Comput Geosci
DOI 10.1007/s10596-016-9609-9

ORIGINAL PAPER

Evaluation concepts to compare observed and simulated deposition areas of mass movements

M. Heiser[1] · C. Scheidl[1] · R. Kaitna[1]

Received: 26 January 2016 / Accepted: 21 December 2016

Abstract The simulation of geophysical mass flows, including debris flows, rock and snow avalanches, has become an important tool in engineering hazard assessment. Especially the runout and deposition behaviour of observed and expected mass flows are of interest. When being confronted with the evaluation of model performance and sensitivity, there are no standard, objective approaches. In this contribution, we review methods that have been used in literature and outline a new approach to quantitatively compare 2D simulations of observed and simulated deposition pattern. Our proposed method is based on the comparison of normalized partial areas which can be plotted in a ternary diagram to visualize the degree of over- and under-estimation. Results can be summed up by a single metric between -1 (no fit) and 1 (perfect fit). This study shall help developers and end-users of simulation models to better understand model behaviour and provides a possibility for comparison of model results, independent of simulation platform and type of mass flow.

Keywords Mass movements · Evaluation · Simulation · Lee-Salle index

Electronic supplementary material The online version of this article (doi:10.1007/s10596-016-9609-9) contains supplementary material, which is available to authorized users.

✉ M. Heiser
micha.heiser@boku.ac.at

[1] Institute of Mountain Risk Engineering, University of Natural Resources and Life Sciences, Vienna, Peter-Jordan Street 82, A-1190, Vienna, Austria

1 Introduction

In mountainous regions with a high density of population, people meet a challenge to find an accurate balance between the imminence of natural hazards and the progress in spatial developments. Thus, a detailed delineation of potentially endangered areas becomes increasingly important for regional and urban planning and various countries published guidelines or recommendations. For example, guidelines of Austria[1], Southtyrol[2], Slovenia[3] and Switzerland[4] require process-related information about, e.g. the pressure, flow depth or flow velocity or deposition depth for different return periods to delineate areas into different hazard zones. This information is normally not available from in situ measurements, and one can only be inferred with the use of simulation models.

In contrast to a forward simulation where constitutive parameters θ are known and the outcome $\hat{\omega}$ is of interest ($\theta \longrightarrow$ Model $\longrightarrow \hat{\omega}$), most event simulations of mass movement processes are conducted as inverse problems where the reference ω but not the model parameters $\hat{\theta}$ are known ($\hat{\theta} \longleftarrow$ Model $\longleftarrow \omega$). The objective of inverse modelling is to find model parameters $\hat{\theta}$ which are able to reproduce the known results or reference of the system under consideration. Back calculation, sometimes also denoted as back analysis, calibration or reverse modelling is essentially the same. By calibrating the model parameter(s) $\hat{\theta}$, the event which is described by reference variables like

[1] BMLFUW-LE.3.3.3/0185-IV/5/2007, 4th February, 2011
[2] Supplemento n. 2 al B.U. n. 35/I-II, 26th August,2008
[3] e.g. Article 83,Waters Act ZV-1, 12th July, 2002
[4] 721.100 Federal law on Hydraulic, 1st January, 2011 and 921.0 Federal law on Forest, 1st Juli, 2013

Published online: 19 January 2017

run out distance, deposition distribution or flow velocity, should be reproduced so that the model outcome $\hat{\omega}$ is close to the reference ω. This approach allows to simulate outcome variables like pressure, velocity or flow depth which were not observed or measured.

This contribution focuses on the evaluation of 2D simulation of gravitational mass movements like debris avalanches, debris flows, lahars, rock(/-ice) avalanches, rockslides, and snow avalanches. The prerequisite conditions for all these geophysical mass flows include an abundant source of material (snow, ice, rock or loose debris), steep slopes and a trigger.

Owned to the complexity of such gravitational mass movements, existing simulation models may be divided into empirical-statistical and analytical-deterministic methods. In recent years, also an increasing number of simulation models based on cellular modelling of mass movement processes was published [e.g. 3, 10, 13, 27, 49, 50].

Empirical-statistical approaches are easy to use but rely on empirical parameters (e.g. a mobility coefficient), which have to be determined by the user, mostly out of a range of "typical" parameters. The use of such models is restricted to conditions similar to those on which their development is based. 2 D empirical models like LAHARZ [29], DFLOWZ [5], TopRunDF [49] or TopFLowDF [50] provide a quick view on potential endangered areas and have been applied for several hazard analyses [17, 21, 40, 45, 54].

More complex simulation models in engineering practice are based on an "equivalent fluid" approach [25], routing such mass flows in one or two dimensions along the talweg or over a digital elevation model solving depth averaged flow equations [11, 16, 26, 35, 36, 39]. Such models, like DAN3D [25], Flo2D [39], RAMMS [11] or SAMOS [47], are physically based and consider the momentum or energy conservation of the flow. A major difficulty of models with a more physical background is the choice of appropriate flow resistance parameters or material rheologies [2, 25, 28] as well as finding an appropriate stopping criteria. Nevertheless, physical models have been applied worldwide for hazard assessments for different kind of mass movements, whereas simulation results respectively evaluation concepts are often based on subjective or expert based rules. A systematic review of 75 peer reviewed articles[5] from 1995 to 2015 which inverse modelled real events of different types of mass flows revealed that there are no standard approaches for testing the accuracy of observed and simulated deposition patterns of mass movement simulation models, unlike in statistical, computational or meteorological modelling where a lot of different error measures exist [1, 4, 14].

[5] A table that lists all articles as well as meta-information used in this study can be found in the supplementary material.

For this reason, an objective evaluation tool is needed to allow comparison amongst different simulation results. We will therefore first identify variables that are typically of interest in simulating geophysical mass flows and subsequently discuss existing evaluation concepts, focusing on the spatial distribution of deposits (i.e. model result). Finally, we suggest a new evaluation concept that allows an objective comparison and visualisation of model performance across platforms and independent of type of process, which is illustrated by two examples.

2 Reference variables of interest

A lot of different reference variables are used to describe mass movements (see Fig. 1). In an ideal situation, the modeler can evaluate or calibrate the model on a wide range of reference variables that describe the flow and deposition behaviour of the process (e.g. 1-D study of [37]). However, in most cases, these combined observations are not available. Nearly 80 % of the studies used the deposition distribution (d_D) as a reference variable for evaluation. We use the term deposition distribution to emphasize that the deposition area can not only be presented by a scalar value but also by the shape and location of the deposited mass. In 25 % of the studies, the descriptive parameter is runout distance (r), which is the length measured between two points, where one is representing the start and the other the end of the mass movement. These points could be for example the centre of gravity or the furthest point of the detachment and deposition zone. The deposition depth (d_z), which is the height of deposition expressed as the mean or maximum height of the observed deposition, was used in 40 % of the studies. The deposition volume (d_V) and the flow velocity (f_v) where only used by nearly 25 % of the studies, where f_v can be measured at different locations of the process (front, distal) an expressed as mean or maximum. Finally, mean velocity ($\bar{v}$), total distance travelled divided by the duration of the movement and flow depth (f_z), which again can be expressed as minimum, maximum or mean, where at least often used. Besides those reference variables, some more specific variables were applied in some studies. The Heim friction coefficient [23] was used by [42, 43]. Total discharge and the front trajectory–position of the debris flow front at a certain time was used by [37]. Different geometric measure of the reference and simulated deposition distribution was used by [42], including deposition length and length to width ratio.

Most studies show qualitative measures like visual comparison and verbal descriptions, e.g. [9, 18, 38]. Fewer studies applied quantitative measures for the comparison between reference and simulation, e.g. [30, 41, 42]; in almost 80 % of the reviewed studies, the deposition area was

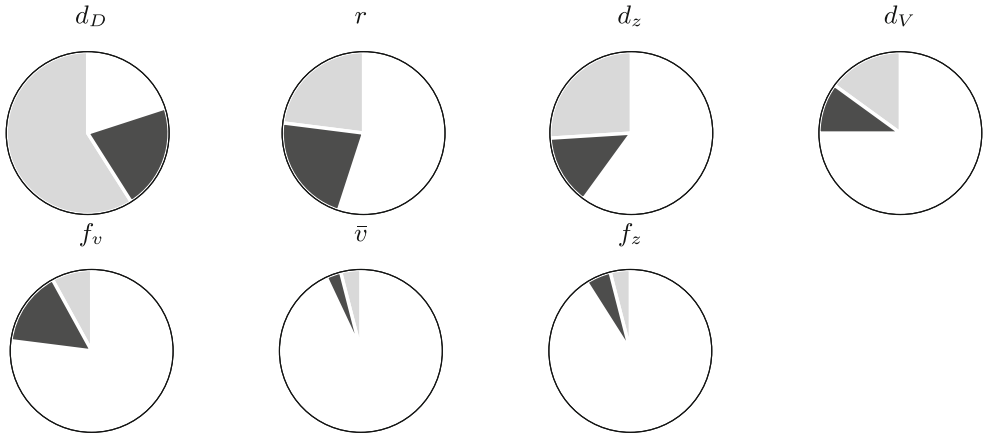


Fig. 1 *Grey areas* indicate a qualitative evaluation approach, *black areas* indicate a quantitative evaluation approach expressed as a ratio of all studies under consideration (see text for explanation of the reference variables). The most striking feature of this figure is the great amount of studies where the deposition distribution was available but only 26 % of those where evaluating it quantitatively. It seems therefore convenient to base an evaluation concept on d_D

observed and available, but only 26 % of those studies used them as a quantitative measure for the error of the simulation. Figure 1 also shows that half of the studies focusing on runout distance provided a quantitative comparison between simulation and observation.

For model comparison and model development (e.g. testing whether higher complexity of simulation models leads to higher accuracies of their outcomes), we argue that objective standardized approaches—as available in the field of statistical science or meteorology [1, 14]—are needed for modelling geophysical mass flows. Since engineering hazard assessment needs spatial information of areas affected by mass movement processes and as we find for most studies the deposition distribution was available, but rarely used for comparison, it seems convenient to base an evaluation concept on d_D.

3 Evaluation concepts for the comparison of deposition area

The quantitative assessment of the similarity between the observed deposition area (or reference area) and the simulated deposition area (Fig. 2) needs the definition of several partial areas (1a)–(1c).

$$X = A \cap \hat{A} \tag{1a}$$

$$U = A \setminus \hat{A} \tag{1b}$$

$$O = \hat{A} \setminus A \tag{1c}$$

Here, X defines the area overlapped by the simulation outcome and the reference, U refers to the magnitude of the underestimated area and O to the magnitude of the overestimated area.

The simplest approach to evaluate simulation results is by estimating the error ϵ_d as the difference between the reference area and the simulated area ($\epsilon_d = A - \hat{A}$), applied for instance by [19]. For a first-order evaluation, this approach gives plausible results, but lacks the capability to detect the shape similarity of the two areas. Similar arguments apply also for the evaluation concept defining ϵ_q as the quotient between simulated area and the reference area ($\epsilon_q = \hat{A}/A$), used by e.g. [15]. For both approaches, an error value of $\epsilon_d = 0$, respectively $\epsilon_q = 1$ does not only refer to a perfect fit but can theoretically also imply no overlapping for areas with equal magnitudes. To avoid this problem [19, 24, 43, 53] used the amount of overlap between reference and simulation as a measure of accuracy. Their error-approach can be defined as $\epsilon_q = X/A$. If ϵ_q—i.e. the error—is solely defined by quotient of overlap and reference area, a perfect fit might also be indicated as long as the simulation is totally overlapping the reference, regardless of the amount of overestimation. For this reason, [19] also used $\epsilon_q = U/A$ and $\epsilon_q = O/A$ to evaluate their simulations.[6] However, they did not combine those three error measures into one number. The combination of the error sources is as important as the decomposition of the error because an ideal evaluation concept should not only offer the possibility to compare different models with respect to an non-flawed error measure but should also give information about the dominating error source (overestimation or

[6] Note that this three quotient are identical to the definition of α, β and γ in Eq. 4.

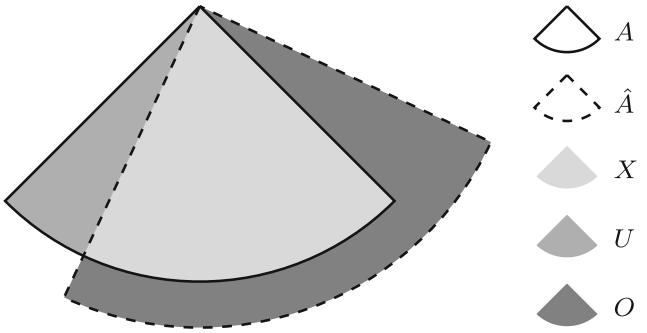


Fig. 2 Schematic figure of a mass movement deposition distribution with superimposed simulation. Here, A refers to the observed deposition area (reference), whereas $\hat{A}$ denotes the total simulated area.The magnitude of the overestimated area is defined by O, whereas U refers to the size of the underestimated area. Finally, X defines the area overlapped by the simulation outcome and the reference

underestimation). [34] introduced the evaluation index I_a which, to a certain amount, satisfy those criteria.

$$I_a = \left(\frac{X}{A} + \frac{X}{\hat{A}}\right) \cdot 2^{-1} \tag{2}$$

Equation 2 can been seen as a simple additive weighting of under- (X/A) and overestimation $(X/\hat{A})$. A perfect fit would yield a value of 1 for both fractions. While underestimation dominates the first fraction, overestimation affects solely the second fraction in Eq. 2. Therefore, the approach of [34] gives information about the amount of over- respectively underestimation caused to the simulation outcome. However, no information about the goodness of overlapping (e.g. the error that a simulation hits the reference) can be given in such a case.

An evaluation concept that is also taking the shape of the deposition distribution into account, has much similarity with classification problems. The objective is to identify those points, in the model domain, which are affected by the process and distinguish them from those which are not affected, i.e. classifying the pixels into affected and not affected ones. The classification can be made with the aid of an indicator function, like that given by Eq. 3.[7]

$$I(\omega) = \begin{cases} 1 \text{ if } \omega \geq \omega_{th} \\ 0 \text{ if } \omega < \omega_{th} \end{cases} \tag{3a}$$

$$I(\hat{\omega}) = \begin{cases} 1 \text{ if } \hat{\omega} \geq \hat{\omega}_{th} \\ 0 \text{ if } \hat{\omega} < \hat{\omega}_{th} \end{cases} \tag{3b}$$

ω_{th} threshold for reference variable, e.g. deposition depth
$\hat{\omega}_{th}$ threshold for simulation variable, e.g. pressure, percentage of momentum

[7]The choice of ω_{th} and $\hat{\omega}_{th}$ is a decision made by the user which is based on his respective needs and should be well documented.

The result is a binary raster of A respectively $\hat{A}$ where all pixels are 1 for values greater than the pre-defined thresholds ω_{th} and $\hat{\omega}_{th}$. The advantage of using a binary raster is that it allows to distinguish between over- and underestimation due to a pixel by pixel comparison. A disadvantage, however, lies in the high number of true negatives, i.e. pixels that are neither affected by the model nor the observation. This might mask the true performance of the model.
Nevertheless, [8] used the idea of deposition distribution as a classification problem which was then adopted to account for the disadvantages stated earlier and denoted with Ω by [49] . The calculation of the Ω-value is shown in Eq. 4. Here, α indicates the fit between simulation and reference. The β value is a measurement for the first source of error, i.e. the underestimation, and finally γ is a measure for the second source of error, i.e. the overestimation. The combination of all three values defines Ω, a value accounting for the simulation accuracy with respect to the deposition distribution.

$$\Omega = \alpha - \beta - \gamma \tag{4}$$

α overlap calculated as X/A
β underestimation calculated as U/A
γ overestimation calculated as O/A

A perfect simulation would totally fit the observed deposition area and is defined with $\hat{A} = A$ and $O = U = 0$. For such a case, $X = \hat{A} = A$ which yields to $\alpha = 1$ respectively $\beta = \gamma = 0$ and finally, following Eq. 4, results in the maximum possible omega value, $\Omega_{\max} = 1$. For the extreme case of no underestimation, i.e. $\beta = 0$, α will get 1, but the value of γ will depend on the amount of overestimation (exceeding 1). In this case, only the size of the overestimated area, relative to the reference area, is determining the value of Ω and so also its lower boundary. This is different to [49] which defined a lower boundary of $\Omega_{\min} = -1$. However, to calculate the minimum of the Ω value, we hypothesise a simulation that totally misses the reference so that $O = \hat{A}$ and $X = 0$. Then the minimum omega value can be calculated by:

$$\Omega_{\min} = -1 - \hat{A}/A \tag{5}$$

Equation 5 refers to the lowest possible value of omega for the simulation under consideration. For this reason, it is clear that $\Omega_{\min}$ strongly depends on the value of γ which in turn depends on the total area of the simulation. Based on this approach, no unbiased comparison between simulations can be made.
A possible solution to compare simulations over different model domains by using Ω is to rescale it based on the difference of its minimum and maximum value. Because the maximum value is defined with $\Omega_{\max} = 1$, the rescaled

omega value (Ω') solely depends on $\Omega_{\min}$ (Eq. 5) as shown in Eq. 6.

$$\Omega' = \frac{\Omega - \Omega_{\min}}{1 - \Omega_{\min}} \tag{6}$$

In contrast to the first discussed error measures, ϵ_d and ϵ_q, the Ω value combines the information about the overlap and the two sources of errors into one number. Because Ω depends on the amount of overestimation a comparison between different simulation outcomes is still impossible. For this reason, we proposed Ω', which is normalized to the particular overestimation, showing a fix range between [0, 1].

4 Proposed evaluation concept

An evaluation concept not depending on the amount of overestimation can be based on the intersection and union of the reference and the simulation. This approach is known as Lee-Salle index in the urban growth modelling community [12, 33]. In Eq. 7, we denote the Lee-Salle index α_T to indicate the similarity to α of Eq. 4. Here, α_T is also a measure of the amount of overlap but with respect to the total area T, union of reference (A) and simulated ($\hat{A}$) area. Similarly, we can define β_T and γ_T which yields to Ω_T.

$$\Omega_T = \alpha_T - \beta_T - \gamma_T \tag{7}$$

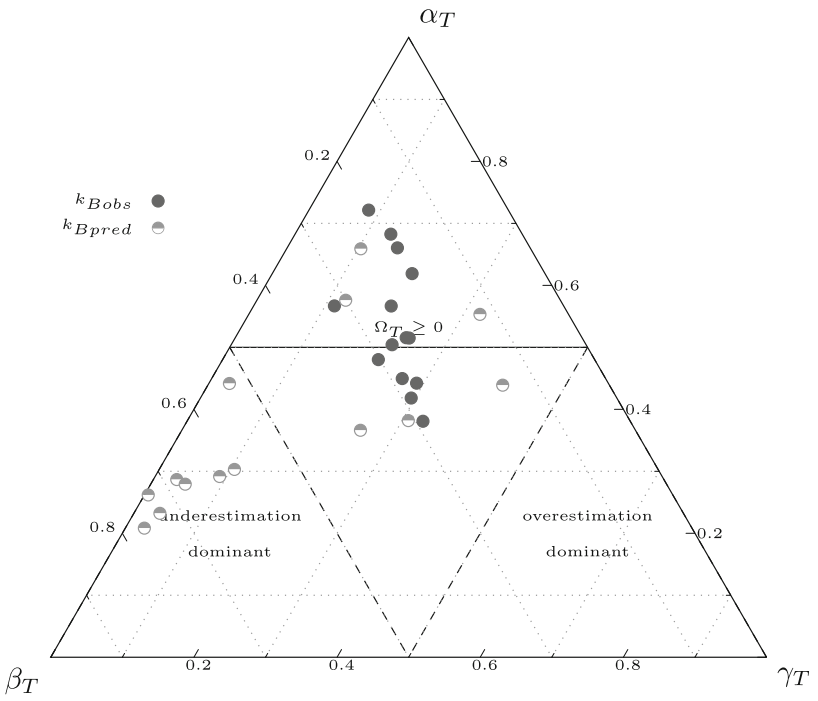
α_T overlap calculated as X/T
β_T underestimation calculated as U/T
γ_T overestimation calculated as O/T

The evaluation concept based on Eq. 7 has the advantage that the possible range of Ω_T is fixed between 1 and -1, indicating a perfect fit or total misfit, regardless of the model domain. A further benefit of this evaluation concept is the possibility to compare different simulation results in a ternary plot (Fig. 3b).To illustrate the evaluation concept, described in this study, two different cases on how the evaluation concept can be applied are discussed.[8]

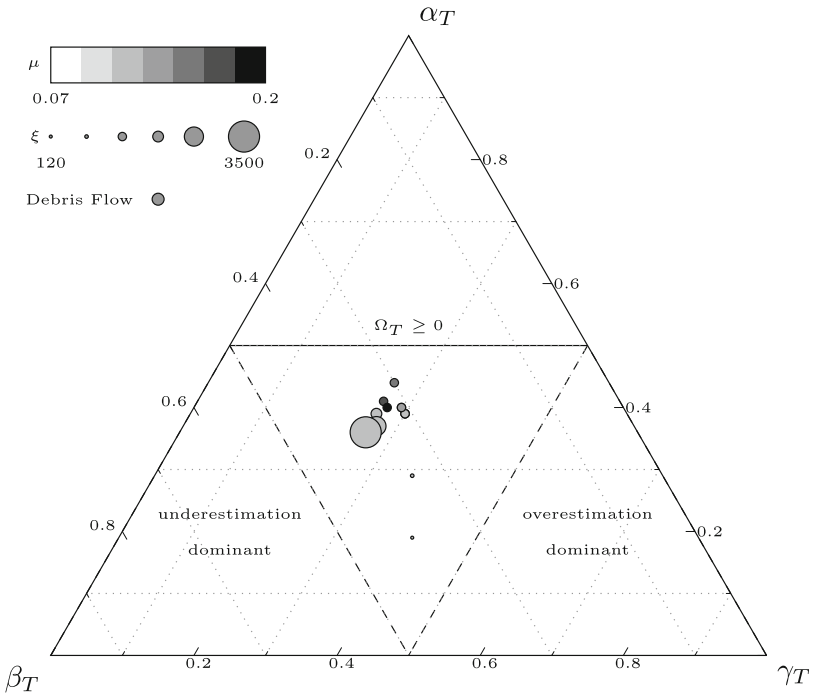
[49] developed an semi-empirical model to simulate the runout of debris-flow events. Their model, TopRunDF, is based on a multiple flow routine and the assumption of geometric similarity between deposited area and volume. TopRunDF can be used to predict inundated areas of debris-flow events on the fan based on a so called mobility coefficient k_B. For best-fit simulations, the mobility coefficent is defined by the observed deposition area as well as the observed deposition volume—denoted by k_{Bobs}. For forward modelling, [49] introduced a new empirical relation to determine the mobility coefficient (k_{Bpred}) as a function of geomorphologic catchment parameters. Figure 3a shows estimated Ω_T values for the simulations of different events. The filled circles refer to the best fit simulations (based on k_{Bobs}) whereas the half-filled circles show the Ω_T values of the predicted simulations of the same events (based on k_{Bpred}).
Although TopRunDF includes only one variable, the evaluation values of the best fit simulations scatter over a wide range. However, no systematic over- or underestimation can be detected. The scatter of the evaluation values of the predicted simulations is even wider but this seems plausible since [46] reported an uncertainty of a factor two, comparing k_{Bpred} with k_{Bobs}. Nevertheless, the evaluation values show that simulations based on k_{Bpred} tend to underestimate the related observations.
Both cases of this example reveal that the proposed evaluation concept can be used to support the identification of potential weaknesses and strengths of the considered simulation tool.

Another example is based on a parameter study from [52]. They used, amongst other, RAMMS-DF (developed at the WSL Institute for Snow and Avalanche Research SLF and the Swiss Federal Institute for Forest, Snow and Landscape Research WSL) to conduct debris-flow runout simulations. Here, the objective was to find those set of basal and internal friction parameters $\theta^* = [\mu^*, \xi^*]$ where the error for the deposition distribution d_D gets a minimum; respectively, the evaluation benchmark Ω_T a maximum. As a first step, [52] performed six simulations with an initially fixed parameter μ, varying only ξ from 120 to 1300. Based on these simulations, they kept the best fit parameter $\xi^* = 300$ constant and conducted seven more simulations with μ ranging from to 0.20. We transformed the simulated deposition patterns of all simulations into binary rasters, using deposition depth as threshold leading to $\hat{\omega}_{th} = 10\,\text{cm}$ (see Eq. 3b). All calculated Ω_T values are drawn in Fig. 3b, showing that a maximum evaluation value of $\Omega_T = -0.13$ was obtained with $\xi^* = 300\,\text{m s}^{-2}$, $\mu^* = 0.13$.
Sometimes, simulations of one event might yield to similar results although different constitutional parameters have been chosen. In other words, if the differences between the Ω_T values are very small, the selection of the best fit simulation will be difficult. In such a case, the proposed evaluation model in combination with the ternary plot might act as a decision support tool, providing the possibility to chose whether over- or underestimation should dominate the simulation error.

[8] Note that the examples are chosen arbitrarily and shall not judge whether one simulation model is superior to the other.



(a) Debris flow simulated with TopRunDF (data from [49]).



(b) Debris flow and snow avalanche simulated with RAMMS (data for debris flow from [52]).

Fig. 3 In **a**, the evaluation concept is used to detect systematic errors—over- or underestimation—in the simulation model, while in **b**, the concept is used for finding those parameters θ^* of the simulation model which are best reproducing the reference (calibration)

5 Conclusion and outlook

Several studies apply simulation models with different complexity approaches for gravitational mass movements. Many of these studies provide somehow information on the plausibility of their simulation outcomes. The review conducted in this article shows that the deposition distribution acts as the most often used reference variable for evaluation. However, existing evaluation concepts with respect to the deposition distribution does only account for

one or a combination of two possible evaluation errors due to overestimation, underestimation and/or overlapping of the simulation outcome and the observed reference. The proposed evaluation concept integrates all three possible errors and, combined with a ternary plot, might act as a simple decision support tool to (i) identify weaknesses and strengths of the simulation model, (ii) to find the best fit simulation setup and (iii) to test whether higher complexity of simulation models are balanced by higher accuracies.

For future applications, the proposed concept could also be used to compare reference and simulated deposition volumes with:

$$\Omega_V = \alpha_V - \beta_V - \gamma_V \tag{8}$$

α_V overlap calculated as V_X / V_T
β_T underestimation calculated as V_U / V_T
γ_T overestimation calculated as V_O / V_T

In Eq. 8, V_T and V_X are the maximal and minimal deposition height integrated over T. V_U and V_O are defined as the absolute difference between reference d_z and simulated deposition height $\hat{d}_z$ integrated over T, for $d_z > \hat{d}_z$ respectively $d_z < \hat{d}_z$. For the discrete case of two raster data sets with n rows and m columns, the integral reduces to a summation over all rows and columns times the pixel area p_A (Eqs. 9–12).

$$V_T = p_A \cdot \sum_{i=1}^{n} \sum_{j=1}^{m} \max\left(d_{z_{i,j}}, \hat{d}_{z_{i,j}}\right) \tag{9}$$

$$V_X = p_A \cdot \sum_{i=1}^{n} \sum_{j=1}^{m} \min\left(d_{z_{i,j}}, \hat{d}_{z_{i,j}}\right) \tag{10}$$

$$V_U = p_A \cdot \sum_{i=1}^{n} \sum_{j=1}^{m} \left| d_{z_{i,j}} - \hat{d}_{z_{i,j}} \right| \text{ for } d_{z_{i,j}} > \hat{d}_{z_{i,j}} \tag{11}$$

$$V_O = p_A \cdot \sum_{i=1}^{n} \sum_{j=1}^{m} \left| d_{z_{i,j}} - \hat{d}_{z_{i,j}} \right| \text{ for } d_{z_{i,j}} < \hat{d}_{z_{i,j}} \tag{12}$$

Although the accurate determination of deposition volumes for evaluation purposes is related to high financial and technical efforts, it might become more popular because of the recent technical advances, which may enhance the availability of dense spatial data of deposition heights in the future. Airborne LiDAR was used by [6, 32, 51] to estimate deposition volumes of debris flows. [31, 44] used terrestrial laser scanning (TLS) to determinate the dynamics of a slow moving landslide and debris flows for subsequent model simulations. Some studies also used a combination of both methods [7, 20, 22]. The relatively new method based on structure from motion as discussed in [55] will further increase the availability of 3-D spatial data, as shown in an UAV study by [48].

It has to be noted that the proposed evaluation concept and all evaluation techniques reviewed for this study are often not strongly influenced by any physical based stopping criteria but are rather based on heuristic stopping assumptions. This could be seen as a fundamental limit for an evaluation of the applied model, which is out of the scope of this study. The intension of the proposed concept is to provide an objective evaluation tool between (somehow) arbitrary observed and simulated deposition areas—whereas the latter is based on the state-of-the-art of the applied model. For a true improvement of model performance and predictive power, more research on the stopping criteria for mass movements is essential.

Acknowledgments Open access funding provided by University of Natural Resources and Life Sciences Vienna (BOKU).

References

1. Armstrong, J.S.: Principles of forecasting: a handbook for researchers and practitioners, vol. 30 Springer Science & Business Media (2001)
2. Barbolini, M., Gruber, U., Keylock, C., Naaim, M., Savi, F.: Application of statistical and hydraulic-continuum dense-snow avalanche models to five real european sites. Cold Reg. Sci. Technol. **31**(2), 133–149 (2000). doi:10.1016/S0165-232X(00)00008-2
3. Barpi, F., Borri-Brunetto, M., Veneri, L.D.: Cellular-automata model for dense-snow avalanches. J. Cold Reg. Eng. **21**(4), 121–140 (2007)
4. Bayarri, M.J., Berger, J.O., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J., Lin, C.H., Tu, J.: A framework for validation of computer models Technometrics (2012)
5. Berti, M., Simoni, A.: Dflowz: a free program to evaluate the area potentially inundated by a debris flow. Comput. Geosci. **67**, 14–23 (2014)
6. Bossi, G., Cavalli, M., Crema, S., Frigerio, S., Luna, B.Q., Mantovani, M., Marcato, G., Schenato, L., Pasuto, A.: Multi-temporal LiDAR-DTMs as a tool for modelling a complex landslide: a case study in the rotolon catchment (eastern italian alps). Nat. Hazards Earth Syst. Sci. **15**(4), 715–722 (2015). doi:10.5194/nhess-15-715-2015
7. Bremer, M., Sass, O.: Combining airborne and terrestrial laser scanning for quantifying erosion and deposition by a debris flow event. Geomorphology **138**(1), 49–60 (2012). doi:10.1016/j.geomorph.2011.08.024

8. Carranza, E.J.M., Castro, O.T.: Predicting lahar-inundation zones: case study in west mount pinatubo, Philippines. Nat. Hazards **37**(3), 331–372 (2006)
9. Ćetina, M., Rajar, R., Hojnik, T., Zakrajšek, M., Krzyk, M., Mikoš, M.: Case study: numerical simulations of debris flow below stože, Slovenia. J. Hydraul. Eng. **132**(2), 121–130 (2006)
10. Chiari, M., Scheidl, C.: Application of a new cellular model for bedload transporting extreme events at steep slopes. Geomorphology **246**, 413–419 (2015)
11. Christen, M., Kowalski, J., Bartelt, P.: Ramms: Numerical simulation of dense snow avalanches in three-dimensional terrain. Cold Reg. Sci. Technol. **63**(1), 1–14 (2010)
12. Clarke, K.C., Hoppen, S., Gaydos, L.: Methods and Techniques for Rigorous Calibration of a Cellular Automaton Model of Urban Growth. In: Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico. Citeseer (1996)
13. D'Ambrosio, D., Spataro, W., Iovine, G.: Parallel genetic algorithms for optimising cellular automata models of natural complex phenomena: an application to debris flows. Comput. Geosci. **32**(7), 861–875 (2006). doi:10.1016/j.cageo.2005.10.027
14. Ebert, E.E.: Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. Meteorol. Appl. **15**(1), 51–64 (2008)
15. Evans, S., Guthrie, R., Roberts, N., Bishop, N.: The disastrous 17 february 2006 rockslide-debris avalanche on leyte island, Philippines: a catastrophic landslide in tropical mountain terrain. Nat. Hazards Earth Syst. Sci. **7**(1), 89–101 (2007)
16. Fraccarollo, L., Papa, M.: Numerical simulation of real debris-flow events. Phys. Chem. Earth Part B **25**(9), 757–763 (2000)
17. Fidej, G., Mikoš, M., Rugani, T., Jež, J., Kumelj, Š., Diaci, J.: Assessment of the protective function of forests against debris flows in a gorge of the slovenian alps iForest - Biogeosciences and Forestry (1), 73–81. doi:10.3832/ifor0994-007 (2015)
18. Granig, M., Sampl, P., Tollinger, C., Jörg, P.: Experiences in Avalanche Assessment with the Powder Snow Avalanche Model Samosat. In: International Snow Science Workshop Proceedings, Davos, Switzerland, pp. 514–518 (2009)
19. Gruber, U., Bartelt, P.: Snow avalanche hazard modelling of large areas using shallow water numerical methods and gis. Environ. Model Softw. **22**(10), 1472–1481 (2007)
20. Haas, F., Heckmann, T., Wichmann, V., Becht, M.: Runout analysis of a large rockfall in the dolomites/italian alps using LIDAR derived particle sizes and shapes. Earth Surf. Process. Landforms **37**(13), 1444–1455 (2012). doi:10.1002/esp.3295
21. Han, Z., Chen, G., Li, Y., Zhang, H., He, Y.: Elementary analysis on the bed-sediment entrainment by debris flow and its application using the topflowdf model. Geomat. Nat. Haz. Risk **7**(2), 764–785 (2016)
22. Heckmann, T., Bimbse, M., Krautblatter, M., Haas, F., Becht, M., Morche, D.: From geotechnical analysis to quantification and modelling using LiDAR data: a study on rockfall in the reintal catchment, bavarian alps, Germany. Earth Surf. Process. Landforms **37**(1), 119–133 (2011). doi:10.1002/esp.2250
23. Heim, A.: Bergsturz Und Menschenleben. Fretz und Wasmuth, Zürich (1932)
24. Hsu, S., Chiou, L., Lin, G., Chao, C., Wen, H., Ku, C.: Applications of simulation technique on debris-flow hazard zone delineation: a case study in hualien county, Taiwan. Nat. Hazards Earth Syst. Sci. **10**(3), 535–545 (2010)
25. Hungr, O.: A model for the runout analysis of rapid flow slides, debris flows, and avalanches. Can. Geotech. J. **32**(4), 610–623 (1995)
26. Hungr, O., McDougall, S.: Two numerical models for landslide dynamic analysis. Comput. Geosci. **35**(5), 978–992 (2009)
27. Iovine, G., D'Ambrosio, D., Gregorio, S.D.: Applying genetic algorithms for calibrating a hexagonal cellular automata model for the simulation of debris flows characterised by strong inertial effects. Geomorphology **66**, 287–303 (2005). doi:10.1016/j.geomorph.2004.09.017
28. Iverson, R.M.: The physics of debris flows. Rev. Geophys. **35**(3), 245–296 (1997)
29. Iverson, R.M., Schilling, S.P., Vallance, J.W.: Objective delineation of lahar-inundation hazard zones. Geol. Soc. Am. Bull. **110**(8), 972–984 (1998)
30. Jakob, M., Anderson, D., Fuller, T., Hungr, O., Ayotte, D.: An unusually large debris flow at hummingbird creek, mara lake, british columbia. Can. Geotech. J. **37**(5), 1109–1125 (2000)
31. Jun, B.H.: Numerical simulation of the topographical change in korea mountain area by intense rainfall and consequential debris flow. Adv. Meteorol. **2016**, 1–11 (2016). doi:10.1155/2016/9363675
32. Kim, H., Lee, S.W., Yune, C.Y., Kim, G.: Volume estimation of small scale debris flows based on observations of topographic changes using airborne LiDAR DEMs. J. Mt. Sci. **11**(3), 578–591 (2014). doi:10.1007/s11629-013-2829-8
33. Lee, D.R., Sallee, G.T.: Theoretical patterns of farm shape and central place location. J. Reg. Sci. **14**(3), 423–430 (1974)
34. Lin, J.Y., Yang, M.D., Lin, B.R., Lin, P.S.: Risk assessment of debris flows in songhe stream, Taiwan. Eng. Geol. **123**(1), 100–112 (2011)
35. Medina, V., Hrlimann, M., Bateman, A.: Application of flatmodel, a 2d finite volume code, to debris flows in the northeastern part of the iberian peninsula. Landslides **5**, 127–142 (2008)
36. Mergili, M., Schratz, K., Ostermann, A., Fellin, W.: Physically-based modelling of granular flows with open source gis. Nat. Hazards Earth Syst. Sci. **12**(1), 187–200 (2012). doi:10.5194/nhess-12-187-2012
37. Naef, D., Rickenmann, D., Rutschmann, P., McArdell, B.: Comparison of flow resistance relations for debris flows using a one-dimensional finite element simulation model. Nat. Hazards Earth Syst. Sci. **6**(1), 155–165 (2006)
38. Nagelisen, J., Moore, J.R., Vockenhuber, C., Ivy-Ochs, S.: Post-glacial rock avalanches in the obersee valley, glarner alps, Switzerland. Geomorphology **238**, 94–111 (2015)
39. O'Brien, J.S., Julien, P.Y., Fullerton, W.: Two-dimensional water flood and mudflood simulation. J. Hydraul. Eng. **119**, 244–260 (1993)
40. Oramas Dorta, D., Toyos, G., Oppenheimer, C., Pareschi, M., Zanchetta, G.: Empirical modelling of the may 1998 small debris flows in sarno (italy) using laharz. Nat. Hazards **40**, 381–396 (2007). doi:10.1007/s11069-006-0035-5
41. Pirulli, M.: On the use of the calibration-based approach for debris-flow forward-analyses. Nat. Hazards Earth Syst. Sci. **10**, 1009–1019 (2010)
42. Preuth, T., Bartelt, P., Korup, O., McArdell, B.: A random kinetic energy model for rock avalanches: eight case studies. J. Geophys. Res. Earth Surf. (2003–2012) **115**(F3), 1–22 (2010)
43. Procter, J.N., Cronin, S.J., Platz, T., Patra, A., Dalbey, K., Sheridan, M., Neall, V.: Mapping block-and-ash flow hazards based on titan 2d simulations: a case study from mt. taranaki, nz. Nat. Hazards **53**(3), 483–501 (2010)
44. Prokop, A., Panholzer, H.: Assessing the capability of terrestrial laser scanning for monitoring slow moving landslides. Nat. Hazards Earth Syst. Sci. **9**(6), 1921–1928 (2009). doi:10.5194/nhess-9-1921-2009
45. Reid, M.E., Coe, J.A., Brien, D.L.: Forecasting inundation from debris flows that grow volumetrically during travel, with

application to the Oregon Coast Range, USA. Geomorphology (2016). doi:10.1016/j.geomorph.2016.07.039
46. Rickenmann, D., Scheidl, C.: Modelle zur Abschätzung des Ablagerungsverhaltens von Murgängen. Wasser Energie Luft **1**, 17–26 (2010)
47. Sampl, P., Zwinger, T.: Avalanche simulation with samos. Ann. Glaciol. **38**(1), 393–398 (2004)
48. Scaioni, M., Longoni, L., Melillo, V., Papini, M.: Remote sensing for landslide investigations: an overview of recent achievements and perspectives. Remote Sens. **6**(10), 9600–9652 (2014). doi:10.3390/rs6109600
49. Scheidl, C., Rickenmann, D.: Empirical prediction of debris-flow mobility and deposition on fans. Earth Surf. Process. Landf. **35**(2), 157–173 (2010)
50. Scheidl, C., Rickenmann, D.: TopflowDF - a Simple Gis Based Model to Simulate Debris-Flow Runout on the Fan. In: Genevois, D.P.A., Hamilton, R. (eds.) 5Th International Conference on Debris-Flow Hazards: Mitigation, Mechanics, Prediction and Assessment (2011)
51. Scheidl, C., Rickenmann, D., Chiari, M.: The use of airborne LiDAR data for the analysis of debris flow events in Switzerland. Nat. Hazards Earth Syst. Sci. **8**(5), 1113–1127 (2008). doi:10.5194/nhess-8-1113-2008
52. Schraml, K., Thomschitz, B., McArdell, B., Graf, C., Kaitna, R.: Modeling debris-flow runout patterns on two alpine fans with different dynamic simulation models. Nat. Hazards Earth Syst. Sci. Discuss. **3**(2), 1397–1425 (2015)
53. Sosio, R., Crosta, G., Frattini, P.: Field observations, rheological testing and numerical modelling of a debris-flow event. Earth Surf. Process. Landf. **32**(2), 290–306 (2007)
54. Vasiliou, A., Maris, F., Varsami, G.: Estimation of Sedimentation to the Torrential Sedimentation Fan of the Dadia Stream with the Use of the ToprunDF and the GIS Models, pp. 207–214. Springer, Berlin (2011)
55. Westoby, M., Brasington, J., Glasser, N., Hambrey, M., Reynolds, J.: structure-from-motion photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology **179**, 300–314 (2012). doi:10.1016/j.geomorph.2012.08.021