., Azzam, R., Wasowski, J., and Sciarra, N., Springer International Publishing, Cham, 1913–1917, https://doi.org/10.1007/978-3-319-09057-3_339, 2015.
- Volkwein, A. and Gerber, W.: Repetitive trajectory testing in Tschamut 2014, <https://doi.org/https://www.doi.org/10.16904/envdat.34>, 2018.
- 1050 Volkwein, A., Schellenberg, K., Labiouse, V., Agliardi, F., Berger, F., Bourrier, F., Dorren, L. K. A., Gerber, W., and Jaboyedoff, M.: Rockfall characterisation and structural protection – a review, *Natural Hazards and Earth System Science*, 11, 2617–2651, <https://doi.org/10.5194/nhess-11-2617-2011>, 2011.
- Volkwein, A., Brügger, L., Gees, F., Gerber, W., Krummenacher, B., Kummer, P., Lardon, J., and Sutter, T.: Repetitive Rockfall Trajectory Testing, *Geosciences*, 8, 88, <https://doi.org/10.3390/geosciences8030088>, 2018.
- 1055 Wang, Y., Jiang, W., Cheng, S., Song, P., and Mao, C.: Effects of the impact angle on the coefficient of restitution in rockfall analysis based on a medium-scale laboratory test, *Natural Hazards and Earth System Sciences*, 18, 3045–3061, <https://doi.org/10.5194/nhess-18-3045-2018>, 2018.
- Wu, S.-S.: ROCKFALL EVALUATION BY COMPUTER SIMULATION, *Transportation Research Record*, 1985.
- Wyllie, D. C.: Rock fall engineering: development and calibration of an improved model for analysis of rock fall hazards on highways and railways, The University of British Columbia, 238 pp., <https://doi.org/10.14288/1.0167542>, 2014a
- 1060 Wyllie, D. C.: Calibration of rock fall modeling parameters, *International Journal of Rock Mechanics and Mining Sciences*, 67, 170–180, <https://doi.org/10.1016/j.ijrmms.2013.10.002>, 2014b..
- Ye, Y., Zeng, Y., Thoeni, K., and Giacomini, A.: An Experimental and Theoretical Study of the Normal Coefficient of Restitution for Marble Spheres, *Rock Mechanics and Rock Engineering*, 52, 1705–1722, [https://doi.org/10.1007/s00603-](https://doi.org/10.1007/s00603-018-1709-5)
- 1065 018-1709-5, 2019.