

Article

Application of Sensitivity Analysis for Process Model Calibration of Natural Hazards

Candace Chow *, Jorge Ramirez and Margreth Keiler

Institute of Geography, University of Bern, 3012 Bern, Switzerland; jorge.ramirez@giub.unibe.ch (J.R.); margreth.keiler@giub.unibe.ch (M.K.)

* Correspondence: candace.chow@giub.unibe.ch

Received: 15 May 2018; Accepted: 12 June 2018; Published: 15 June 2018



Abstract: Sensitivity analysis (SA) describes how varying inputs to a model subsequently varies its outputs. Its inclusion can support the systematic calibration of numerical models to back-calculate intensity properties of past torrent events that would otherwise be difficult or impossible to collect during their occurrence. Sensitivity analysis for model calibration is assessed with the back-calculation of a known torrent event. In particular, FLO-2D, a cell-based numerical model, is used to simulate the 2005 debris flow event that occurred in Brienz, Switzerland. Under 4000 simulations were completed with ranges of physically reasonable parameter values. Model results were compared in 3-dimensions with both sediment deposition extents (x, y) and estimated deposition heights (z) from available post-event images. The comparisons highlighted that more accurate input and validation data, namely the flow behavior of hazardous processes and post-event deposition heights, are required to produce stronger agreements between simulated and observed results. Furthermore, the application of SA for model calibration supports systematic exploration of large parameter spaces characteristic of complex phenomena like natural hazard events. These findings demonstrated how important model input factors can be identified, which provide guidance for future data collection efforts to capture both the rheology and the spatial distribution of hazards more accurately.

Keywords: sensitivity analysis; calibration; physical vulnerability; risk analysis; summary scalar variables

1. Introduction

Torrents are defined as steep waterways in mountainous environments [1]. Hazardous torrent processes are characterized by the rapid propagation of large quantities of available sediments, debris, and water from an upslope source via a transit zone, to a downslope depositional area where human settlements may be established. In theory, torrent processes can be further differentiated as debris flows, hyperconcentrated flows, or fluvial sediment transport, based on the respective characteristics and dominant processes of each event [2,3].

While conventional approaches based the classification process on flow behavior alone [4], peak discharge has since been recommended as a complimentary criterion [3,5]. Observable flow characteristics reflect variable concentrations of water and sediment, which provides insight into the internal physics of the flow. For instance, debris flows typically produce thicker, more hummocky and lobate depositions, and are characterized as very rapid to extremely rapid flows of saturated, non-plastic materials along steep channels, headed by a coarse surge front [3]. While debris floods are associated with the transport of considerable quantities of coarse sediments, the flows are generally characterized as thin, wide sheets of materials [3]. Relative to debris floods, hyperconcentrated or sediment-laden flows transport relatively less, albeit still notable quantities of fine sediment in suspension [6,7]. In particular, mudflows have been defined as very to extremely rapid flows of

saturated plastic debris in a channel, and are characterized by significantly larger quantities of water content with respect to the amount of solid source materials; the plastic index is greater than 5% [8]. Following the definition by Bradley and McCutcheon [9], mudflows also have sufficient viscosity to transport sizable boulders, and natural and anthropogenic debris within a matrix of smaller-sized particles. It is possible for the flow behavior of a single torrent event to evolve as it occurs, depending on the types, velocity, and quantities of materials propagated downwards. In this respect, events may often be more realistically described as a combination or evolution of the aforementioned torrent process types.

The distribution of these materials downslope from the active zone is further influenced by underlying site characteristics (e.g., topography, presence of confined or unconfined preferential pathways). Due to the combined effects of composition, flow behavior and site-specific characteristics, torrent processes are associated with variable peak discharges, sediment transport capacities, momentums, and subsequently differential potentials to cause damage to elements at risk upon impact.

Assessing the physical vulnerability of elements at risk (e.g., affected buildings) due to these processes is a part of consequence analysis within risk assessment, where the intensity of a given event is related to the damages sustained [10,11]. The design and implementation of effective risk mitigation strategies is dependent on the results of such analyses. The estimation of expected direct losses, as a result of the hazard process patterns with respect to the properties of exposed elements, is possible with the derivation of representative physical vulnerability functions [12].

However, a persistent challenge in vulnerability studies on torrent processes is the high uncertainty and limited amounts of direct, field-based observation data that is available, since its collection during the occurrence of these types of events is difficult or impossible [13]. Detailed analyses have been conducted for events that result in high losses in both Austria and Switzerland. These analyses generally report on the triggering and boundary conditions of the particular torrent process, the process evolution, the extent of runout zones, and estimates of eroded materials deposited downslope (e.g., [14,15]). However, comprehensive information on processes, damage patterns, and their interactions with structural building properties have not been adequately documented to date.

To address this challenge, proxies have been adapted to inform about event intensities, including, but not limited to sediment deposition heights, velocities, and impact pressures [16,17]. It is of interest to replicate past events with process models to determine if simulated intensity proxy data can be considered in further consequence analysis. Following Mazzorana et al. [12], process modeling is the first three of five steps to accurately assessing the physical vulnerability of the built environment. A range of recognized methods have been applied to different process models including empirical [7,18–23], empirical-statistical combined with simple flow equations [24], topographic gradient-based [25], numerical-based with the integration of shallow water equations [26–41], and smoothed particle hydrodynamics (SPH) or Lagrangian [42–46] (see References [47,48] for review). Of the numerical models, 1- (e.g., DAN-W [28]; DFEM-1D [49]) or 2- (e.g., FLO-2D [27]; RAMMS-DF [35,36]; TopRunDF [7]; MassMov2D [34]) dimensional runout modeling approaches can be adapted. Rickenmann [50] provides a comprehensive overview of the advantages and limitations of each type of modeling approach as a combination of how flows are propagated and distributed on the alluvial fan. Furthermore, flows are represented as either single-phased, homogenous fluids with constant rheological properties, or dual-phased, heterogeneous matrices. Each approach has specific data requirements, which may limit its application to data scarce case studies. The different model approaches also highlight tradeoffs between minimizing computational time and various degrees of accuracy in simulated results. Consequently, the choice of a suitable model is determined by data availability and how representative a given approach is to representing a complex, heterogeneous 3D problem, while minimizing the computation time required to return a solution.

Furthermore, known sources of uncertainties can have adverse implications on model results [51]. Sources associated with debris flow models include, but are not limited to, the quality of model input and calibration data, initial and boundary conditions, how accurately the model structure

represents events of interest, the sensitivity of defined parameters, and the calibration method applied. Accurate topographic representation is imperative to model both the propagation of debris flows within the torrent channel and the lateral distribution of materials that exceed bankfull conditions onto the alluvial fan [52]. For example, Rickenmann et al. [32] demonstrated that debris flow model results are sensitive to the presence of local topographic features, which can divert flows and consequently determine where material is deposited on alluvial fans. Moreover, infrequent generation of topographic inputs, which are used in debris flow models, may not necessarily reflect the same conditions when the event occurred. This temporal mismatch results in modelling with topography that does not accurately represent initial conditions and can include differences in channel slope steepness, channel width, and inaccurate representation of mitigation structures along the channel [53]. Additional uncertainties in debris flow modelling stem from the lack of basic data required to reconstruct the characteristics of debris flow events. These may include the point where the debris flow event was initiated, the hydrograph peak and duration [16,53], and the volume of material at the start of the event and further entrained during its course [54]. In lieu of actual data, working assumptions are made to reconstruct ranges of plausible data values, which contains inherent uncertainties that are introduced in the simulated results. For example, Rickenmann et al. [32] attributed sources of model errors to the lack of detailed data needed to describe a debris flow consisting of multiple surges, which was described with the use of a simpler, single surge hydrograph instead. The degree of model complexity can also be a source of uncertainty. For instance, sediment entrainment during a debris flow changes the volume of materials and flow behavior [55,56]. Simpler debris flow models do not replicate entrainment, but partly consider this process by including additional quantities of material at the debris initiation point, while more complex models explicitly replicate the spatially and temporally-distributed erosion of channel materials as the flow is propagated [57]. However, while a more complex model that includes entrainment can better replicate debris flow heights near the point of initiation [57], this additional process introduces another source of uncertainty into the model with the introduction of additional parameters (e.g., erosion rate), which require values that may not be known. Event and site-specific parameters for debris flow models are estimated through calibration, and efforts usually focus on flow resistance [48] and rheological parameters [58]. These parameters are associated with wide ranges of plausible values in literature and may be difficult to determine which are most representative for a given event. The majority of debris flow model research uses a trial and error approach to calibration [58], and may not arrive to optimal parameter sets because calibration is time consuming and the process may be ended prematurely [52]. Methods exist to efficiently derive parameter sets (e.g., genetic algorithms), but these methods are rarely, if ever, employed in debris flow modelling.

Given the inherent complexities that characterize natural hazard processes and the contributing sources of uncertainties, modeling past torrent events requires the exploration of relatively large parameter spaces, even with model simplifications. Sensitivity analysis (SA) describes how varying inputs in a numerical model subsequently varies its outputs [59]. The inclusion of SA for model calibration supports a better understanding of model behavior, its parameterization, and its associated uncertainties [60].

In this respect, the inclusion of SA offers several advantages as a part of a sound model calibration and evaluation framework. Firstly, it is instrumental in reducing the number of parameters that require calibration. Calibration is an example of an inverse problem, where the most optimal agreement between simulated and observed reference data is obtained with parameter combinations and values that result in higher model performance. Secondly, SA supports the identification of the degree of influence that input factors have on model simulation results. Thirdly, SA can highlight limitations of model calibration due to residual sources of uncertainties once parameter uncertainties are accounted for [59]. Consequently, the inclusion of SA into the modelling process identifies optimal results while highlighting application-specific limitations with greater efficiency. Furthermore, epistemic uncertainty about natural hazard phenomena that occur in complex systems remains prevalent [61]. In particular, the combination of model and input data limitations compromises the ability to learn about the

most influential parameter(s) within the systems of interest to a sufficiently high degree of accuracy. Uncertainties due to imperfect knowledge about initial conditions and simplified representations of model inputs are generally addressed with ensembles of model predictions, where each simulation represents a different choice of parameter combinations and values. A structured statistical approach to assess parametric uncertainty and model performance is preferable, where decisions about parameters can be made in a transparent and explicit way, using methods that can be easily understood [62]. This study addresses the two aforementioned challenges, firstly, relating to the large parameter spaces needed to be explored to fully capture the complexities of hazardous torrent events, and secondly, evaluating model performance to gain a better understanding of epistemic uncertainties.

In response to the first challenge, the utility of SA for model calibration is assessed with the back-calculation of the 2005 debris flow event that occurred in Brienz, Switzerland. Taking all of these requirements into consideration, the FLO-2D [27], a simplified, physically-based, 2-dimensional numerical model, was used to model the event. Single-phased models like FLO-2D are commonly used to simulate debris flows by researchers and practitioners working in the risk community (Table 1) as a computationally efficient first step to gain insight into these complex processes [32,63]. In the study conducted by De Blasio et al. [38], it was suggested that in principle, the presence of larger blocks interspersed within a mud matrix may behave comparably to a pure Bingham fluid for certain types of flows. This working assumption was also adopted to minimize the amount of computation time to support the evaluation of a wider model parameter space. Furthermore, the application of more complex models often requires input data and initial conditions that exceed that which is available [64,65]. With respect to specific model data requirements, FLO-2D results have been reported to be strongly influenced by topography [58]; the availability of a high resolution SwissALTI3D digital elevation model (DEM) [66] satisfied this requirement. Previous studies (e.g., [58]) have also cited that FLO-2D is capable of generating accurate runout distances and capturing the distribution of materials across the fan through back-calculation. Finally, to support subsequent investigations of building vulnerability with simulated hazard intensities, FLO-2D is capable of generating sediment deposition (flow) heights, in addition to flow velocities and impact pressures.

Model performance is assessed against post-event observations of sediment deposition extent, sediment deposition heights, and a single point estimate of flow velocity for close to 4000 completed simulations. Focusing only on studies conducted with the commonly applied FLO-2D model, Table 1 presents an overview of published literature on the assessment of simulated results specific to the back-calculation of torrent events. To date, the performance of simulated torrent events has generally been assessed by visual comparison, with limited instances where quantitative or hybrid approaches are applied to consider validation data in 3-dimensions. In particular, while comparisons of observed sediment deposition extents are more common, the additional inclusion deposition heights is limited. For instance, only three known studies with FLO-2D assessed results with both post-event observations of sediment deposition extent and deposition heights. From these cases, sample sizes of observed points were limited ($n < 20$), and no further statistical analyses were published at the time this manuscript was prepared. Furthermore, other studies that quantified model performance presented the percent of over or under prediction by sediment deposition extent or only visually displayed the simulated results without quantification.

In light of these findings, statistically-based performance metrics are applied in this study to address the second challenge. This supports the quantitative formulation of aggregated uncertainties related to data quality, parameterization, and model suitability; it is a step towards effectively identifying priorities for data collection and future modeling efforts based on quantitative methods. In effect, the methods proposed in this study combines SA for calibration together with the evaluation of model performance and behavior with a set of statistically-based metrics to support the process in a systematic and efficient way.

Table 1. Summary of published literature that assessed FLO-2D model results.

Cases from Literature	Sources of Rheology Values		Model Validation			Reference Data	
	Lab Analyses of Post-Event Field Samples	From Literature	Visual Comparison	Quantification	Sediment Deposition Extent	Sediment Deposition Heights	Number of Sample Points
Sosio et al. [67]	✓	✓	✓	✓	✓	✓	18
Quan Luna et al. [16]		✓	✓		✓	✓	13
Lin et al. [68]	✓	✓	✓	✓	✓	✓	8
D’Agostino et al. [69]	✓	✓	✓	✓	✓		
Boniello et al. [70]	✓	✓	✓	✓	✓		
Rickenmann et al. [32]		✓	✓		✓		

2. Materials and Methods

2.1. Site Description

The proposed framework is applied to a well-documented, high magnitude debris flow event that occurred along the Glyssibach in Brienz, Switzerland in 2005. A detailed description of the study site and characteristics of the debris flow was obtained from post-event documentation (i.e., Lokale Lösungsorientierte Ereignisanalyse or LLE) prepared by local experts and published in Reference [71].

Brienz is located at the fan of the Glyssibach, a 1.98 km² alpine catchment situated between 563 m and 2039 m above sea level. The underlying bedrock on location is Lower Cretaceous limestone. The combination of marl and siliceous limestone found in the vicinity and contributions of weathered schist provided sources of sediment for the torrent processes.

Between 19 and 23 August 2005, continuous rainfall varying between 6 and 16 mm/h fell in the upstream catchment. This resulted in 320 mm of rain over 72 h, which gradually transported finer sediments downwards along the Glyssibach into Lake Brienz, prior to the onset of the main debris flow event. An eyewitness account described the northern bank, where the channel meets the lake, reaching its retention capacity. The subsequent backpropagation of sediment was deposited upwards along the main channel until the limit, which is identified in Figure 1. While solid materials did not breach the confines of the channel, witnesses reported small quantities of water spilling onto the banks. This observation indicates reduced carrying capacity prior to the occurrence of the main event.

The debris flow was triggered by a landslide in the early morning on 23 August 2005 and lasted for approximately 15 min. The maximum discharge (Q_{\max}) was estimated to be between 140 and 160 m³/s. Available sediment was entrained and approximately 72,000 m³ of total bulk volume was subsequently propagated along the main channel into the settlement, deposited on the alluvial fan, and subsequently flowed into Lake Brienz. Flow velocities between 6 and 10 m/s at the Glyssibrücke was estimated with the superelevation approach. This is based on the forced vortex equation, where the difference in surface elevation of a debris flow is determined as it travels within a known channel curvature [73]. Debris flow materials were described by local experts to be relatively homogeneous, resulting in a viscoplastic mud matrix composed primarily of fine fractions of silt and sand. In particular, the mud matrix of the 2005 event is estimated to have been comprised of 30–70% water; 5–8% of the total sediment volume is comprised of clay. The channelized mud matrix was capable of entraining coarse, unconsolidated materials (e.g., vehicles, large woody debris) and transported boulders between 3 and 5 m in diameter to the bottom of the alluvial fan. For this reason, the 2005 event is referred to as a debris flow.

2.2. Methods

Generally, a pre-screening stage is carried out to optimize modeling efforts by focusing on a subset of inputs based on relative importance [74,75]. Figure 2 outlines the required inputs and expected outputs of the dynamic FLO-2D simulation model following the pre-screening stage. Both sets of boundary conditions and model parameter values reflect characteristics of the study site and the 2005 debris flow event described in Section 2. These characteristics were translated into ranges of physically plausible parameter values used to define the parameter space of interest; Table 2 summarizes key parameters and associated ranges of values identified from various sources. In particular, full ranges of acceptable input values are defined in the FLO-2D model reference manuals. Truncated a priori ranges of these values were further defined based on a combination of literature-based values from debris flow modeling studies with FLO-2D and values from expert knowledge about the debris flow that occurred specifically in Brienz. Certain identified parameters are set as initial conditions with fixed values based on recommendations in the literature, with respect to the degree of uncertainty in required model parameters. For example, accurate definition of uncertain flow behavior was prioritized over simulations designed to investigate the impact of changing the spatial resolution of the computational

grid. Furthermore, the amount of time that would be required to evaluate both input factors and model parameters was also a consideration. These considerations effectively minimized the parameter space and optimized the search for parameter combinations that best match the observed reference data.

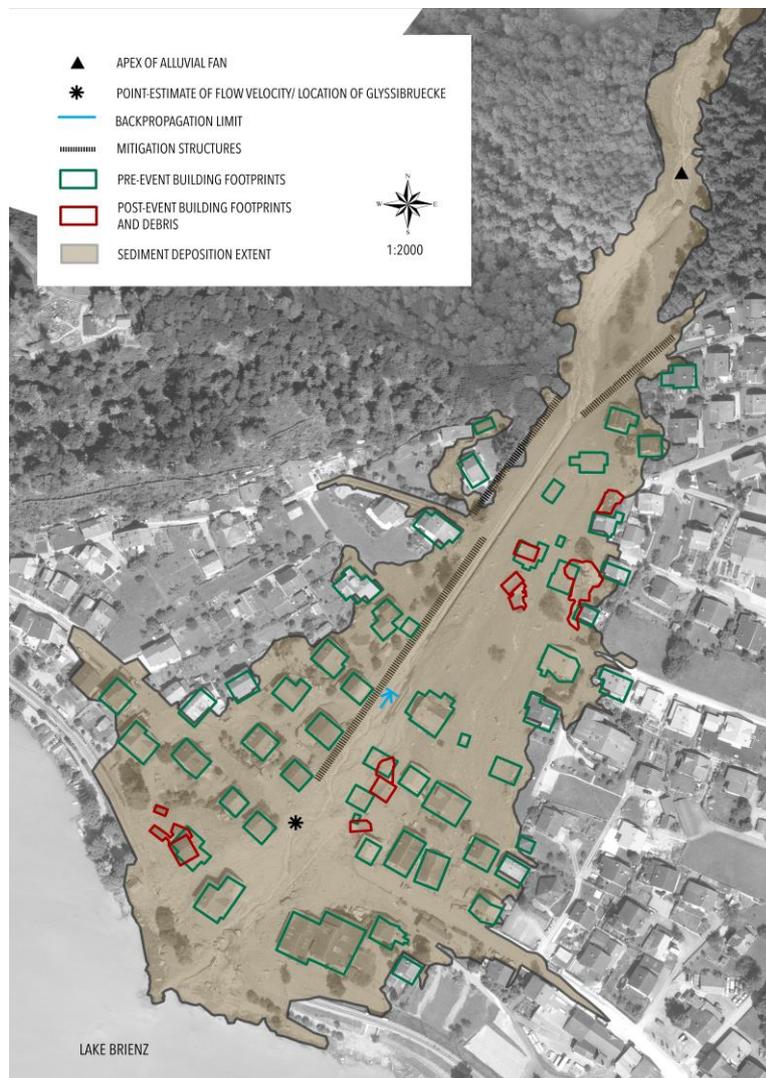


Figure 1. Vectorized post-event orthophotograph [72] capturing the distribution of mud and debris in Brienz.

In this study, a global sensitivity analysis was conducted, where the model output was obtained by varying inputs across their entire feasible space rather than around a reference value. One-at-a-time (OAT) or all-at-a-time (AAT) methods can be applied to calibrate each input factor independently or all input factors simultaneously [76]. In this particular study, the AAT method was initially applied to explore the wider parameter space by considering model output sensitivity to both direct changes to input factors and to joint input interactions; this corresponds to about 45% of the total simulations completed for this study or 1764 model runs. The OAT method was applied to specific parameters once the most influential input factors were identified from evaluations of AAT results. Global sensitivity analysis was performed semi-automatically with the implementation of Sikuli scripts [77] coupled with an executable batch file program that was provided by the developers of FLO-2D. Due to constraints with the batch program (i.e., only a maximum of 20 simulations can be run in a series within the batch program at a time), fully automated calibration was not supported.

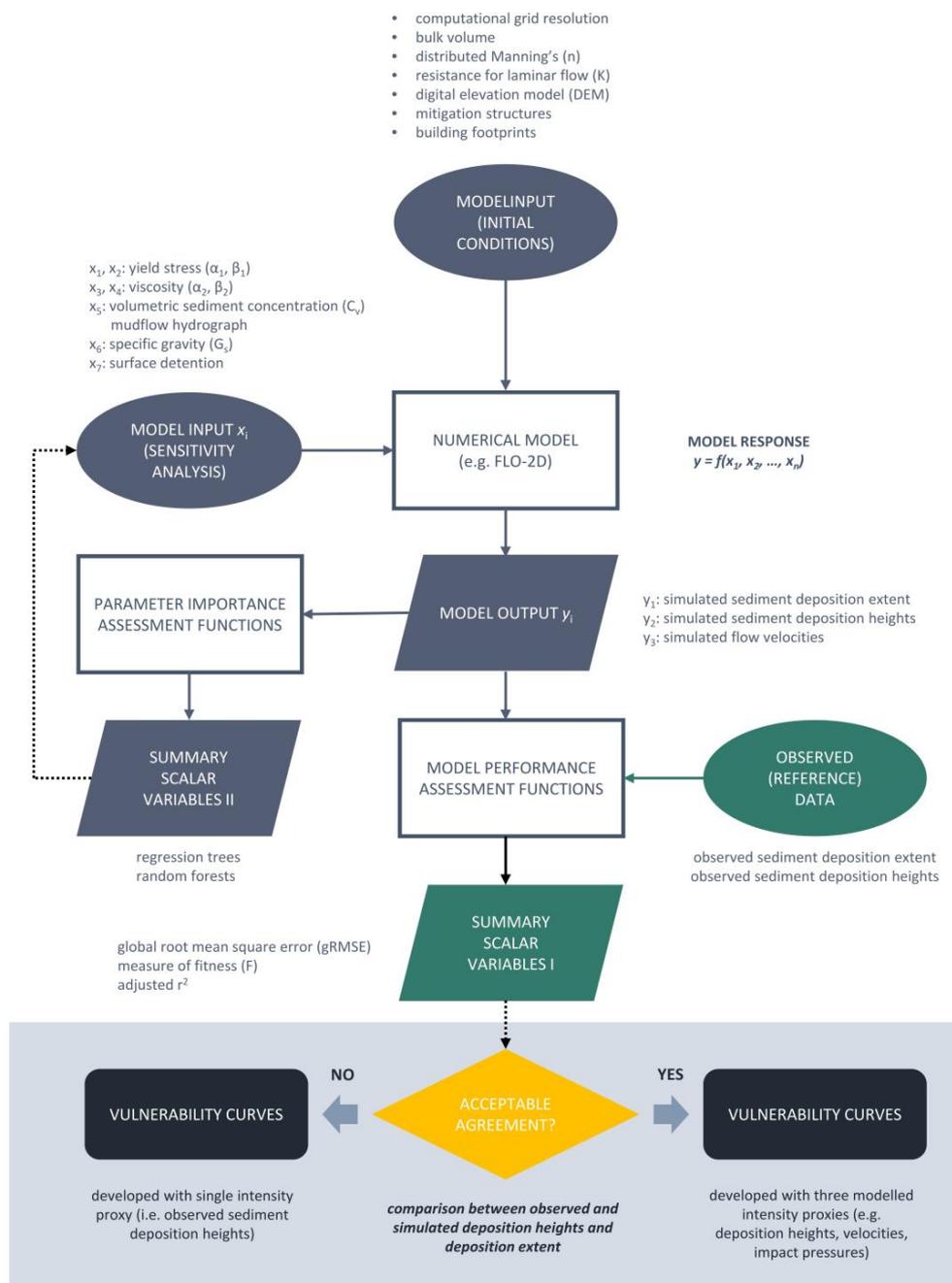


Figure 2. Schematic illustrating how the different sources of data are fed into the dynamic numerical model to generate simulated outputs; summary scalar variables are subsequently produced to support model performance evaluations and provide insight about the potential use of simulated intensity proxies to further develop physical vulnerability curves.

A distinction is made between model outputs and their translation into summary scalar variables with different performance assessment functions; these variables (e.g., global root mean square error, fitness measure F) quantify model performance by comparing simulated results to observed reference data [59]. In particular, completed simulations were evaluated against reference observation data to determine whether higher agreements were obtained and new sets of input factor values were defined based on the results. This procedure was repeated iteratively in phases until no further gain in model performance was detected with respect to the available event data. Consequently, the evaluation of model outputs with summary scalar variable describe how influential variations of input factors are.

Furthermore, the results provide insight about the feasibility of using simulated intensity proxies to further develop physical vulnerability curves (e.g., [16]).

Table 2. Overview of required inputs in FLO-2D and feasible ranges of values specific to the 2005 debris flow event that occurred in Brienz.

Initial Conditions	Input Factors for Calibration	Description	Symbol	Units	Feasible Ranges of Values	
					for FLO-2D Model	for 2005 Debris Flow in Brienz
✓		computational grid resolution		m		5
✓		sediment volume		m ³		70,000–72,000
✓		(distributed) Manning's values	n	m ^{-1/3} ·s		forested areas (0.33), channels and streets (0.01), sparsely vegetated settlement areas (0.08)
✓		resistance parameter for laminar flow	K		24–50,000	determined based on floodplain grid element's Manning's <i>n</i> value
	x ₁	yield stress, α ₁	τ _y	Pa	0–∞	unknown; see Table 3 for potentially applicable values
	x ₂	yield stress, β ₂			0–∞	
	x ₃	viscosity, α ₁	η	Pa·s	0–∞	
	x ₄	viscosity, β ₂			0–∞	
	x ₅	volumetric sediment concentration (determines mud hydrograph)	C _v		0.03–0.90	0.30–0.70
	x ₆	specific gravity	G _s		2.5–2.8	2.5–2.8
	x ₇	surface detention		m	0.01–0.50	model-specific input factor; calibrated with feasible model ranges

2.3. Data Sources

2.3.1. Initial Conditions and Fixed Parameters

In this study, pre-screening began with the identification of accepted ranges of required input values for the numerical model [78]. Truncated ranges of physically feasible values were defined to reflect the specific 2005 debris flow event based on expert consultation. Additionally, unknown values were identified from literature describing potentially similar study sites and processes. Information from past SA studies with FLO-2D (e.g., [67–69,79]) was compiled; findings highlighted the degree to which the model outputs varied with the definition of certain conditions and parameters. Certain inputs were defined as initial conditions with fixed input values, while others were defined as input factors with a range of plausible values, based on the findings on relative parameter importance, on available information about the 2005 event and the amount of computational time required to explore a given number of parameters. For the 2005 mudflow in Brienz, initial conditions were defined with data about the underlying topography, the bulk volume of sediment and water propagated downslope, the inclusion of mitigation structures and building footprints into the topography, surface roughness, and the spatial resolution of model grid elements (Figure 2; Table 2). Sensitivity analysis for model calibration was applied to the defined input factors, which effectively reduced the size of the original problem by minimizing the model parameter space.

A swissALTI3D digital elevation model (DEM; [66]) was obtained at 2 m spatial resolution and was resampled to 5 m. While decreasing the spatial resolution of grid elements limits the ability to perform highly detailed analyses, the results provided an overview of reasonable reproductions of the simulated process with greater computational efficiency. To account for any notable changes represented in DEM since the occurrence of the mudflow event, additional or modified mitigation structures were changed to reflect the topographic conditions in 2005 based on information from expert consultation and review of pre-event orthophotographs. Furthermore, the DEM was adapted to

include witness accounts of sediment backpropagation up to channel bankfull conditions until the limit of the blue arrow indicated in Figure 2, prior to the onset of the main debris flow event.

While some studies have calibrated for Manning's coefficient n (e.g., [79]), which describes surface roughness as a function of land cover types, fixed distributed values were assigned to forested areas ($0.33 \text{ m}^{-1}/3 \cdot \text{s}$), channels and streets ($0.01 \text{ m}^{-1}/3 \cdot \text{s}$), and sparsely vegetated settlement areas ($0.08 \text{ m}^{-1}/3 \cdot \text{s}$) in this study [80].

2.3.2. Input Factors for Calibration

Following the definition of initial conditions, site-specific ranges of physically feasible input factor values are defined based on expert knowledge and from published literature. This ensures that simulated outputs, produced by varying these constrained values, are realistic for a given process and event. Different combinations of input factors represent the simplified physics that govern the flow behavior of the process, how the materials spread laterally across the alluvial fan and downslope from the defined triggering location.

For all of the simulations, the same total bulk volume is transported downslope. However, the volumetric sediment concentration (C_v) input factor is calibrated for, which is the ratio between solid materials and water. Within the solid component, different grain size distributions can be defined. This reflects the importance of grain size on energy dissipation in granular flows, which translates to different effects on geophysical flow mobility [81].

An inflow mud hydrograph is required for each simulation and is generated based on the defined C_v and discharge data. The amount of precipitation that initiated the debris flow could not be determined from rain or discharge records most proximal to the study site. Consequently, an expert-based estimate of maximum discharge from the main channel was used to produce a series of mud hydrographs based on variations of different volumetric sediment concentrations.

2.3.3. Calibration of Rheological Values

Flow behavior is partially characterized by the amount of volumetric concentration of a fluid matrix, C_v , where V_s and V_w represent the respective sediment and pore water volumes in a given mixture [82].

$$C_v = V_s / (V_s + V_w) \quad (1)$$

Empirical data on relatively homogeneous, single-phased geophysical flows have shown that the amount of interstitial fluid and its associated properties determine both the viscosity and yield strength [83–85]. In general, both viscosity

$$\eta = \alpha_1 e^{\beta_1 C_v} \quad (2)$$

and yield stress

$$\tau_y = \alpha_2 e^{\beta_2 C_v} \quad (3)$$

increase exponentially with the volumetric concentration of fine sediments.

The empirical coefficients or material parameters, α_i and β_i , are key parameters within the FLO-2D model that define the internal resistance of the mudflow materials. In particular, the study conducted by D'Agostino et al. [69] demonstrated that the simulated results are highly dependent on the correct definition of rheological parameters. While these values are typically determined by lab-based analyses of representative field samples collected shortly post-event (e.g., [67,70,82,83]) to describe process flow behavior, it may not always be possible to obtain such samples. In lieu of post-event field samples, rheological values have been calibrated from existing values reported in the literature (e.g., [67,70]).

Consequently, in addition to the aforementioned input factors, rheological values were also calibrated for this study due to the lack of more detailed information about the sediments that

were transported. A selection of rheological values (x_{1-4}) that was compiled from the literature for consideration and subsequent calibration in FLO-2D (Table 3).

Specific to FLO-2D, the additional calibration of an input factor referred to as sediment detention (x_7) modifies the defined rheology by introducing a uniform height threshold (in m) to each computational grid cell, so that the volume of materials propagating downslope is retained until the defined height is exceeded.

Table 3. Summary of selected rheological values from the literature to evaluate with the FLO-2D model with respect to the 2005 debris flow event, in lieu of location-specific post-event field samples; a unique ID is assigned to each group of values.

ID	Sources from Literature	Sample Name	Yield Stress (T_Y)		Viscosity (η)	
			α_1	β_1	α_2	β_2
			Dynes/cm ²	Poises		
A	O'Brien and Julien [83]	Glenwood 4	0.00172	29.5	0.000602	33.1
		Glenwood 1	0.0345	20.1	0.00283	23
		Glenwood 3	0.0765	16.9	0.648	6.2
		Glenwood 2	0.000707	29.8	0.00632	19.9
B	O'Brien and Julien [83]	Aspen pit 1	0.181	25.7	0.036	22.1
		Aspen pit 2	2.72	10.4	0.0538	14.5
		Aspen natural soil	0.152	18.7	0.00136	28.4
		Aspen mine fill	0.0473	21.1	0.128	12
		Aspen natural soil source	0.0383	19.6	0.000495	27.1
		Aspen mine fill source	0.291	14.3	0.000201	33.1
C	Sosio et al. [67]	Scenario A	0.0013	23	0.0000283	19
		Scenario B	0.000093	23.5	0.000183	19
		Scenario C	0.0011	21.8	0.0000283	18.2
		L1 (deposit sample)	0.000127	22.8	0.0000297	18.8
		L2 (source area sample)	0.0004	22	0.000203	18
D	D'Agostino and Tecca [69]	Rio Dona	0.05	22	0.0015	22
		Fiames	0.152	18.7	0.0075	14.39
E	Lin et al. [68]	Nan-Ping-Kern	0.2433	4.116	0.001386	3.372
		Jun-Kern; Err-Bu	0.7299	12.348	0.004158	10.116
		Shan-Bu	0.4055	6.86	0.00231	5.62
		Fong-Chu	0.85155	14.406	0.004851	11.802
		Tung-Fu Community	0.811	13.72	0.00462	11.24
		Her-Ser 1	1.0543	17.836	0.006	14.612
		Chui-Sue River	1.2652	16.464	0.924	14.612
F	Boniello et al. [70]	Fella sx debris flow 1	0.000005	42.01	0.00000002	42.23
		Fella sx debris flow 2	0.0383	19.6	0.000495	27.1

2.3.4. Observed Reference Data for Model Verification

Observed reference data was based on contributions of volunteered geographic information (VGI) from local residents, experts, and authorities. On the building level, available data included post-event photographs taken by residents. Point-based estimates of sediment deposition heights (Figures 3 and 4; in meters) and degrees of loss were determined for each affected building.

All height estimations are associated with a confidence level, primarily based on the quality of the post-event photograph and the ability to determine deposition heights accurately. Figure 4 shows three examples of photographs of affected buildings in Brienz. In the top left, sediment deposition heights can be observed on the outer wall of the building, however, the location of the ground (i.e., reference) cannot be clearly determined.

As a result, this estimated observed height would be assigned a lower confidence level (2 or 3: somewhat uncertain) than the height estimated from the photograph in the top right (1: very certain), where the ground can be observed. While sediment deposition heights can be observed in the orthophotograph (bottom) for several buildings, the associated confidence level (4: highly uncertain) is

much lower due to the angle at which the photograph is taken. Figure 4 shows the spatial distribution of estimated sediment deposition heights with respect to associated confidence levels. Post-event orthophotographs [72] provide a synoptic view of the affected site and building level damages. The sediment deposition extent (Figure 3) was delineated from an orthophotograph that shows the limits of where the debris flow materials were deposited in Brienz.

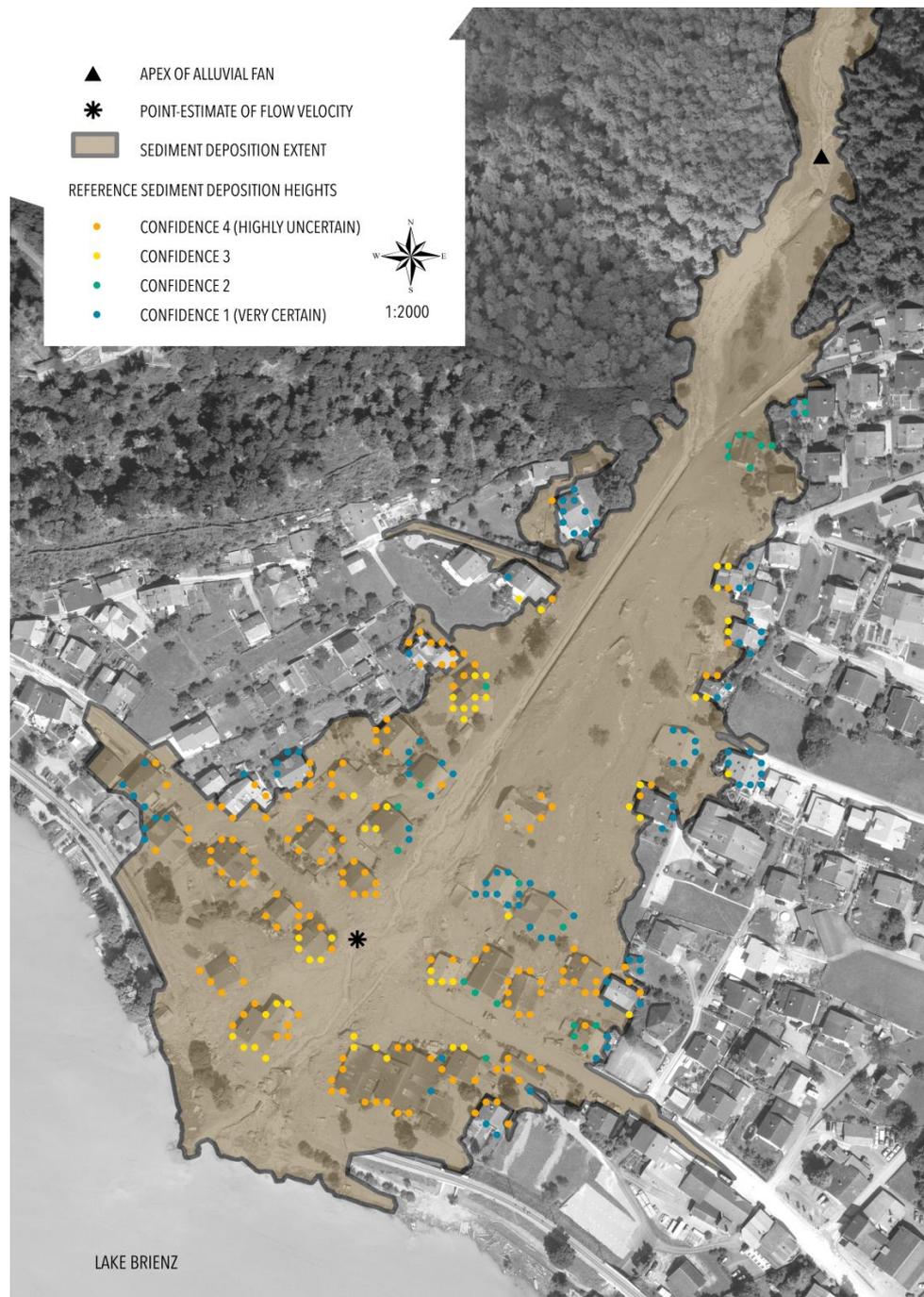


Figure 3. Two main sets of reference data are available for the study site including sediment deposition extent and 317 point estimates of sediment deposition heights that are associated with spatial coordinates and confidence levels.

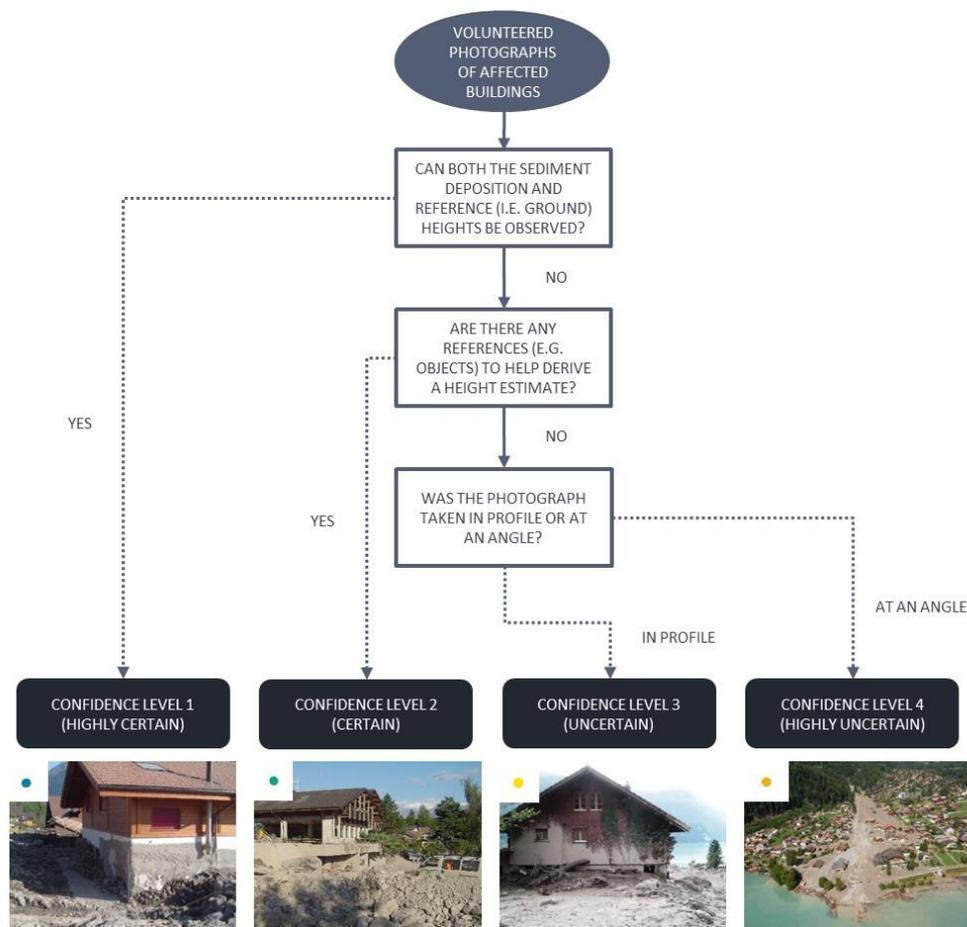


Figure 4. Selected examples of post-event photographs of affected buildings in Brienz; sediment deposition heights estimated from the photographs are associated with variable levels of confidence.

In summary, three sources of reference data were available to support model validation, namely the sediment deposition extent, estimates of sediment deposition heights distributed across the alluvial fan area, and a point-estimate estimate of flow velocity (i.e., between 6 and 10 m/s) at the Glyssibrück.

2.4. Model Description

FLO-2D [27] is a two-dimensional, physically-based distributed flood routing model that is also capable of simulating non-Newtonian sediment flows. In particular, cumulative shear stress in hyperconcentrated sediment flows is simulated, which includes debris flows, mudflows and mud floods [58]. Cumulative shear stress in the model is comprised of five components, namely, cohesive yield stress, Mohr–Coulomb shear, viscous, turbulent, and dispersive shear stress. A quadratic rheological model function of sediment concentration adds turbulent and dispersive terms to the original Bingham equation when the aforementioned components are expressed as shear rates [86]. Furthermore, the model is driven with the definition of a mudflow hydrograph and conserves the defined volume of sediment and water over a square grid-based, user-defined computational domain based on a cell storage approach. Variable combinations of model parameter values results in a specific rheology or flow behavior, which effectively describes how the mud matrix is distributed over a given landscape. In particular, the geophysical flows are modeled as simplified Bingham fluids, which is defined by yield stress and viscosity coefficients [86]. The rheology can be further calibrated with the definition of the model-specific sediment detention input factor.

2.5. Performance Metrics

While numerical models have been previously applied to replicate hazardous events, it is instrumental to gain a sound understanding about simulated outputs, in addition to model- and calibration-specific limitations, with respect to available data. Performance metrics provide a means to assess these components in a systematic and reproducible way. In this study, two sets of performance assessments are carried out to address two different objectives. In particular, the assessments provide insight into model performance with respect to validation data and parameter importance on simulated results.

Model performance is evaluated on the degree of agreement between simulated results with observed data, which is quantified by a series of summary scalar variables. In this study, a three-dimensional evaluation was conducted with validation data describing both the sediment deposition extent (x , y) and sediment deposition heights (z). Furthermore, a single, point-based comparison of simulated and estimated flow velocities was performed to determine whether the simulated outputs were within a reasonable range.

A measure of fitness, F , quantifies the binary (i.e., yes or no) agreement between observed and simulated sediment deposition extent per grid cell as a percentage over the study site. Both instances of over- and under-prediction in the simulated extent were penalized with respect to the observed extent [87].

$$F[\%] = \frac{\text{extent}_{obs} \cap \text{extent}_{sim}}{\text{extent}_{obs} \cup \text{extent}_{sim}} \times 100 \quad (4)$$

The global root mean square error (gRMSE) provides insight on the overall agreement between observed and simulated sediment deposition heights and is expressed in meters. Once the summary scalar variables were computed, model outputs were ranked by highest F scores and lowest gRMSE values to identify combinations of input factors that matched the validation data most closely.

Models of complex phenomenon, such as natural hazards, are often characterized by a large number of input factors that may interact in non-linear ways [88]. To minimize the uncertainty associated with SA findings, a larger number of simulations is required to increase the sampling density [89]. This is especially important for the exploration of high dimension parameter spaces, so that there is a higher chance of capturing complex interactions among parameter sets. In this study, 3876 simulations representing different input factor combinations were completed. Further assessment of input factor combinations was conducted with two statistical approaches, namely regression trees and random forests.

Regression trees [90] begin with a root node and end with terminal nodes or leaves. The root node of the regression tree accounts for the total number of completed simulation runs or training set of data that are being considered in the analysis, with respect to a defined objective function (e.g., higher F or lower gRMSE). The training set is recursively partitioned into subsets of simulations that reflect combinations of input factor values that satisfy the objective function. The complexity parameter, cp , is a value that defines the amount by which splitting the current node improves the relative error. In the case of ANOVA splitting, if the overall r^2 is increased by the defined cp value at each step, the regression tree continues to grow. In this study, two regression tree models were generated, based on average gRMSE and F scores, respectively.

A random forest is based on a set of independently generated regression trees [91]. To further the analysis, a measure of predictor importance (%IncMSE) was returned after a random forest was grown. If a predictor is important in the model, randomly assigning it realistic values (i.e., permuting predictor values over the dataset) should result in a worse prediction with the same model. A comparison can be made with respect to a predictive measure (i.e., mean squared error or MSE) by comparing the model prediction results based on original and permuted datasets. It is expected that predictions with the original dataset are better than ones made with the permuted dataset.

Consequently, %IncMSE represents the increase in the mean squared error (MSE) of predictions, which is estimated by out-of-bag cross validation. First, a regression forest is grown. The MSE

associated with the original model is computed and assigned to MSE0. Each predictor variable is then permuted a new model MSE(x) is calculated. %IncMSE of the predictor is determined by: $(\text{MSE}(x) - \text{MSE0})/\text{MSE0} \times 100\%$, or the difference between the reference model MSE0 and the new model MSE(x). The values are scaled to support the comparison over multiple predictors. The predictor is considered to be relatively important if the %IncMSE is higher.

While random forests do not have the same explanatory power as individual regression trees [90,92,93], they offer the advantages of limiting overfitting while minimizing errors due to bias. This is because an individual regression tree is sensitive to changes in the training data and splitting criteria, which can result in different tree structures and subsequent explanations [90]. Thus, results from both statistical methods are complementary and provide insight on model behavior as a function of the relative sensitivity of individual parameters and the effects of their joint interactions.

3. Results

3.1. Model Performance Assessments

A total of 3876 simulations were completed with the FLO-2D model, each representing the evaluation of a different combination of the seven input factors described in Table 2. Summary scalar variables were calculated to determine how closely simulations matched the observed debris flow event. For all completed simulations, the *F* scores ranged between 24.28% and 53.42%, which represent the lowest and the highest percent of binary agreement between the simulated and observed sediment deposition extent. The global root mean square error ranged between 0.60 and 1.16 m when considering only estimated reference heights associated with very high confidence (gRMSE1), and between 0.87 and 1.23 m when data points of all confidence levels were considered (gRMSE4). Based on this result, only gRMSE1 was subsequently used to rank the performance of simulation runs.

Tables 4 and 5 present the highest overall performing simulations ranked based on *F* and gRMSE1 scores, respectively. In general, the simulations with lower gRMSE1 scores are associated with a difference between simulated and reference sediment deposition heights of approximately 0.10 m lower than height differences associated with the top performing simulations ranked on the *F* score. However, the corresponding agreement between simulated and observed extents is 8–9% lower.

Furthermore, Table 6 presents the highest performing simulations from each group of rheological values that were evaluated in this study (Table 2) based on *F* scores. The ranges provide a first impression about model sensitivity to different parameter combinations, especially to definitions of rheology (i.e., input factors x_{1-4} and x_7).

In general, *F* scores were low (i.e., a maximum of 53% agreement matching the observed extent) and gRMSE1 values were high (i.e., >0.6 m difference between the simulated and observed heights). Summary scalar variables associated with the highest performing simulations for each type of rheology (i.e., A–F) are presented in Table 6. Figure 5 illustrates the simulated results that were generated based on the definition of two different rheologies (i.e., B and F) and input factor combinations. While there are certain similarities in terms of deposition extent (i.e., fairly close reproduction on the right side of the alluvial fan towards the mountains and under-prediction of extent in the bottom-left side), there are notable differences in the spatial distribution of sediment heights within the modelled extent. Additionally, the flow behavior that is defined in each simulation affects the flow velocities across the alluvial fan. The linkages among flow behavior, sediment deposition extent and heights, and flow velocity are further exemplified in the comparison of the two selected simulations. While the modelled output of B53 captures the sediment deposition heights relatively better than the output of F300, the maximum velocity is notably lower than that estimated at the Glyssibrücke (i.e., 2–4 m/s versus 6–10 m/s); the point-based simulated velocity of the F300 output matched that of the estimate provided by local experts. However, the agreement between simulated and estimated deposition heights is lower and a larger quantity of simulated debris flow materials was deposited into Lake Brienz.

Table 4. Overview of top 10 performing simulations based on highest *F* fitness measurement values.

		Input Factors for Model Calibration							Summary Scalar Variables I			
		x_1, x_2 : Yield Stress		x_3, x_4 : Viscosity		x_5 : Volumetric Sediment Concentration	x_6 : Specific Gravity	x_7 : Surface Detention	<i>F</i>	gRMSE1	gRMSE4	
Rank	Simulation Run ID	α_1	β_1	α_2	β_2				(%)	Very Certain Data Points (m)	All Data Points (m)	
1	B53	Aspen pit 1	0.181	25.7	0.036	22.1	0.45	2.8	0.80	53.42	0.71	0.92
2	B54	Aspen pit 1	0.181	25.7	0.036	22.1	0.45	2.8	1.00	53.34	0.72	0.92
3	B52	Aspen pit 1	0.181	25.7	0.036	22.1	0.45	2.8	0.60	53.22	0.71	0.92
4	E48	Chui-Sue River	1.2652	16.464	0.924	14.612	0.50	2.5	1.40	53.15	0.87	0.94
5	B34	Aspen pit 1	0.181	25.7	0.036	22.1	0.45	2.65	1.00	53.02	0.72	0.93
6	E103	Chui-Sue River	1.2652	16.464	0.924	14.612	0.50	2.65	1.40	52.99	0.87	0.94
7	B41	Aspen pit 1	0.181	25.7	0.036	22.1	0.45	2.8	0.50	52.98	0.71	0.92
8	E158	Chui-Sue River	1.2652	16.464	0.924	14.612	0.50	2.8	1.40	52.97	0.87	0.93
9	B33	Aspen pit 1	0.181	25.7	0.036	22.1	0.45	2.65	0.80	52.96	0.71	0.93
10	E49	Chui-Sue River	1.2652	16.464	0.924	14.612	0.50	2.5	1.50	52.95	0.84	0.92

Table 5. Overview of the top 10 performing simulations based on lowest gRMSE1 values calculated with very certain reference height data only.

		Input Factors for Model Calibration							Summary Scalar Variables I			
		x_1, x_2 : Yield Stress		x_3, x_4 : Viscosity		x_5 : Volumetric Sediment Concentration	x_6 : Specific Gravity	x_7 : Surface Detention	<i>F</i>	gRMSE1	gRMSE4	
Rank	Simulation Run ID	α_1	β_1	α_2	β_2				(%)	Very Certain Data Points (m)	All Data Points (m)	
1	E112	Chui-Sue River	1.2652	16.464	0.924	14.612	0.30	2.80	1.00	44.66	0.60	0.88
2	A6	Glenwood 4	0.00172	29.5	0.000602	33.1	0.35	2.50	1.00	42.71	0.60	0.90
3	E2	Chui-Sue River	1.2652	16.464	0.924	14.612	0.30	2.50	1.00	44.72	0.60	0.87
4	E57	Chui-Sue River	1.2652	16.464	0.924	14.612	0.30	2.65	1.00	44.81	0.60	0.87
5	B1	Aspen pit 1	0.181	25.7	0.036	22.1	0.30	2.50	0.80	42.84	0.61	0.89
6	B21	Aspen pit 1	0.181	25.7	0.036	22.1	0.30	2.65	0.80	42.69	0.61	0.89
7	B41	Aspen pit 1	0.181	25.7	0.036	22.1	0.30	2.80	0.80	42.58	0.61	0.89
8	A26	Glenwood 4	0.00172	29.5	0.000602	33.1	0.35	2.65	1.00	42.57	0.61	0.90
9	B26	Aspen pit 1	0.181	25.7	0.036	22.1	0.35	2.65	1.00	43.18	0.61	0.92
10	B46	Aspen pit 1	0.181	25.7	0.036	22.1	0.35	2.80	1.00	43.29	0.61	0.92

Table 6. Comparison of parameter combinations and summary scale variables for the top performing simulations runs for each of the six sets of rheologies obtained from published literature to back-calculate the 2005 debris flow in Brienz.

Model Inputs for Calibration									
Rank	Simulation Run ID		x_1, x_2 : Yield Stress		x_3, x_4 : Viscosity		x_5 : Volumetric Sediment Concentration	x_6 : Specific Gravity	x_7 : Surface Detention
			α_1	β_1	α_2	β_2			
1	B53	Aspen pit 1	0.181	25.7	0.036	22.1	0.45	2.80	0.80
2	E48	Chui-Sue River	1.2652	16.464	0.924	14.612	0.50	2.50	1.40
3	A55	Glenwood 4	0.00172	29.5	0.000602	33.1	0.45	2.80	1.50
4	D50	Rio Dona	0.05	22	0.0015	22	0.30	2.80	1.30
5	F300	Fella sx debris flow 2	0.0383	19.6	0.000495	27.1	0.30	2.65	1.30
6	C60	Scenario A	0.0013	23	2.83×10^{-5}	19	0.30	2.80	1.30
Summary Scalar Variables I									
Rank	Simulation Run ID		F	gRMSE1	gRMSE4	Adjusted r^2			Simulated Maximum Velocity (m/s)
			(%)	Very Certain Data Points (m)	All Data Points (m)	Very Certain Data Points	All Data		
1	B53	Aspen pit 1	53.42	0.71	0.92	0.24	0.07		2–4
2	E48	Chui-Sue River	53.15	0.87	0.94	0.08	0.04		0.5–1
3	A55	Glenwood 4	52.03	0.86	0.93	0.11	0.06		1–2
4	D50	Rio Dona	46.27	0.72	0.96	0.10	0.06		2–8
5	F300	Fella sx debris flow 2	46.11	0.72	0.95	0.11	0.06		6–8
6	C60	Scenario A	45.73	0.72	0.97	0.09	0.01		6–8

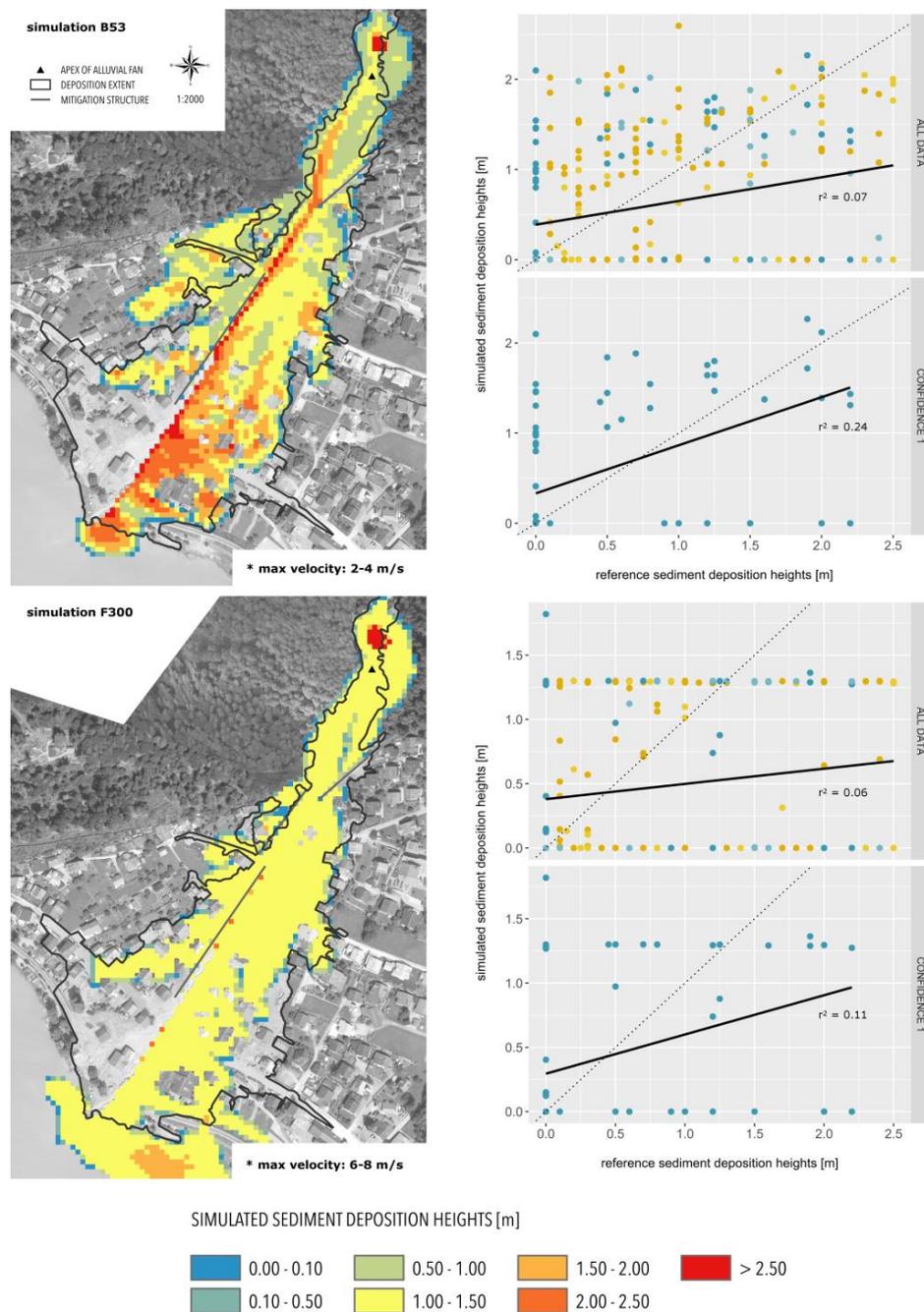


Figure 5. The two selected maps exemplify the agreement between simulated and observed sediment deposition extents, in addition to the spatial distribution of the simulated deposition heights for two of the best performing simulation runs; simulated maximum velocities provide a quick comparison with the point-based velocity estimate of 6–8 m/s that was determined by local experts at the Glyssibrücke.

Scatterplots with simple linear regression models compared the performance of simulated sediment deposition heights against all of the reference height data and a subset consisting of only very certain height data (Figure 5). The adjusted r^2 value [76] associated with using all of the data ranged from 0.01 to 0.07, while the adjusted r^2 value associated with the subset was relatively higher and ranged from 0.09 to 0.24 (Table 6). For the highest performing model output (i.e., B53), 24.57% of the variance between simulated and reference heights could be explained, an increase of about 17% from using all available observation heights associated with variable confidence levels.

As the model outputs did not reproduce the debris flow event in a highly satisfactory way (e.g., $F > 75\%$; $gRMSE < 0.5$ m), it is difficult to justify further inclusion of simulated results to develop physical vulnerability curves. However, the wide ranges of both F scores and $gRMSE$ values indicate that some parameter combinations did perform better than others.

3.2. Parameter Importance and Model Behaviour Assessments

To investigate model behavior as a function of the relative sensitivity of individual parameters and the effects of their joint interactions, two statistical methods described in Section 2.2 were applied.

Two were regression trees grown to support this study. The resulting tree models can change according to variable definitions of cp values; this describes the range of input factor combinations that generally produce simulations that have higher or lower agreements with observed data. For instance, in both of the regression trees, simulated outputs with higher degrees of agreement with observed data were produced by increasing the surface detention input factor from the default value of 0.03 m to a certain limit before performance no longer improves. In Figure 6, the blue circle indicates the cluster of simulation runs ($n = 21$) that is associated with the lowest averaged $gRMSE1$ values (0.68 m). These simulations have defined surface detention values < 1.4 m, a volumetric sediment concentration $< 47\%$, and rheological parameters where the viscosity coefficient $\beta_2 < 28$ and yield stress coefficient $\alpha_1 < 356 \times 10^6$. In Figure 7, the cluster of simulation runs ($n = 228$) indicated by the blue circle produced results with the highest averaged F scores (44% agreement between observed and simulated extents). These simulations were defined by surface detention values < 0.7 m, a volumetric sediment concentration $< 32\%$, and the viscosity coefficient $\beta_2 \geq 23$.

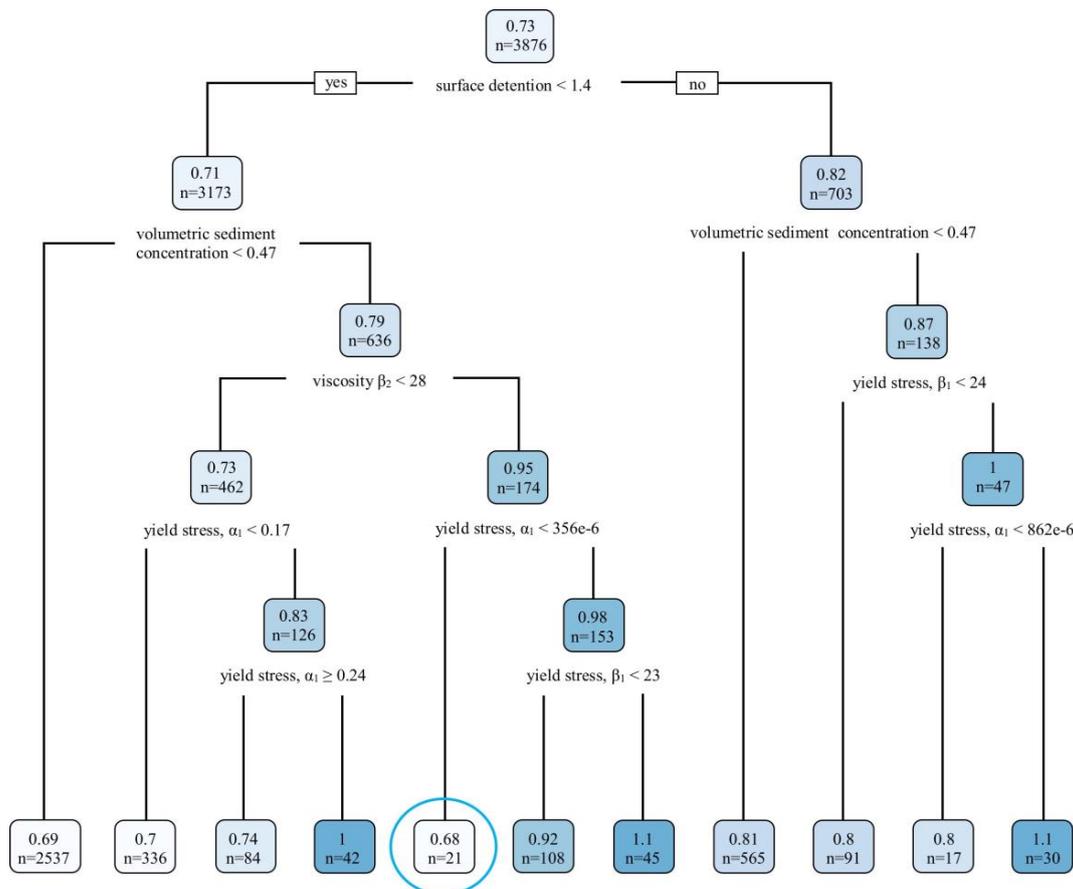


Figure 6. Regression tree based on average $gRMSE1$ values; $minsplit = 0$, $cp = 0.02$.

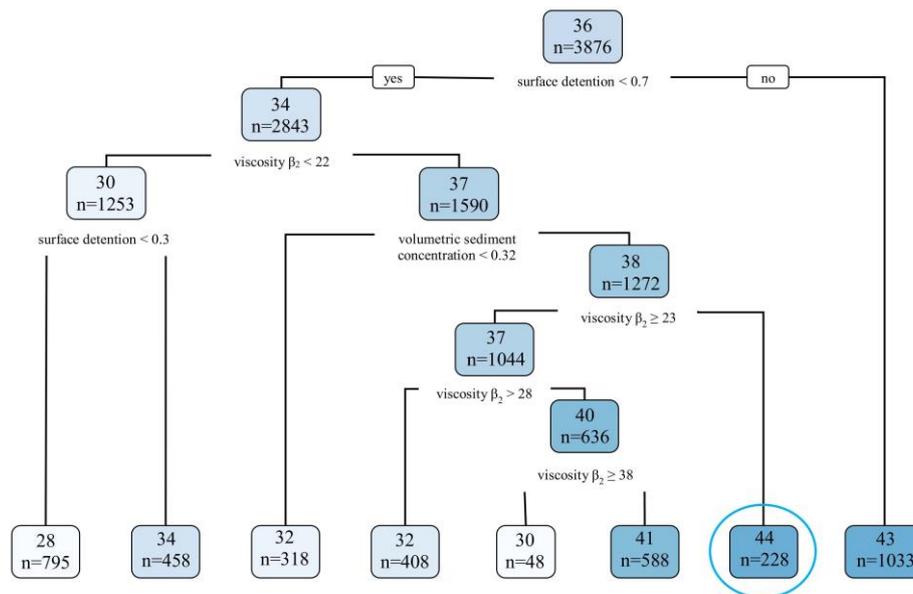


Figure 7. Regression tree based on averaged F statistic values; $cp = 0.02$.

Based on the results of the random forest model (Figure 8), the predictors within the FLO-2D model are identified and ranked. Relative parameter importance is determined based on highest %IncMSE. In this study, volumetric sediment concentration, surface detention, viscosity coefficients, and yield stress coefficients are the most important input factors in the FLO-2D model, respectively. Specific gravity is associated with a negative %IncMSE. This indicates that the randomly permuted predictor values performed better than the original values. Consequently, while the variability observed in Tables 4 and 5 show notable variability in the specific gravity values associated with top performing simulations, model outputs are insensitive to these changes and the predictor is not considered to play an important role in the model with respect to the contributions of other input factors.

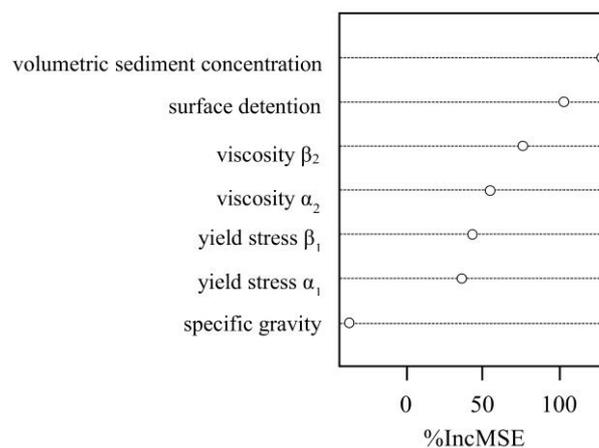


Figure 8. The results of the random forest analysis ranks model input factors based on importance.

4. Discussion

4.1. Sensitivity Analysis for Model Calibration

The utility of conducting SA for the calibration of a process model to reproduce a past event was assessed with a series of performance metrics. In natural hazard studies, observations or reference

data have inherent uncertainties attributed to a range of sources such as lack of standardized data collection and interpretation guidelines, pre-processing errors, and uncertainties introduced with spatial averaging. Furthermore, sources of uncertainties also exist with respect to model limitations and working assumptions. With the acknowledgement of these uncertainties, absolute corroboration or rejection of a model on the basis of performance metrics alone has its limitations. Nevertheless, the results of a sensitivity analysis provide useful guidance on the identification of parameter importance on model behavior and effectively reduces the amount of parameter space to explore.

Based on the results presented in Tables 3 and 4, it can be observed that the accurate definition of rheological parameters is the basis to capturing representative flow behavior. Additional input factors such as surface detention and volumetric sediment concentration were identified by both regression trees and the random forest model to have a notable effect on improving the overall agreement between simulated and reference data through further calibration. In particular, sediment detention values effectively modify rheological properties by introducing a height threshold to each computational grid cell, so that the debris flow materials are retained for a longer time before further propagating downwards as the detention height threshold is exceeded. In effect, the surface detention value modifies the flow viscosity. Volumetric sediment concentration values are directly used to generate mud hydrographs, which influence the definition of flow characteristics. The results of the sensitivity analysis helped to refine reasonable range of values for the two input factors for further calibration, highlight the importance of rheological properties, in addition to showing that variations in specific gravity did not result in notable changes in model outputs. Through the investigation of parameter importance and their joint interactions, we can continue to build a knowledge base for a more efficient locally-based calibration within sub-regions of the parameter space that may be of interest [59]. With respect to processing modelling past torrent events with FLO-2D, it is recommended that in situ sediment samples be collected to represent flow behavior as accurately as possible.

4.2. Model Performance Assessment: Potentials and Limitations

Several working assumptions should be considered to determine if the FLO-2D model is suited to replicating torrent event and the potential limitations if it is applied. Firstly, certain gravity-driven geophysical flows that comprise of solid particles within a fluid (e.g., snow avalanches, lava, mud and debris flows) are assumed to behave as relatively homogeneous, single-phase flows at a macroscopic scale [28]. In reality, the materials that comprise the matrix may be more accurately described as multi-phased. Furthermore, in the case of Brienz, the matrix was also capable of transporting boulders between 3 and 5 m in diameter. Since particles within granular flows experience frictional and collisional interactions, discrete element modeling is needed to account for these types of interactions in order to effectively describe flow behaviors [81]. The treatment of such geophysical flows as a single-phase phenomenon effectively reduces the complexity of the problem [94,95]. Consequently, both pore pressure and inter-grain frictional effects are considered to be negligible and bulk rheological properties (i.e., constant dynamic viscosity and yield stress) are assumed to apply to the entire matrix in FLO-2D. Furthermore, the theory and working assumption of viscoplasticity may not be valid for describing all geophysical flows, where Coulomb plasticity or other models may describe the behavior more accurately [96,97]. In the study conducted by Bertolo and Bottino [98], the authors expressed that given the complexity of geophysical flows, multiple rheological models may be necessary to capture the full range of characteristics of a specific debris flow. This recommendation also applies to different events occurring at the same location.

Under the assumption of homogeneity, the geophysical flow modelled in FLO-2D should primarily be comprised of fine materials. Ideally, laboratory-based assessments of collected field samples can then be conducted to infer rheological properties about the materials [60]. However, field samples of the torrent event of interest may not be available. For these cases, it may be possible to simulate an event using rheological values cited in literature from materials with comparable mineralogical compositions and/or grain size distributions (e.g., [69]).

Model results provide insight about the feasibility of simulating past natural hazard events with limited field-based model input data. The study focused on determining whether ranges of FLO-2D input factor values could be adequately estimated from available data and whether model calibration could generate simulated results with a strong agreement with the reference data. In particular, the definition of flow behavior based on rheological values from literature highlighted the limitations of this approach. For this study, where the exact grain size distribution is unknown and the estimated range of the volumetric concentration is wide (i.e., 30–70%), the event was simulated based on assumptions of homogeneity. Rheological parameter values obtained from literature varied in terms of their mineralogical compositions. Furthermore, it is important to note that flow behavior within a torrent can change within the duration of a given event. Consequently, the defined range of parameters captures generalized flow physics. In lieu of additional information to constrain available rheological values in literature, those that were associated with comparable grain size distributions to the study site were favored.

In a study by Sosio et al. [99], shear stress was observed to vary for flows comprised of greater amounts of solid content at relatively higher volumetric concentrations ($C_v = 45\text{--}63\%$). This variability has been attributed to the local effects of air bubbles or the formation of grain clusters in the mixture, which violate the continuum assumption for mud matrices characterized by a higher C_v . This could provide some explanation for the lower agreement of simulated and observed extents when using literature-based values corresponding to materials from torrent events characterized by a lower C_v to calibrate for higher C_v values (i.e., 45–70%) in this study.

Additionally, O'Brien and Julien [84] observed that for mudflow matrices with < 20% volumetric sand concentrations, the viscosity corresponds with that of the silt-clay fraction of the mixture. This relationship is important since rheological analysis to determine viscosity and yield stress values is only conducted with the fine fraction of a given mixture. The specific maximum threshold of grain size that is included in a given rheological analysis is independently determined by each working group. For example, lab-based rheological analyses were conducted based on grain sizes < 0.425 mm and < 0.063 mm from field samples, in studies conducted by Sosio et al. [67] and Boniello et al. [70], respectively. This highlights potential limitations with determining the rheological properties of a complex event, such as the 2005 debris flow, based only on the defined fine fraction of the associated grain size distribution. Uncertainty is further introduced when rheological values derived with this reductionist approach are applied to calibrate a different study site such as Brienz. Furthermore, considering the working assumptions of homogeneity while reviewing the range of materials that are identifiable in post-event photographs, it is clear that the impact of grain sizes beyond the defined fine fraction and larger pieces of debris within the mud matrix cannot be accounted for with this specific model. However, simulated results are generated based on the consistent application of rules. Consequently, outputs will have a consistent degree of error attributed to problem oversimplification or the effects of spatial averaging. This is in contrast to variable, often human errors that can be introduced from data collection and subjective interpretation stages, especially when multiple parties and participants are involved.

The FLO-2D model respects mass conservation. Consequently, over-prediction of the sediment deposition extent is related to the consistent under-prediction of sediment deposition heights. Additionally, the accurate definition of the flow behavior through the rheological parameters characterizes the velocity at which the flow propagates downwards and the manner in which it spreads laterally across the alluvial fan. Lower performing simulations were generally characterized by an over-prediction of the extent, low deposition heights and high velocities attributed to flows with low viscosity (Figure 5, simulation F300). On the contrary, relatively higher performing simulations were generally characterized by a combination of under- and over-prediction of the extent, relatively higher deposition heights and low velocities attributed to more viscous flows (Figure 5, simulation B53).

The results of this study illustrate that while the visual agreement between simulated and observed deposition extents may look promising, quantifying the degree of agreement with a three-dimensional approach provides a clearer picture about the feasibility of reproducing a past event under the aforementioned constraints. In particular, this addresses past observations that while process models often require calibration, many simulated results have yet to be comprehensively evaluated against field events [32]. In the case of reproducing past torrent events, additional agreement between simulated outputs and reference deposition heights and/or available flow velocities is necessary to determine the accuracy of the defined flow behavior. The current study effectively extends the analysis initiated by past studies that were conducted with the consideration of only a limited number of sediment deposition heights (i.e., $n < 20$). It is a prerequisite to accurately define flow behavior to be able to generate representative intensity proxies (e.g., velocities and impact pressures) via process modeling. Only then can the inclusion of accurately simulated intensity proxies contribute to the derivation of physical vulnerability curves for consequence analysis.

5. Conclusions

Based on the analysis results to date, sources of uncertainties from the model, input data and reference data limit further inclusion of simulated intensity proxies from this study for further development of physical vulnerability curves. Nevertheless, the proposed method and tools support the quantitative representation of aggregated uncertainties related to data quality, parameterization and model suitability. In light of these findings, the following recommendations may be of interest to researchers and practitioners in the natural hazards and risk community:

The findings of the sensitivity analysis demonstrated how the most influential input factors within a given model can be identified. This provides guidance on setting priorities for future data collection and process modeling efforts. In particular, the definition of flow behavior is a key prerequisite to obtaining more highly accurate simulated results and is predicated upon acquiring representative post-event field samples. Collecting field samples of materials to determine rheology in the lab generates much higher agreement in modelled results than the use of rheology reported in literature from other study sites. Additionally, Tiranti and Deangeli [100] presented a method of interest that predicts probable rheologies of alpine debris flows, based on the availability of source area data such as lithology.

Exploration of alternative methods to determine deposition heights in a consistent way to reduce uncertainties about validation data accuracy (i.e., reduce the range of confidence levels) is imperative before a more accurate understanding of process and model behavior is possible. For instance, advances in satellite, especially after 2015, enables researchers to produce highly accurate and up to date digital terrain models that reflect the landscape characteristics well, especially with shorter temporal return periods where images captured immediately following event occurrence become more readily available. Differences in the landscape before and after a hazard event can be calculated from a pre-event and post-event image acquired immediately over the affected area. Cavalli et al. [101] presented a method to detect geomorphic changes in mountain catchments based on increasingly available high-resolution DEM data. Alternatively, the difference between pre- and post-event orthophotographs containing height information can also be used to calculate the sediment deposition height across an area of interest.

Exploration of alternative process models with different working assumptions and approaches and consideration of models that can be calibrated automatically to support the efficient exploration of larger parameter spaces is instrumental. Models with the latter characteristic can reach optimal solutions under reduced computational times.

In summary, the conclusions of this study are twofold. Firstly, the results support previous findings (e.g., [102]) that the main limitation of the FLO-2D model is predicated on the use of simplified Bingham equations to represent complex debris flow physics. In doing so, the granular flow of heterogeneous materials such as debris and rock fragments within an interstitial mud [103] is reduced

to a single-phased, homogenous flow. While the model may effectively support other types of investigations, it is not suitable for back calculating past hazardous events, where highly accurate simulated intensity outputs are required for further analysis. Secondly, the study demonstrated the utility of SA to focus efforts on areas of parameter space associated with higher model performance, rather than evaluating all potential parameter combinations. It is especially useful to have large numbers of simulation runs completed under a shorter amount of time when exploring a complex problem characterized by a large parameter space. The application of statistical tools can support data exploration to gain insights from modeled results. Finally, results from SA provide guidance for future data collection efforts based on the identification of important parameters and parameter combinations for a given model.

Author Contributions: Conceptualization, C.C. and J.R.; Formal analysis, C.C.; Funding acquisition, M.K.; Methodology, C.C.; Project administration, M.K.; Scripting, C.C. and J.R.; Supervision, M.K.; Validation, C.C.; Visualization, C.C.; Writing—original draft, C.C.; Writing—review & editing, C.C., J.R., and M.K.

Funding: This study was funded by the Swiss National Science Foundation (SNSF number 159899).

Acknowledgments: Post-event photographs volunteered by affected residents in Brienz were compiled by Heinz Pfiffner from the Kantonale Verwaltung Amt für Wald (KAWA) in 2005. Markus Zimmermann provided insight on pre- and post-event conditions in Brienz used to define initial conditions and parameter values for process modeling with FLO-2D.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Eisebacher, G. Mountain torrents and debris flows. *Episodes* **1982**, *4*, 12–17.
2. Varnes, D. Slope movement types and processes. In *Landslides: Analysis and Control, TRB Special Report 176: Transportation Research Board*; National Academy of Sciences: Washington, DC, USA, 1978; pp. 11–33.
3. Hungr, O.; Evans, S.; Bovis, M.; Hutchinson, J. A review of the classification of landslides of the flow type. *Environ. Eng. Geosci.* **2001**, *7*, 221–238. [[CrossRef](#)]
4. Costa, J. Physical Geomorphology of Debris Flows. In *Developments and Applications of Geomorphology*; Springer: Berlin, Germany, 1984; pp. 268–317.
5. Takahashi, T.; Sawada, T.; Suwa, H.; Mizuyama, T.; Mizuhara, K.; Wu, J.; Tang, B.; Kang, Z.; Zhou, B. Japan-China joint research on the prevention from debris flow hazards. In *Research Report*; Japanese Ministry of Education, Science and Culture, International Scientific Research Program No. 03044085; Japanese Ministry of Education: Tokyo, Japan, 1994.
6. Wan, Z.; Wang, Z. Hyperconcentrated flow. In *IAHR Monograph Series*; Balkema: Rotterdam, The Netherlands, 1994.
7. Scheidl, C.; Rickenmann, D. Empirical Prediction of Debris-Flow Mobility and Deposition on Fans. *Earth Surf. Process. Landf.* **2010**, *35*, 157–173. [[CrossRef](#)]
8. Jakob, M.; Hungr, O. *Debris-flow Hazards and Related Phenomena*; Praxis Publishing Limited: Chichester, UK, 2005; p. 14.
9. Bradley, J.; McCutcheon, S. The effects of high sediment concentration on transport processes and flow phenomena. In *International Symposium on Erosion, Debris Flow and Disaster Prevention*; Toshindo: Tsukuba, Japan, 1985; pp. 219–225.
10. Pappathoma-Köhle, M.; Kappes, M.; Keiler, M.; Glade, T. Physical vulnerability assessment for alpine hazards: State of the art and future needs. *Nat. Hazards* **2011**, *58*, 645–680. [[CrossRef](#)]
11. Totschnig, R.; Fuchs, S. Mountain torrents: Quantifying vulnerability and assessing uncertainties. *Eng. Geol.* **2013**, *155*, 31–44. [[CrossRef](#)] [[PubMed](#)]
12. Mazzorana, B.; Simoni, S.; Scherer, C.; Gems, B.; Fuchs, S.; Keiler, M. A physical approach on flood risk vulnerability of buildings. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 3817–3836. [[CrossRef](#)]
13. Pappathoma-Köhle, M.; Zischg, A.; Fuchs, S.; Glade, T.; Keiler, M. Loss estimation for landslides in mountain areas—An integrated toolbox for vulnerability assessment and damage documentation. *Environ. Model. Softw.* **2015**, *63*, 156–169. [[CrossRef](#)]

14. Hübl, J.; Kienholz, H.; Loipersberger, A. DOMODIS—Documentation of Mountain Disasters. Available online: http://www.baunat.boku.ac.at/fileadmin/data/H03000/H87000/H87100/Downloads/Domodis_Deutsch.pdf (accessed on 14 June 2018).
15. 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.* **2008**, *8*, 1113–1127. [[CrossRef](#)]
16. Quan Luna, B.; Blahut, J.; van Westen, C.; Sterlacchini, S.; van Asch, T.; Akbas, S. The application of numerical debris modelling for the generation of physical vulnerability curves. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 2047–2060. [[CrossRef](#)]
17. Jakob, M.; Stein, D.; Ulmi, M. Vulnerability of buildings to debris flow impact. *Nat. Hazards* **2012**, *60*, 241–261. [[CrossRef](#)]
18. Scheidegger, A. On the prediction of the reach and velocity of catastrophic landslides. *Rock Mech.* **1973**, *5*, 231–236. [[CrossRef](#)]
19. Körner, H. Reichweite und Geschwindigkeit von Bergstürzen und Fliesslawinen. *Rock Mech.* **1976**, *8*, 225–256. [[CrossRef](#)]
20. Rickenmann, D. Empirical relationships for debris flows. *Nat. Hazards* **1999**, *19*, 47–77. [[CrossRef](#)]
21. Legros, F. The mobility of long runout landslides. *Eng. Geol.* **2002**, *63*, 301–331. [[CrossRef](#)]
22. Berti, M.; Simoni, A. Prediction of debris flow inundation areas using empirical mobility relationships. *Geomorphology* **2007**, *90*, 144–161. [[CrossRef](#)]
23. Magirl, C.; Griffiths, P.; Webb, R. Analyzing debris flows with the statistically calibrated empirical model LAHARZ in southeastern Arizona, USA. *Geomorphology* **2010**, *119*, 111–124. [[CrossRef](#)]
24. Scheidl, C.; Rickenmann, D. TopFlowDf—A simple GIS based model to simulate debris-flow runout on the fan. In Proceedings of the 5th International Conference on Debris-Flow Hazards: Mitigation, Mechanics, Prediction and Assessment, Padua, Italy, 14–17 June 2011; pp. 253–262.
25. Gruber, S. A mass-conserving fast algorithm to parameterize gravitational transport and deposition using digital elevation models. *Water Resour. Res.* **2007**, *43*, W06412. [[CrossRef](#)]
26. Takahashi, T. *Debris Flows: Mechanics, Prediction and Countermeasures*; CRC Press: Rotterdam, The Netherlands, 1991.
27. O'Brien, J.; Julien, P.; Fullerton, W. Two-dimensional water flood and mudflow simulation. *J. Hydraul. Eng.* **1993**, *119*, 244–261. [[CrossRef](#)]
28. Hungr, O. A model for the runout analysis of rapid flow slide, debris flow, and avalanches. *Can. Geotech. J.* **1995**, *32*, 610–623. [[CrossRef](#)]
29. Brufau, P.; Garcia-Navarro, P.; Ghilardi, R.; Natale, L.; Savi, F. 1-D mathematical modelling of debris flow. *J. Hydraul. Res.* **2000**, *38*, 435–446. [[CrossRef](#)]
30. Medina, V.; Hurlimann, M.; Bateman, A. Application of FLATModel, a 2D finite volume code, to debris flows in the Northeastern part of Iberian Peninsula Experimental evidences and numerical modelling of debris flow initiated by channel runoff. *Landslide* **2008**, *5*, 127–142. [[CrossRef](#)]
31. Liu, K.; Wang, M. Numerical modelling of debris-flow with application on hazard area mapping. *Comput. Geosci.* **2006**, *10*, 221–240. [[CrossRef](#)]
32. Rickenmann, D.; Laigle, D.; McArdell, B.M.; Hübl, J. Comparison of 2D debris-flow simulation models with field events. *Comput. Geosci.* **2006**, *10*, 241–264. [[CrossRef](#)]
33. Armanini, A.; Fraccarollo, L.; Rosatti, G. Two-dimensional simulation of debris flows in erodible channels. *Comput. Geosci.* **2009**, *35*, 993–1006. [[CrossRef](#)]
34. Begueira, S.; Van Asch, T.; Malet, J.; Grondahl, S. A GIS-based numerical model for simulating the kinematics of mud and debris flows over complex terrain. *Nat. Hazards Earth Syst. Sci.* **2009**, *9*, 1897–1909. [[CrossRef](#)]
35. Christen, M.; Bartelt, P.; Kowalski, J. Back calculation of the In den Arelen avalanche with RAMMS: Interpretation of model results. *Ann. Glaciol.* **2010**, *51*, 161–168. [[CrossRef](#)]
36. Christen, M.; Kowalski, J.; Bartelt, P. RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. *Cold Reg. Sci. Technol.* **2010**, *63*, 1–14. [[CrossRef](#)]
37. Pirulli, M. On the use of the calibration-based approach for debris-flow forward-analyses. *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 1009–1019. [[CrossRef](#)]
38. De Blasio, F.; Elverhøia, A.; Issler, D.; Harbitz, C.; Brync, P.; Lienc, R. Flow models of natural debris flows originating from overconsolidated clay materials. *Mar. Geol.* **2004**, *213*, 439–455. [[CrossRef](#)]

39. Mergili, M.; Schratz, K.; Ostermann, L.; Fellin, W. Physically-based modelling of granular flows with Open Source GISNat. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 187–200. [[CrossRef](#)]
40. Pudasaini, S. A general two-phase debris flow model. *J. Geophys. Res.* **2012**, *117*, F03010. [[CrossRef](#)]
41. Rosatti, G.; Begnudelli, L. Two dimensional simulations of debris flows over mobile beds: Enhancing the TRENT2D model by using a well-balanced generalized Roe-type solver. *Comput. Fluids* **2013**, *71*, 179–185. [[CrossRef](#)]
42. McDougall, S.; Hungr, O. A model for the analysis of rapid landslide motion across three-dimensional terrain. *Can. Geotech. J.* **2004**, *41*, 1084–1097. [[CrossRef](#)]
43. Pastor, M.; Haddad, B.; Sorbino, G.; Cuomo, S.; Demptric, V. A depth integrated-coupled SPH model for flow like landslides and related phenomena. *Int. J. Numer. Anal. Methods Geomech.* **2008**, *33*, 143–172. [[CrossRef](#)]
44. Deangeli, C.; Segre, E. Cellular automaton for realistic modelling of landslides. *Nonlinear Process. Geophys.* **1995**, *2*, 1–15.
45. D'Ambrosio, D.; Di Gregorio, S.; Iovine, G.; Lupiano, V.; Rongo, R.; Spataro, W. First simulations of the Sarno debris flows through Cellular Automata modelling. *Geomorphology* **2003**, *54*, 91–117. [[CrossRef](#)]
46. Deangeli, C. Laboratory granular flows generated by slope failures. *Rock Mech. Rock Eng.* **2008**, *41*, 199–217. [[CrossRef](#)]
47. Gregoretti, C.; Degetto, M.; Boreggio, M. GIS-based cell model for simulating debris flow runout on a fan. *J. Hydrol.* **2016**, *534*, 326–340. [[CrossRef](#)]
48. Schraml, K.; Thomschitz, B.; McArdell, B.; Kaitna, R. Modeling debris-flow runout patterns on two alpine fans with different dynamic simulation models. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 1483–1492. [[CrossRef](#)]
49. 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.* **2006**, *6*, 155–165. [[CrossRef](#)]
50. Rickenmann, D. Debris-flow hazard assessment and methods applied in engineering practice. *Int. J. Eros. Control Eng.* **2016**, *9*, 80–90. [[CrossRef](#)]
51. Mannschatz, T.; Dietrich, P. Chapter 2—Model Input Data Uncertainty and Its Potential Impact on Soil Properties. In *Sensitivity Analysis in Earth Observation Modelling*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 25–52.
52. Hürlimann, M.; Rickenmann, D.; Medina, V.; Bateman, A. Evaluation of approaches to calculate debris-flow parameters for hazard assessment. *Eng. Geol.* **2008**, *102*, 152–163. [[CrossRef](#)]
53. Hussin, H.; Quan Luna, B.; Van Westen, C.; Christen, M.; Malet, J.-P.; van Asch, T. Parameterization of a numerical 2-D debris flow model with entrainment: A case study of the Faucon catchment, Southern French Alps. *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 3075–3090. [[CrossRef](#)]
54. D'Aniello, A.; Cozzolino, L.; Cimorelli, L.; Della Morte, R.; Pianese, D. A numerical model for the simulation of debris flow triggering, propagation and arrest. *Nat. Hazards* **2015**, *75*, 1403–1433. [[CrossRef](#)]
55. Iverson, R. The physics of debris flows. *Rev. Geophys.* **1997**, *35*, 245–296. [[CrossRef](#)]
56. McDougall, S.; Hungr, O. *Modelling of landslides which entrain material from the path*, *Can. Geotech. J.* **2005**, *42*, 1437–1448.
57. Frank, F.; McArdell, B.; Huggel, C.; Vieli, A. The importance of entrainment and bulking on debris flow runout modeling: Examples from the Swiss Alps. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 2569–2583. [[CrossRef](#)]
58. Bertolo, P.; Wieczorek, G. Calibration of numerical models for small debris flows in Yosemite Valley, California, USA. *Nat. Hazards Earth Syst. Sci.* **2005**, *5*, 993–1001. [[CrossRef](#)]
59. Pianosi, F.; Beven, K.; Freer, J.; Hall, J.; Rougier, J.; Stephenson, D.; Wagener, T. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environ. Model. Softw.* **2016**, *79*, 214–232. [[CrossRef](#)]
60. Song, X.; Zhang, J.; Zhan, C.; Xuan, Y.; Ye, M.; Xu, C. Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical frameworks. *J. Hydrol.* **2015**, *523*, 739–757. [[CrossRef](#)]
61. Beven, K.; Aspinall, W.; Bates, P.; Borgomeo, E.; Goda, K.; Hall, J.; Page, T.; Phillips, J.; Simpson, M.; Smith, P.; et al. Epistemic uncertainties and natural hazard risk assessment. 2. What should constitute good practice? *Nat. Hazards Earth Syst. Sci. Discuss.* **2017**. [[CrossRef](#)]
62. Rougier, J.; Sparks, S.; Hill, L. *Risk and Uncertainty Assessment for Natural Hazards*; University Press: Cambridge, UK, 2013.

63. Stancanelli, L.; Foti, E. A comparative assessment of two different debris flow propagation approaches—Blind simulations on a real debris flow event. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 735–746. [CrossRef]
64. Von Boetticher, A.; Turowski, J.; McArdell, B.; Rickenmann, D.; Kirchner, J. DebrisInterMixing-2.3: A finite volume solver for three-dimensional debris-flow simulations with two calibration parameters—Part 1: Model description. *Geosci. Model Dev.* **2016**, *9*, 2909–2923. [CrossRef]
65. Huang, Y.; Qin, X. Uncertainty analysis for flood inundation modelling with random floodplain roughness field. *Environ. Syst. Res.* **2014**, *3*, 9. [CrossRef]
66. Swisstopo. swissALTI^{3D}. Available online: https://shop.swisstopo.admin.ch/en/products/height_models/alti3D (accessed on 14 June 2018).
67. Sosio, R.; Crosta, G.; Frattini, P. Field observations, rheological testing and numerical modelling of a debris-flow event. *Earth Surf. Process. Landf.* **2007**, *32*, 290–306. [CrossRef]
68. Lin, M.; Wang, K.; Huang, J. Debris flow run off simulation and verification? Case study of Chen-You-Lan Watershed, Taiwan. *Nat. Hazards Earth Syst. Sci.* **2005**, *5*, 439–445. [CrossRef]
69. D’Agostino, V.; Tecca, P. Some considerations on the application of FLO-2D model for debris flow hazard assessment. *WIT Trans. Ecol. Environ.* **2006**, *90*, 159–170.
70. Boniello, M.; Calligaris, C.; Lapasin, R.; Zini, L. Rheological investigation and simulation of a debris-flow event in the Fella watershed. *Nat. Hazards Earth Sci.* **2010**, *10*, 989–997. [CrossRef]
71. NDR Consulting Zimmermann/Niederer + Pozzi Umwelt AG. *Lokale Lösungsorientierte Ereignisanalyse Glyssibach. Bericht zum Vorprojekt*; Tiefbauamt des Kantons Bern: Thun, Switzerland, 2006.
72. Swisstopo. Orthophotos. Available online: https://shop.swisstopo.admin.ch/en/products/images/ortho_images/SWISSIMAGE_RS (accessed on 14 June 2018).
73. Chow, V. *Open-Channel Hydraulics*; McGraw-Hill: New York, NY, USA, 1959.
74. Ziliani, L.; Surian, N.; Coultard, T.; Tarantola, S. Reduced-complexity modeling of braided rivers: Assessing model performance by sensitivity analysis, calibration, and validation. *J. Geophys. Res. Earth Surf.* **2013**, *118*, 2243–2262. [CrossRef]
75. Skinner, C.; Coulthard, T.; Schwanghart, W.; Van De Wiel, M.; Hancock, G. Global Sensitivity Analysis of Parameter Uncertainty in Landscape Evolution Models. *Geosci. Model Dev. Discuss.* **2017**. [CrossRef]
76. Yin, P.; Fan, X. Estimating R² shrinkage in multiple regression: A comparison of different analytical methods. *J. Exp. Educ.* **2001**, *69*, 203–224. [CrossRef]
77. Sikuli/SikuliX Documentation for version 1.1+ (2014 and Later). Available online: <http://www.sikuli.org/> (accessed on 18 June 2016).
78. FLO-2D. FLO-2D Reference Manual. Available online: <http://www.flo-2d.com/wp-content/uploads/2014/04/Manuals-Pro.zip> (accessed on 11 February 2016).
79. 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.* **2010**, *10*, 535–545. [CrossRef]
80. Arcement, G.; Schneider, V. Guide for Selecting Manning’s Roughness Coefficients for Natural Channels and Flood Plains. In *United States Geological Survey Water-Supply Paper 2339*; United States Government Printing Office: Washington, DC, USA, 1989.
81. Cagnoli, B.; Piersanti, A. Combined effects of grain size, flow volume and channel width on geophysical flow mobility: Three-dimensional discrete element modeling of dry and dense flows of angular rock fragments. *Solid Earth* **2017**, *8*, 177–188. [CrossRef]
82. Di Santolo, A.; Pellegrino, A.; Evangelista, A. Experimental study on the rheological behaviour of debris flow. *Nat. Hazards Earth Syst. Sci.* **2010**, *10*, 2507–2514. [CrossRef]
83. O’Brien, J.; Julien, P. Laboratory analysis of mudflow properties. *J. Hydraul. Eng.* **1988**, *114*, 877–887. [CrossRef]
84. Major, J.; Pierson, T. Debris flow rheology: Experimental analysis of fine-grained slurries. *Water Resour. Res.* **1992**, *28*, 841–857. [CrossRef]
85. Coussot, P.; Piau, J. The effects of an addition of force-free particles on the rheological properties of fine suspensions. *Can. Geotech. J.* **1995**, *32*, 263–270. [CrossRef]
86. FLO-2D. Data Input Manual PRO. Available online: <http://www.flo-2d.com/wp-content/uploads/2014/04/Manuals-Pro.zip> (accessed on 11 February 2016).
87. Bates, P.; De Roo, A. A simple raster-based model for flood inundation simulation. *J. Hydrol.* **2000**, *236*, 54–77. [CrossRef]

88. Pappenberger, F.; Iorgulescu, I.; Beven, K. Sensitivity analysis based on regional splits and regression trees (SARS-RT). *Environ. Model. Softw.* **2006**, *21*, 976–990. [[CrossRef](#)]
89. Hofer, E. Sensitivity analysis in the context of uncertainty analysis for computationally intensive models. *Comput. Phys. Commun.* **1999**, *117*, 21–34. [[CrossRef](#)]
90. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
91. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
92. Breiman, L. Statistical modeling: The two cultures. *Stat. Sci.* **2001**, *16*, 199–215. [[CrossRef](#)]
93. Iorgulescu, I.; Beven, K. Nonparametric direct mapping of rainfall-runoff relationships: An alternative approach to data analysis and modeling? *Water Resour. Res.* **2004**, *40*, W08403. [[CrossRef](#)]
94. Ancey, C.; Meunier, M. Estimating bulk rheological properties of flowing snow avalanches from field data. *J. Geophys. Res.* **2004**, *109*. [[CrossRef](#)]
95. Sosio, R.; Crosta, G. Rheology of concentrated granular suspensions and possible implications for debris flow modeling. *Water Resour. Res.* **2009**, *45*. [[CrossRef](#)]
96. Ancey, C. Plasticity and geophysical flows: A review. *J. Non-Newton. Fluid Mech.* **2007**, *142*, 4–35. [[CrossRef](#)]
97. Iverson, R. The debris flow rheology myth. In *Third International Conference on Debris-Flow Hazard Mitigation: Mechanics; Prediction and Assessment*, American Society of Civil Engineers: Reston, VA, USA, 2003.
98. Bertolo, P.; Bottino, G. Debris-flow event in the Frangerello Stream-Susa Valley (Italy)—Calibration of numerical models for the back analysis of the 16 October, 2000 rainstorm. *Landslides* **2008**, *5*, 19–30. [[CrossRef](#)]
99. Sosio, R.; Crosta, G.; Frattini, P.; Valbuzzi, E. Caratterizzazione reologica e modellazione numerica di un debris flow in ambiente alpino. *G. Geol. Appl.* **2006**, *3*, 263–268.
100. Tiranti, D.; Deangeli, C. Modeling of debris flow depositional patterns according to the catchments and source areas characteristics. *Front. Earth Sci.* **2015**, *3*, 1–14. [[CrossRef](#)]
101. Cavalli, M.; Goldin, B.; Comiti, F.; Brardinoni, F.; Marchi, L. Assessment of erosion and deposition in steep mountain basins by differencing sequential digital terrain models. *Geomorphology* **2017**, *291*, 4–16. [[CrossRef](#)]
102. Iverson, R.; Denlinger, R. Flow of variably fluidized granular masses across three-dimensional terrain 1. Coulomb mixture theory. *J. Geophys. Res.* **2001**, *106*, 537–552. [[CrossRef](#)]
103. McCoy, S.; Tucker, G.; Kean, J.; Coe, J. Field measurement of basal forces generated by erosive debris flows. *J. Geophys. Resour. Earth Surf.* **2013**, *118*, 589–602. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).