Uncertainty analysis of the estimation of stony debris flow rainfall threshold: the application to the Backward Dynamical Approach

Marta Martinengo[1], Daniel Zugliani[1], and Giorgio Rosatti[1]

[1]Department of Civil, Environmental and Mechanical Engineering, University of Trento, Trento, Italy

Abstract. Rainfall thresholds, namely rainfall intensity-duration conditions beyond which the probability of debris flow occurrence is considered significant, can be used as a forecasting tool in debris-flow early warning system. Many uncertainties may affect the thresholds calibration and, in turn, the reliability and effectiveness of this tool. The purpose of this study is to assess the uncertainty in the determination of the rainfall threshold for stony debris flow based on the Back Dynamical Approach (BDA) (Rosatti et al., 2019), an innovative method to estimate the rainfall duration and averaged intensity strictly related to measured debris flow. The uncertainty analysis has been computed performing two Monte Carlo cascade simulations: (i) to assess the variability in the estimate of rainfall conditions due to the uncertainty of some of the BDA parameters and (ii) to quantify the impact of this variability on the threshold parameters, obtained by using the frequentist method. Then, the deviation between these analysis outcomes and the values obtained in Rosatti et al. (2019) has been examined. The results highlight that the variability in the rainfall condition estimate is strongly related to the debris flow characteristics and the hyetograph shape. Depending on these features, the spreading of the obtained distributions can take both low and high values. Instead, the threshold parameters are characterised by a low statistical spreading. Finally, the consistency between the outcome of this study and the results obtained in Rosatti et al. (2019) has been proved and the critical issues related to the rainfall condition estimation have been discussed.

1 Introduction

Debris flows are very intense phenomena that affect mountain regions and have a significant impact on the territory in which they occur, causing damages and, in some cases, casualties (Fuchs et al. (2013), Cánovas et al. (2016)). For this reason, debris flow risk management, based on both active and passive mitigation strategies, is crucial to reduce the effects of the phenomenon on the territory. An early warning system is an example of a passive mitigation tool (Huebl and Fiebiger, 2005) as it allows to activate prevention measures (e.g. evacuation sets out in the civil protection plans) before the expected event occurs.

In this last framework, the forecast of the possible occurrence of debris flow is mainly based on rainfall thresholds (Chien-Yuan et al. (2005), Segoni et al. (2018)), namely rainfall conditions beyond which the probability of debris flow occurrence is considered significant. Usually, rainfall thresholds are power laws that link the rainfall duration to the rainfall cumulated or intensity (Nikolopoulos et al., 2014) and the relevant coefficients are calibrated on historical data. A considerable literature deals with this topic (e.g. Aleotti (2004), Guzzetti et al. (2007), Jakob et al. (2012), Pan et al. (2018)).

One of the critical issues of the rainfall threshold calibration is the uncertainty related to both data and models parameters (Gariano et al., 2019). Here with the term "model", we indicate generically a set of operations that, given some input data and model parameters, provide an output. The uncertainties derive mainly from data error measurements (e.g. in rainfall) and from the estimation of the parameters related to the chosen models used to prepare the data and to calibrate the threshold. The result

30 is an uncertainty framework that can significantly impact the threshold estimation resulting in a not reliable forecasting tool. In this context, uncertainty analysis is required. In general, this type of analysis allows to assess and control the output variability of a given model due to input uncertainties (Helton et al. (2006), Marino et al. (2008), Pisoni et al. (2018)) and to describe the degree of goodness of the results (Coleman and Steele, 2018). In the present case, the model is the threshold estimation and the relevant uncertainty outcomes provide useful information concerning the reliability and robustness of the threshold.

35 This work aims to assess the uncertainty related to the debris-flow threshold estimation based on the Backward Dynamical Approach (BDA), introduced by Rosatti et al. (2019). Given a rainfall event related to a debris flow occurrence, the BDA allows to quantify the rainfall volume strictly pertaining to the surveyed debris flow. The rainfall volume is defined as the amount of water needed to convey downstream as a mixture, the volume of sediments deposited by the debris flow. It is estimated on the basis of a schematic description of debris flows dynamics. Once the rainfall volume is calculated, the relevant rainfall

40 conditions (duration D and averaged intensity I) determining that volume is defined based on the event hyetograph. Finally, the threshold calibration is performed applying the frequentist method (Brunetti et al. (2010) and Peruccacci et al. (2012)) to the (I, D) set, obtained considering all the known and well documented events in a homogeneous area.

In literature, there are some studies related, for example, to the variability of calibrated threshold due to the uncertainty in rainfall estimation (e.g. Nikolopoulos et al. (2014), Rossi et al. (2017)) and in rainfall temporal resolution (Marra, 2019). On

45 the contrary, this work focuses on the analysis of how the uncertainty in some input parameters affects the BDA outputs and threshold estimation and not on the rainfall.

The uncertainty analysis can be performed using different strategies and methods (e.g. Helton et al. (2006), Coleman and Steele (2018), Hofer (2018)). In this paper, we have chosen the Monte Carlo (MC) approach. It consists in the production of a large set of model outputs, obtained picking the model inputs within given ranges. Then, the statistical analysis of the

50 output set gives the required uncertainty estimation. In this work, two MC cascade simulations have been performed to carry out the uncertainty analysis. A first MC application is used to assess the uncertainty in the BDA outputs estimation. Then, the effects of these outputs variability on threshold calibration are quantified with a further MC computation. Finally, the results are compared with the reference values obtained in Rosatti et al. (2019) to assess their consistency.

The paper structure is the following. A brief description of the BDA method and the calibrated threshold obtained by Rosatti

55 et al. (2019) for a study area is presented in Sect. 2. The method used to assess the uncertainty propagation is described in Sect. 3. The results of the application of the method obtained are presented in Sect. 4. In Sect. 5, discussion and conclusion end the paper.

2 The Backward Dynamical Approach

The Backward Dynamical Approach is an innovative method proposed by Rosatti et al. (2019) that defines the rainfall volume
strictly relevant to a stony debris flow in a physical-based way. As described in Rosatti et al. (2019), the BDA starts from the
knowledge of the deposited volume V_{dep} occupied by the sediments, surveyed after a debris flow event. Thanks to a simplified
volumetric description of the phenomenon dynamic (Fig.1), the rainfall volume pertaining to the debris flow V_r^{DF}, defined as
the volume of water necessary to convey downstream as a mixture the volume of sediments V_{dep}, can be express as a function
of this latter volume:

$$V_r^{DF} = \frac{c_b - c}{c} V_{dep} \tag{1}$$

where c_b is the concentration of the sediment in the bed, constant and assumed equal to 0.65 (Takahashi, 2014), and c is a
reference volumetric solid concentration of the given debris flow. According to Takahashi (1978), c can be estimated as

$$c = \min\left(\frac{i_f}{\Delta(\tan\psi - i_f)}, 0.9\,c_b\right) \tag{2}$$

where i_f is the average slope of the last 50 m of the river bed before the deposition zone, ψ is the dynamic friction angle of the
sediments and $\Delta = (\rho^s - \rho^l)/\rho^l$ is the sediment relative submerged density where ρ^l and ρ^s are, respectively, the liquid and
solid constant density.

V_r^{DF} can also be express as product between the rainfall volume per unit area E and the event basin area A_b:

$$V_r^{DF} = E A_b \tag{3}$$

from which, a dynamical expression for the rainfall volume per unit area can be obtained:

$$E = \frac{1}{A_b}\frac{c_b - c}{c} V_{dep} \tag{4}$$

On the other hand, E can be obtained from the forcing of the phenomenon, namely in the hyetograph. Under the assumption
of uniform rainfall over the basin, the hydrological expression for E is:

$$E = \int_{t_1}^{t_2} i(t)\,dt \tag{5}$$

where $i(t)$ is the measured rainfall intensity, t_1 and t_2 are the start and end times related to the debris-flow duration. In the
absence of event detailed data, the instant of maximum intensity during the event t_{max} is assumed to be the debris flow
triggering time (Iadanza et al., 2016) and t_1 and t_2 are assumed to be:

$$\begin{cases} t_1 = t_{max} - \Delta t_1 \\ t_2 = t_{max} + \Delta t_2 \end{cases} \tag{6}$$

Because of the measurement technique, $i(t)$ is a piecewise constant function on time intervals δt. Therefore, the integral in
Eq. (5) cannot equal exactly the value of E, and an approximated value must be considered. The values of $\Delta t_1 = n_1 \delta t$ and
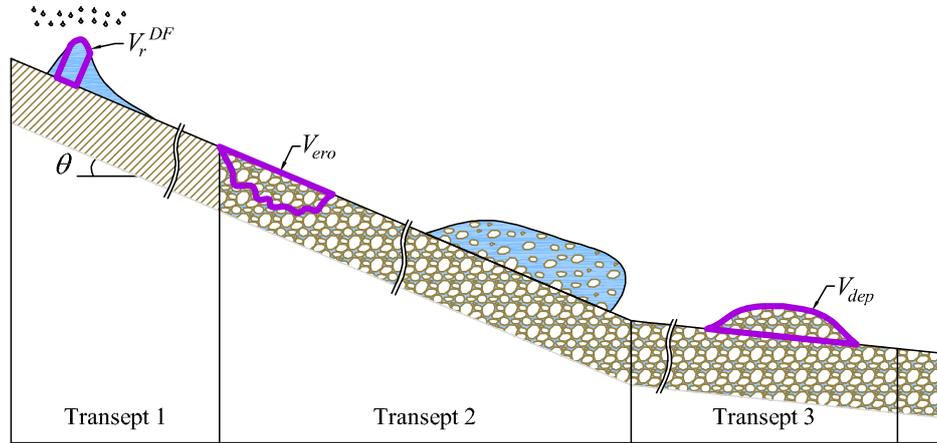
Figure 1. Conceptual Lagrangian volumetric description of debris flow dynamic from Rosatti et al. (2019). The scheme is divided into three transepts: transept 1 is characterised by the runoff formation; the bed material erosion and the achievement of equilibrium conditions occur in transept 2; transept 3 is characterised by the deposition of sediments with water entrapment. V_r^{DF} is the rain volume pertaining to the debris flow, V_{ero} is the bed volume variation related to the erosion, V_{dep} is the deposited volume occupied by the sediments and θ is the inclination angle of the bed with respect to a reference horizontal direction.

$\Delta t_2 = n_2 \delta t$ are computed considering the minimum number of intervals (i.e. n_1 and n_2) necessary to exceed (or equal) the value of E:

$$\min \left(\int_{t_1}^{t_2} i(t)\,dt \right) \geq E \tag{7}$$

For further details on the computation of Δt_1 and Δt_2, please refer to Rosatti et al. (2019).

Therefore, the duration D and the average intensity I related to the event become

$$D = \Delta t_1 + \Delta t_2 \tag{8}$$

$$I = \frac{\min \left(\int_{t_1}^{t_2} i(t)\,dt \right)}{D} \tag{9}$$

2.1 The BDA-based threshold for a study area

As described in Rosatti et al. (2019), the BDA methodology has been applied to obtain a stony debris flow rainfall threshold for the Trentino-Alto Adige/Südtirol region (Italy) (Fig. 2). A dataset composed of 84 debris flow events has been considered to calibrate the threshold. Firstly, the (I, D) couples have been computed applying the BDA method. Then, the threshold has been obtained by using the frequentist method (e.g. Brunetti et al. (2010) and Peruccacci et al. (2012)). According to this method,

Figure 2. Location of Trentino-Alto Adige/Südtirol region (Italy), the study area used in Rosatti et al. (2019).

the threshold is a straight line in the log-log ID plane parallel to the power law

$$I = \hat{a} D^{-b} \tag{10}$$

obtained fitting the events rainfall conditions. The threshold exponent is therefore equal to b while the intercept is computed by setting a value of the non-exceedance probability of the dataset events, namely by imposing the occurrence probability of debris flows related to rainfall conditions located below the threshold. The threshold equation is then:

$$I = a D^{-b} \tag{11}$$

in which $a < \hat{a}$. For more details on the frequentist method, we refer the reader to the above-mentioned references.

In Rosatti et al. (2019) the non-exceedance level has been set equal to 5% and the resulting calibrated threshold is:

$$I = 6.2 D^{-0.67} \tag{12}$$

3 Method

As briefly presented in the introduction, to perform a comprehensive study, the uncertainty analysis of the BDA-base threshold estimation is divided into three parts. A first analysis examines the uncertainty propagation in the computation of the BDA

Table 1. Uncertainty probability distributions of the input parameters established to perform the LHS. $i_{f,r}$, $A_{b,r}$ and $V_{dep,r}$ are the reference values (i.e. the values used in the calibration phase) of the average slope, of the basin area and of the sediments respectively.

Input parameter/data	Probability distribution function
i_f	$\text{Uniform}(i_{f,r}(1-5\%\sqrt{3}), i_{f,r}(1+5\%\sqrt{3}))$
A_b	$\text{Uniform}(A_{b,r}(1-5\%\sqrt{3}), A_{b,r}(1+5\%\sqrt{3}))$
V_{dep}	$\text{Uniform}(V_{dep,r}(1-5\%\sqrt{3}), V_{dep,r}(1+5\%\sqrt{3}))$
ψ	$\text{Uniform}(32°, 38°) \sim \text{Uniform}(35°(1-5\%\sqrt{3}), 35°(1+5\%\sqrt{3}))$

outputs (i.e. (I, D) couple, Eq. (8) and (9), and related values of c, Eq. (2) and E, Eq. (4)) starting from the uncertainty of some input parameters. Subsequently, the effects of this propagation on the frequentist method threshold estimation are analysed. Finally, a comparison between the results of the two previous analyses and the reference outcomes, obtained in Rosatti et al. (2019), is carried out. In the following sections, we present the details of each part.

3.1 BDA outputs uncertainty analysis

The main sources of uncertainties in the BDA outputs are attributed to some of the input parameters and data, namely the average slope i_f, the basin area A_b, the deposited volume V_{dep} and the dynamical friction angle ψ. As described in Rosatti et al. (2019), during the calibration phase, i_f and A_b related to each event have been estimated performing GIS analysis based on the available debris flow data (e.g. location of the deposited area, technical reports...) and V_{dep} have been provided by regional agencies. Instead, due to the scarcity of sediments information, ψ has been assumed to be 35° for all debris flow events. As stressed in the Introduction, the rainfall intensities $i(t)$ associated with the event are assumed to be certain. Future analysis will assess and study also the uncertainties related to this piece of data.

For each event, the uncertainty propagation of i_f, A_b, V_{dep} and ψ is carried out with MC approach using the Latin Hypercube Sampling (LHS) procedure (McKay et al. (2000), Helton and Davis (2003) and Helton et al. (2006)) for the generation of the sample. The implementation of this procedure is schematized in Fig. 3 and is composed of three main steps. First, the characterization of the uncertainty in the parameters, namely the probability distribution function (pdf) of their values, has to be defined (Fig. 3(a)). Lacking certain data concerning the pdfs, according to Marino et al. (2008), all the input parameters are assumed to be uniformly distributed. In particular, for each event, the means of the parameter distributions are set equal to the reference values, namely to the values used in the calibration phase (indicated with subscript r). The dynamical friction angle is the only parameter with a validity range. For stony debris flow, according to Lane (1953) and Blijenberg (1995), it varies from 32° to 38°. Therefore, assuming 35° as the mean of the ψ distribution, this range can be obtained by imposing a variation coefficient CV (i.e. the ratio between the standard deviation and the mean) equal to about 5%. Since being greater than zero is the only constraint of the other parameters, for homogeneity this CV value is considered suitable for all parameters. The uncertainties characterization of the parameters is summarized in Table 1.

Second, the parameter samples, namely the ordered sets of parameters values in the form $(i_f, A_b, V_{dep}, \psi)$, must be obtained.
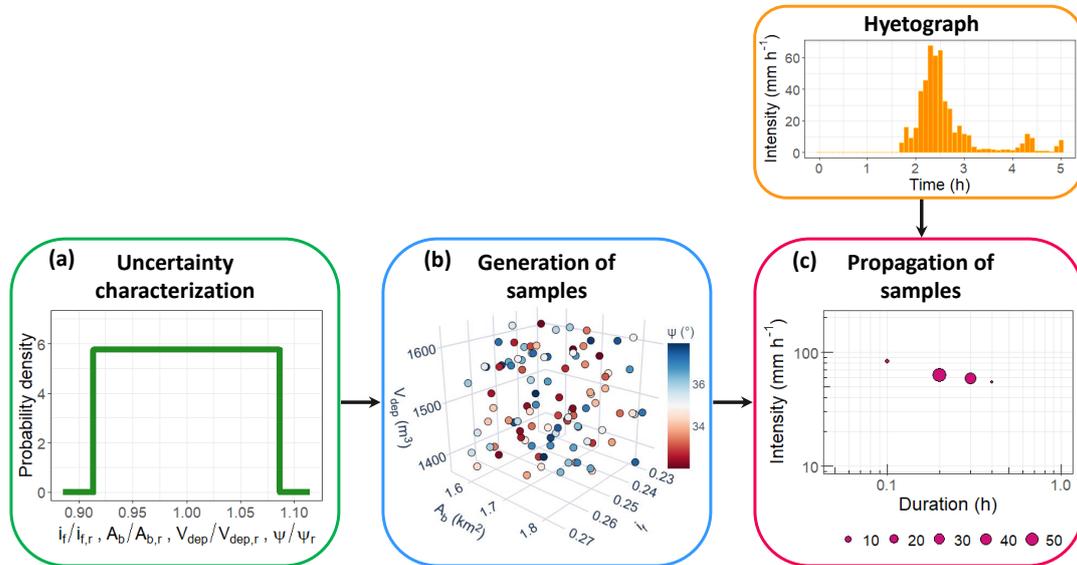
Figure 3. Scheme of the uncertainty analysis performed with Monte Carlo and Latin Hypercube Sampling method for the (I, D) couples computation. For each event, the procedure consists of three steps. (a) Uncertainty characterization of the input parameters: the nondimensional form of the input parameters, obtained dividing the parameters by the reference value used in the calibration phase and indicated with a subscript r, are assumed to be uniformly distributed. (b) Samples generation performing the LHS. (c) Propagation of samples through the BDA method introducing the event hyetograph. The dots size in the log-log ID plane indicates the absolute frequency of obtaining the (I, D) couples.

These samples are generated by using the LHS (Fig. 3(b)), introduced by McKay et al. (2000). This method produces N samples starting with a division of each parameter range into N disjoint intervals of equal probability. Then, one value is randomly selected within every interval, thus obtaining N values for each parameter. These values are then arranged in a matrix composed of N rows and k columns, where k is the number of the parameters (four, in the specific case). In each

140 column, the N values relevant to a single parameter are inserted in random order (Helton et al., 2006). Each row of this matrix gives one of the N parameter samples. According to Marino et al. (2008), to ensure the accuracy, the sample size N should be at least greater than k. In this study, N is set to 100 and the (100×4) LHS matrix is generated for each event, based on the previously established pdfs.

Finally, the BDA outputs are obtained starting from each parameter sample, resulting in 100 (I, D) couples for each event (Fig.

145 3(c)), together with the related concentrations c and rainfall volumes per unit area E.

The obtained results are then analysed in term of relative uncertainty, quantified through the computation of the CV of the output distributions. The CV, by definition, is a standardized measure of uncertainty (Håkanson, 2000) and allows to compare the relative variability of the results independently of their measurement units (Abdi, 2010) and of their means. For this reason, the CV is chosen as statistical measure to compare the relative variability between the outputs (between both the same output

150 of different events and different outputs of the same event) and to understand how the uncertainty of an output variable changes depending on the event characteristics. In particular, the CVs are computed for all the BDA outputs.

Moreover, to evaluate how the (I, D) points move in the ID plane, the absolute uncertainty associated with the (I, D) couples estimation is quantified throughout the computation of the 95% uncertainty intervals. The extremes of the latter are defined as the 2.5 and 97.5 percentiles of the D and I distributions. The length of the uncertainty interval allows evaluating the

155 absolute variability of the D and I distributions.

3.2 Threshold uncertainty analysis

To figure out the effects of the BDA outputs variability on the threshold computation, namely on the two constants a and b of Eq. (11), the uncertainties propagation is estimated with a further Monte Carlo procedure. In this case, 5000 samples are generated selecting randomly one of the one hundred (I, D) couples previously obtained for each event. The result is 5000

160 (a, b) couples, i.e. 5000 different thresholds.

Consistently with the BDA output analysis, the relative uncertainty associated with the a and b distributions is quantified through the computation of the CVs. In addition, to analyse the threshold absolute uncertainty for each combination of threshold parameters, the intensity values are calculated for fixed durations selected within the threshold relevant time interval, from five minutes to six hours with a five minutes time step. In this way, for each duration we obtain an intensity frequency distribu-

165 tion composed of 5000 samples. Then, the 2.5 and 97.5 percentiles of these distributions are chosen as upper and lower bounds of the threshold uncertainty.

3.3 Comparison between MC outputs and reference values

In this analysis, we compare the MC means (I, D) couples, the threshold obtained with the mean values of a and b and the threshold uncertainty bounds with the results of the original calibration phase (reference values) in order to assess the

170 differences between the two approaches.

As regards the rainfall conditions, according to Marra (2019), the bias of duration B_D and intensity B_I are computed for each event as:

$$B_D = \frac{D_m}{D_r} \qquad B_I = \frac{I_m}{I_r} \qquad (13)$$

where the subscripts r and m represent respectively the reference value and the mean of the MC output distribution.

175 For what concerns the threshold, firstly the differences between the MC intensities $I_{MC,k}$, where k stands for mean, upper and lower bounds, and the reference threshold ones $I_{t,r}$ (Eq. (12)) are carried out for the same fixed durations used to define the uncertainty bounds:

$$Diff = I_{MC,k} - I_{t,r} \qquad (14)$$

Subsequently, the percentage changes, defined as:

180 $$\% \, change = \left(\frac{I_{MC,k} - I_{t,r}}{I_{t,r}} \right) \cdot 100 \qquad (15)$$

are computed to figure out how much the MC results deviate relatively from the calibrated threshold.

4 Results

4.1 Propagation of uncertainty in the BDA outputs

The method described in the previous section has been applied to the 84 debris flow events used to calibrate the threshold for
the Trentino-Alto Adige/Südtirol region (Italy) in Rosatti et al. (2019). The values of the CVs for the BDA outputs related to
each event are shown in Table 2. The analysis of their variability allows to make the following observations:

- D is the output characterised by distributions with the largest and most variable CVs with respect to all the other outputs.
  In fact, the events CV_D vary between 0% and 157.5%. The reason for this behaviour will be clarified further on;

- the distributions of I are characterised by a lower spread with respect to the D distributions, being all the CV_I values
  within 0% and to 30.0%. The reason of this behaviour is connected to the fact that I, by definition, is an average and
  therefore the effects of the input parameters variability are smoothed by the mean operator. Also the reason why CV_I
  can be zero will be explained further on;

- also the concentration distributions show a low spread. As c is characterised by an upper bound (i.e. $0.9\,c_b$), the CV_c
  is strictly related to the proximity of c_r to this maximum value and consequently, according to Eq. (2), to the value of
  $i_{f,r}$. As shown in Fig. 4, until the $i_{f,r}$ is less than about 0.3, the CV_c tends to go up by increasing the $i_{f,r}$ since c_r is
  sufficiently smaller than $0.9\,c_b$. Instead, if the $i_{f,r}$ is between about 0.3 and 0.4, the CV_c tends to decrease by increasing
  the $i_{f,r}$ as c reaches the maximum and it is equal to $0.9\,c_b$ for an increasing number of i_f samples. Finally, the CV_c
  becomes 0% if the $i_{f,r}$ is grater than about 0.4 as c is always equal to $0.9\,c_b$ independently from the values of the i_f
  samples;

- the volume per unit area distributions shows CV_Es that vary between 7.1% and 64.6%. It is worth noting that high
  uncertainty in the estimation of E does not necessarily imply large CV_D and/or CV_I (e.g. event 12) and vice versa (e.g.
  event 38).

To better understand the variability of D and I, we classify the events into three categories based on the CV_D values:

1. events with zero variability: $CV_D = 0\%$;

2. events with low variability: $0\% < CV_D \leq 30\%$;

3. events with high variability: $CV_D > 30\%$.

The first category comprises 48 events for which the 100 simulations of the MC method has provided as final outputs always
the same (I, D) couple. For these events, the propagation of the input parameters uncertainties does not effect the (I, D) couple
estimation resulting in $CV_D = CV_I = 0\%$. This type of result is related to two conditions:

Table 2. Coefficients of variation of the outputs of the BDA method related to each event expressed as a percentage. CV_D is the coefficient of variation of the duration distribution, CV_I of the intensity distribution, CV_c of the concentration distribution and CV_E of the rainfall volume per unit area distribution.

Event	CV_D	CV_I	CV_c	CV_E	Event	CV_D	CV_I	CV_c	CV_E	Event	CV_D	CV_I	CV_c	CV_E
1	0.0	0.0	0.0	7.1	29	0.0	0.0	12.6	26.6	57	0.0	0.0	10.6	15.6
2	34.6	5.5	11.6	18.3	30	0.0	0.0	12.5	26.9	58	157.5	14.9	14.1	35.4
3	34.2	7.9	13.3	30.7	31	0.0	0.0	0.0	7.1	59	0.0	0.0	10.6	15.9
4	23.0	6.2	11.1	17.8	32	0.0	0.0	11.3	18.4	60	0.0	0.0	0.4	8.8
5	7.1	1.5	9.1	12.6	33	13.8	4.1	11.2	17.2	61	16.1	4.4	9.5	13.4
6	0.0	0.0	0.0	7.1	34	30.9	9.7	13.1	28.1	62	0.0	0.0	11.1	18.1
7	39.6	13.5	13.8	36.3	35	24.8	22.7	9.8	14.2	63	0.0	0.0	7.0	55.5
8	26.4	2.8	12.0	20.9	36	0.0	0.0	7.9	58.4	64	7.8	1.0	0.0	7.1
9	0.0	0.0	0.0	7.1	37	0.0	0.0	3.6	34.2	65	0.0	0.0	14.0	43.5
10	26.6	15.7	10.3	14.7	38	74.7	30.0	10.1	14.4	66	0.0	0.0	10.0	14.4
11	0.0	0.0	12.9	27.0	39	0.0	0.0	0.0	7.1	67	110.1	28.2	13.6	35.5
12	0.0	0.0	11.6	55.9	40	0.0	0.0	0.0	7.1	68	32.9	4.6	13.8	35.9
13	0.0	0.0	0.0	7.1	41	35.3	15.1	12.4	24.7	69	37.1	12.3	13.9	36.2
14	15.6	2.6	9.5	13.2	42	0.0	0.0	13.3	30.0	70	0.0	0.0	13.7	36.1
15	0.0	0.0	11.2	17.8	43	0.0	0.0	0.0	7.1	71	0.0	0.0	6.6	54.8
16	0.0	0.0	0.0	7.1	44	0.0	0.0	1.0	15.0	72	0.0	0.0	12.9	54.7
17	0.0	0.0	13.3	34.6	45	25.3	10.0	12.5	22.8	73	16.0	0.6	12.0	21.6
18	0.0	0.0	12.7	25.0	46	0.0	0.0	0.0	7.1	74	0.0	0.0	13.2	29.5
19	13.3	3.0	10.4	15.1	47	16.7	4.9	9.0	12.1	75	0.0	0.0	13.2	31.4
20	41.7	4.7	7.1	54.9	48	0.0	0.0	10.8	16.4	76	37.7	18.9	8.9	12.3
21	0.0	0.0	8.8	11.7	49	15.9	4.3	9.7	13.1	77	31.1	9.4	9.5	13.1
22	0.0	0.0	11.0	17.8	50	39.9	11.0	14.3	42.4	78	0.0	0.0	14.2	38.7
23	0.0	0.0	9.3	13.0	51	25.1	8.6	11.2	18.9	79	0.0	0.0	13.0	26.4
24	17.3	5.7	9.5	13.3	52	0.0	0.0	9.6	63.1	80	25.2	9.4	13.4	31.1
25	0.0	0.0	0.0	7.1	53	35.4	1.6	10.1	14.5	81	0.0	0.0	9.8	64.6
26	62.4	16.9	14.1	44.8	54	0.0	0.0	0.0	7.2	82	0.0	0.0	0.7	10.1
27	72.1	25.7	12.9	26.3	55	22.6	10.6	9.3	12.7	83	0.0	0.0	11.4	18.4
28	0.0	0.0	12.9	26.2	56	0.0	0.0	0.0	7.1	84	16.1	1.0	9.4	13.2
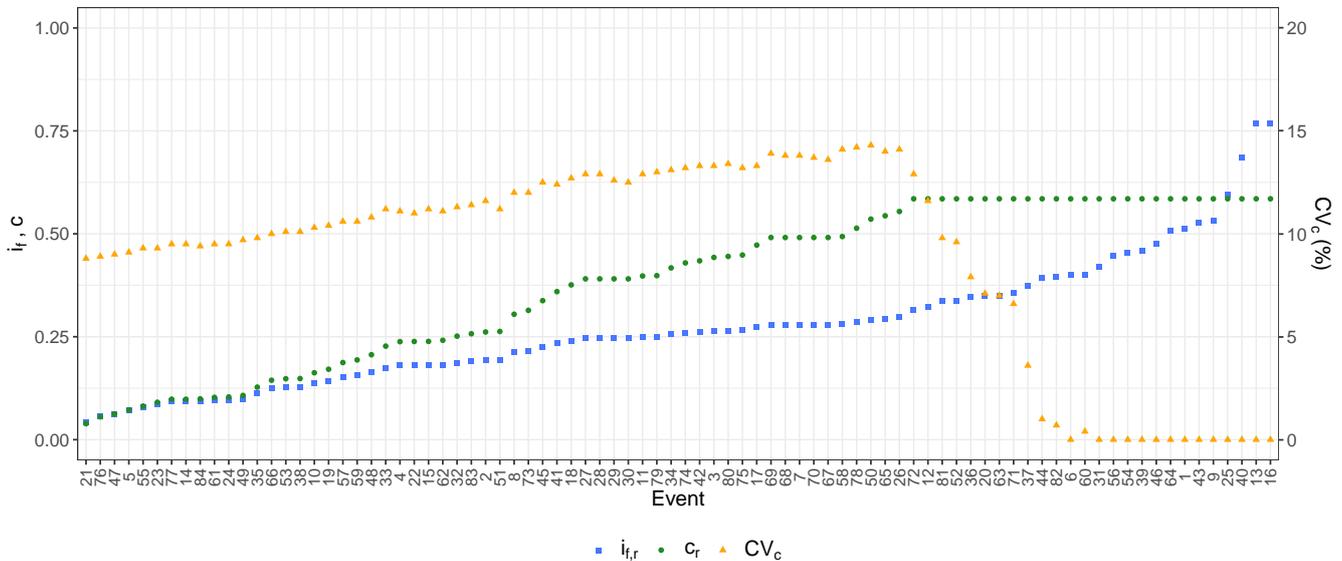
Figure 4. Comparison of the reference slope $i_{f,r}$, reference concentration c_r and CV_c trends. The events have been sorted with decreasing slopes.

210   – regardless of the input parameters variability, the concentration is always equal to $0.9\,c_b$. In this case also the variation coefficient of the concentration is equal to zero (e.g. events 1, 13 and 25) and the variation of E is only due to the propagation of the basin area and deposited volume uncertainties (Eq. (4)), namely $CV_E \simeq 7.1\%$. For these 14 events, such a low variation in E implies the constant computation of the same (I, D) couple;

   – despite CV_c is not zero and the CV_E is greater than 7.1% (e.g. event 28), the condition of Eq. (7) is satisfied, in all the

215   100 simulations, with the same time instants of the hyetograph. 34 events fall into this condition.

The events that belong to the second category are characterised by I and D distributions relatively low spread around their mean values. For each of these 19 events, the uncertainty of the inputs implies the computation of more (I, D) couples which, however, are relatively close to the mean value.

The third category includes 17 events for which the input parameters uncertainties propagation results in high relative uncer-

220 tainty in the (I, D) couples computations. In general, these events are related to a hyetograph characterised by an intensity peak and low values around it. This sharp decrease in intensity implies that small variations in E require large variations in the time intervals in order to satisfy Eq. (7). Event 58 is the extreme case of this condition: the duration distribution of this event has an extreme much greater than the mean value (Fig. 5). This is due to the presence of zero intensity temporal instants in the middle of the precipitation event that must be considered to reach the higher values of E. For this reason, for some combinations of

225 the input parameters, D is much greater than the mean value and these extremes entail an increase in the standard deviation, namely in the CV_D. The effects of this condition on I are fewer thanks to the mean computation carried out to obtain this

Table 3. Mean, standard deviation, variation coefficient CV and mean 95% confidence interval CI of the parameter a and b computed performing the Monte Carlo simulations.

Parameter	Mean	Standard deviation	CV (%)	95% mean confidence interval
a	6.0056	0.3882	6.46	0.0108
b	0.6834	0.0199	2.91	0.0006

output.
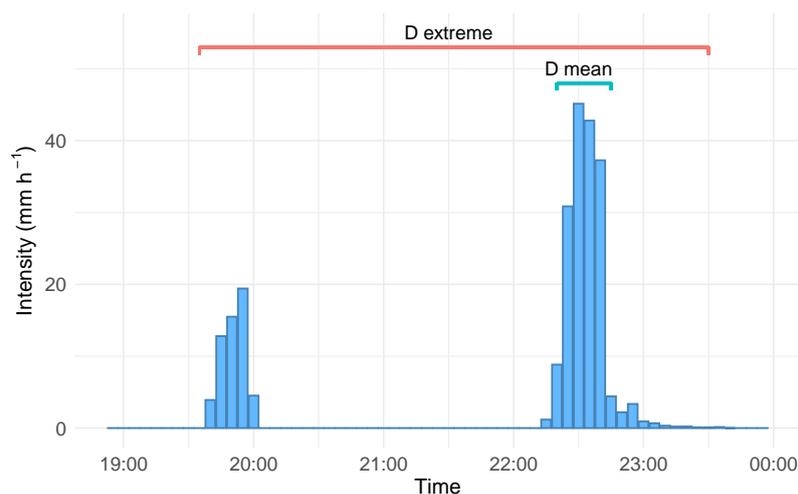
Figure 5. Hyetograph of event 58. In term of D, the extreme value (3.75 h) is much greater than the mean (0.32 h).

As regards the absolute uncertainty associated with the (I, D) couples estimation, the results are shown in Fig. 6: the

230 horizontal and vertical lines represent respectively the 95% uncertainty intervals related to D and I. As evident, the length of the intervals varies greatly depending on the event. In term of intensity, the maximum length is 61.3 mm h^{-1} and is reached with the event 67 while, for the duration, the event 38 is characterised by the maximum length that is equal to 3.51 h.

4.2 Uncertainty in threshold computation

The main statistical quantities concerning the output a and b parameters distributions are given in Table 3. In particular, the

235 variation coefficients are respectively 6.46% and 2.91%. These low values highlight that the relative uncertainties associate to a and b are very small. The low spread nature of the thresholds parameters is also evident in the scatter plot and in the bivariate $3D$ histogram, respectively shown in Fig. 7(a) and 7(b).

Moreover, the lower and upper bounds of the threshold uncertainty, evaluated as explained in Sect. 3.2, are shown in Fig. 8. The uncertainty area is substantially symmetric to the threshold computed with the mean values of the distributions of a and b
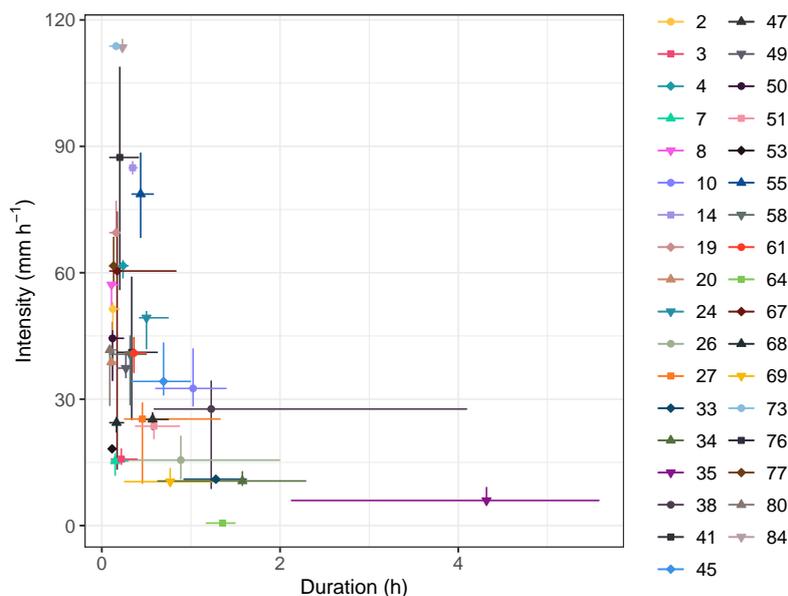
Figure 6. Plot of the uncertainty in the (I, D) evaluation obtained with the Monte Carlo simulations. The dots are the mean values and the horizontal and vertical lines are respectively the duration and intensity absolute uncertainties (i.e. the 2.5 and 97.5 percentiles). To make the graph clearer, the points with uncertainty intervals equal to zero have not been represented and the linear scale is used for both axis.
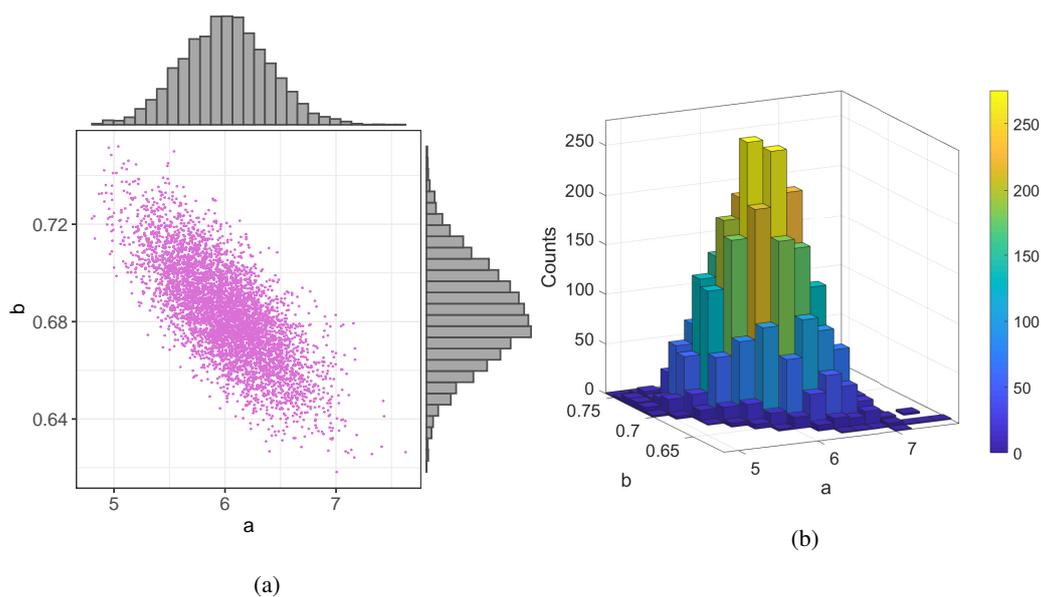
Figure 7. Values a and b of Eq. (11) obtained performing 5000 MC simulations: (a) scatter-plot and $2D$ histogram and (b) bivariate $3D$ histogram

240 (Table 3). The bandwidth decreases monotonically by increasing the duration and varies between 5.61 mm h^{-1} and 0.64 mm h^{-1}. This means that the highest absolute uncertainties are related to shorter durations.
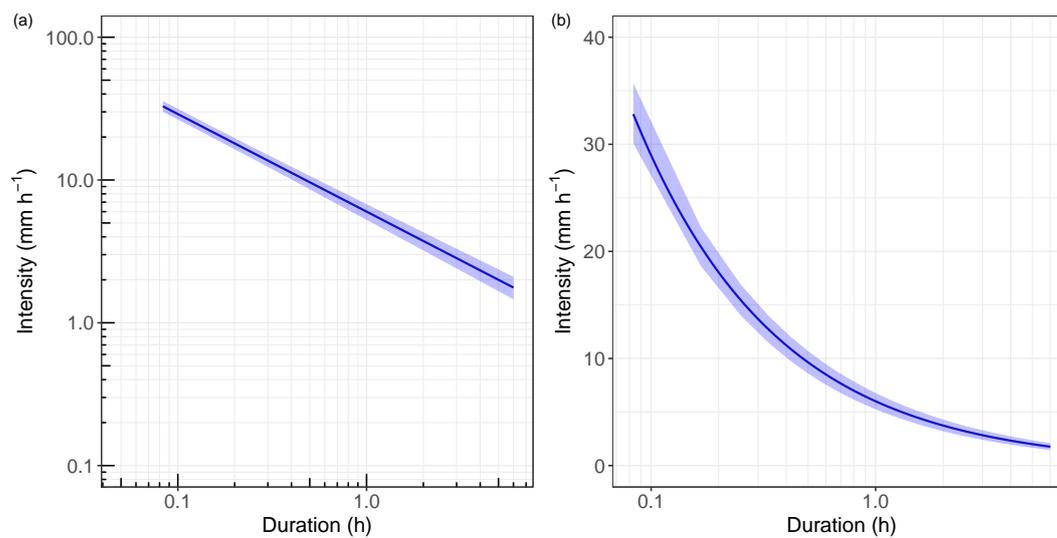
Figure 8. (a) Log-log and (b) semi-log plot of the threshold absolute uncertainty. The blue line is the rainfall threshold obtained using the mean value of a and b (Table 3). The shaded area represents the threshold uncertainty whose upper and lower bounds have been computed considering the 2.5 and 97.5 percentiles of the intensity distributions for fixed durations.

4.3 MC means and reference values

As regards the events rainfall condition, the biases between the reference values and the MC output distributions means (Eq.
245 (13)) are shown in Fig. 9. B_D deviates between 0.8 and 1.5 (Fig. 9(a)) while B_I between 0.86 and 1.15 (Fig. 9(b)). Consistently with the analysis of the CV described in Sect. 4.1, most events (48) are characterised by $B_D = B_I = 1$. Moreover, most of the remaining events have $B_D > 1$ and $B_I < 1$: the MC (I, D) mean couples tend to be located lower and more to the right than the reference ones. This condition highlights an underestimation of some mean rainfall conditions with respect to the reference ones.

250 As regards the threshold, the differences between the mean, lower and upper bounds MC intensities and the reference threshold ones (Eq. (14)) as a function of D are shown in Fig. 10(a). For almost all durations, the threshold obtained with the MC mean values of a and b is slightly lower than the reference one: a positive difference occurs only for the first time interval. Consistently with the obtained a and b mean values (Table 3) and the B_D and B_I trends, in the log-log ID plane the mean threshold is respectively slightly more downward translated and clockwise rotated than the reference one. Instead, the upper
255 and lower bounds are respectively always higher and lower than the reference threshold. This means that the latter is contained in the MC threshold uncertainty area. This behaviour is also evident in Figure 10(b) where the percentage changes (Eq. (15))
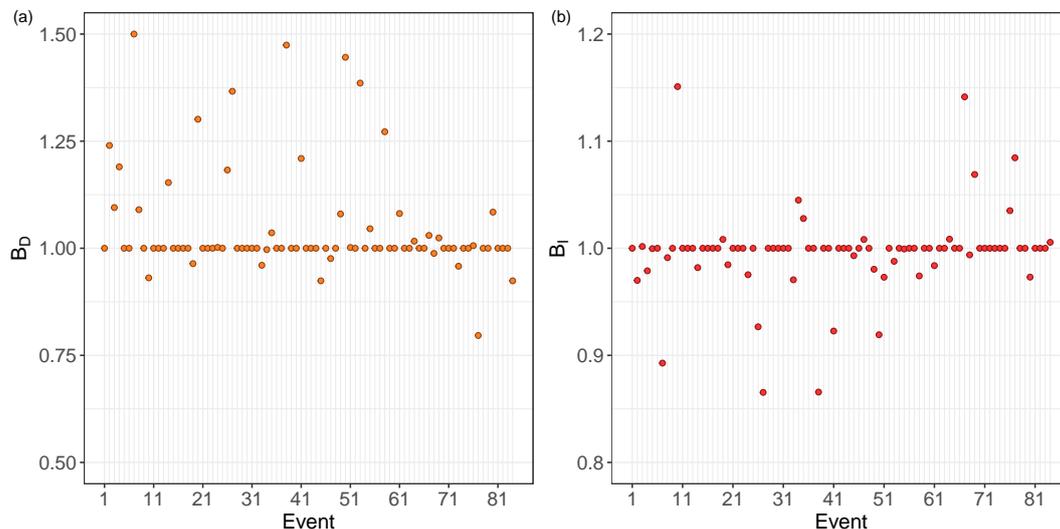
Figure 9. Bias of (a) duration B_D and (b) intensity B_I between the mean values of the variable distributions obtained performing the MC simulations and the reference values for each event.

are plotted: the mean threshold deviates between 0.14% and -5.44%, the upper bound between 8.06% and 12.31% and the lower bound between -8.34% and -22.94% from the reference one.
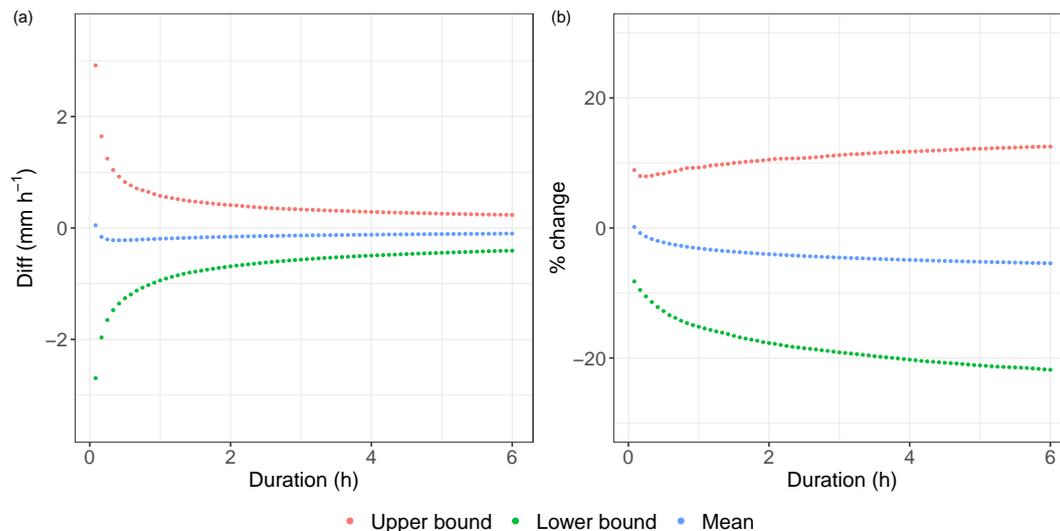
Figure 10. (a) Difference between the MC intensities (upper bound, lower bound and values obtained with the mean values of a and b) and the reference threshold ones as a function of the duration; (b) percentage change of the MC intensities respect the reference threshold ones, as a function of the duration.

5   Discussion and conclusions

260   This study aimed to figure out the effect of the uncertainty of some parameters on the computation of BDA outputs and on the subsequent threshold calibration.

In summary, the results obtained for the BDA outputs analysis has highlighted that the uncertainty of the inputs may result in a both low and high variation of the outputs distributions depending on the event characteristics and on the variable considered. The main uncertainties are related to the computation of D, as evident from Table 2 and Fig. 6, and are mainly due to the

265   piecewise constant nature and shape of the hyetograph. Indeed, given the discrete nature of $i(t)$, in some cases, even a small change in E may result in a relatively large variation of D (e.g. event 38), namely in the number of temporal instants necessary to achieve the needed rainfall volume. This is especially true if the hyetograph is characterized by time intervals with zero intensity within the event. However, although some events are characterised by a high uncertainty in the estimation of the (I, D) couple, the threshold computation is not affected by high variability. Both the threshold parameters have low variation

270   coefficient (Table 3), are low spread within the ab plane (Fig. 7) and the upper and lower bounds of the uncertainty area are close to the mean threshold (Fig. 8). This low uncertainty is mainly due to zero variability of 48 events outputs: since the (I, D) points of these events are located in the same positions in all the 5000 MC simulations, they propagate zero uncertainty in the threshold computation.

Moreover, the deviation between the MC results and the reference values has been analysed both in term of (I, D) couple

275   estimation and threshold computation. As evident from Fig.9, most events are characterised by B_D and B_I equal or close to 1. However, the duration bias of some events is high. Despite this, the reference threshold is very close to the threshold computed with the MC mean value of a and b and is contained in the uncertainty area (Fig.10). This means that if the uncertainty analysis has been done during the calibration of the threshold and the mean threshold had been accepted, there would have been no large variations.

280   As evident, the most critical part of the study is the computation of the (I, D) couples. High variability of some I and D output distributions points out the presence of possible critical issues in the estimation of the rainfall conditions associated with the event, as also highlight in Rosatti et al. (2019). The critical issues may be due to three main conditions. Firstly, some combinations of the input parameters, obtained with the LHS method, may be not representative of the analysed event. This condition leads to inconsistent estimates of the rainfall volume and, consequently, of D and I. Secondly, the analysed debris

285   flow may have had a different dynamic than the simplified one on which the BDA is based. For instance, the hypothesized equilibrium condition may have not been reached by some events due to particular conditions, such as lack of sediments, the presence of non-erodible zones or initiation of a debris flow caused by slope failure. This leads to an overestimation or underestimation of the concentration and, therefore, to a not reliable rainfall condition estimation. Finally, according to Marra et al. (2014), the radar data may be affected by some sources of error. During the calibration phase, to avoid the source of

290   error related to the beam shielding, typical phenomenon for mountain regions (Germann et al., 2006), the events have been filtered out based on the radar signal power. Events with radar signal weakened more than 90% have been excluded (Rosatti et al., 2019). However, other sources of error, such as signal attenuation in heavy rain or wet radome attenuation, have not been

considered in the current study. If present, these errors propagate in the computation of V_r^{DF} (Eq. (7)) implying an over or underestimation of the available rainfall volume. This results in a wrong estimate of the event (I, D) couple. Further analysis

295 will assess how these three conditions affect the (I, D) couple estimation and, consequently, the threshold calibration.

However, the BDA method seems to be robust enough to provide a reliable rainfall threshold: the input parameters uncertainties result in a low variation of the threshold estimation. Further analysis will also assess the robustness and the reliability of the threshold to forecast the possible occurrence of debris flow.

Author contributions. MM designed the experiments, performed the analysis and wrote the paper. DZ and GR supervised the study, analysed
300 the results and wrote the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work has been carried out within the project "Progetto WEEZARD: un sistema integrato di modellazione matematica a servizio della sicurezza nei confronti di pericoli idrogeologici in ambiente montano" (CARITRO Foundation – Cassa di Risparmio di Trento e Rovereto). We thank Prof. Siboni for providing valuable suggestions and Ripartizione Opere Idrauliche and Ufficio Idrografico,
305 Provincia Autonoma di Bolzano (Italy) and Servizio Bacini Montani and Ufficio Previsioni e Pianificazione, Provincia Autonoma di Trento (Italy) for supplying radar and debris flow data.

References

Abdi, H.: Coefficient of variation, Encyclopedia of research design, 1, 169–171, 2010.

Aleotti, P.: A warning system for rainfall-induced shallow failures, Engineering Geology, 73, 247 – 265, https://doi.org/https://doi.org/10.1016/j.enggeo.2004.01.007, Rainfall-triggered landslides and debris flows, 2004.

Blijenberg, H.: In-situ strength tests of coarse, cohesionless debris on scree slopes, Engineering Geology, 39, 137–146, 1995.

Brunetti, M., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Natural Hazards & Earth System Sciences, 10, 2010.

Cánovas, J. B., Stoffel, M., Corona, C., Schraml, K., Gobiet, A., Tani, S., Sinabell, F., Fuchs, S., and Kaitna, R.: Debris-flow risk analysis in a managed torrent based on a stochastic life-cycle performance, Science of the total environment, 557, 142–153, 2016.

Chien-Yuan, C., Tien-Chien, C., Fan-Chieh, Y., Wen-Hui, Y., and Chun-Chieh, T.: Rainfall duration and debris-flow initiated studies for real-time monitoring, Environmental Geology, 47, 715–724, 2005.

Coleman, H. W. and Steele, W. G.: Experimentation, validation, and uncertainty analysis for engineers, John Wiley & Sons, 2018.

Fuchs, S., Keiler, M., Sokratov, S., and Shnyparkov, A.: Spatiotemporal dynamics: the need for an innovative approach in mountain hazard risk management, Natural hazards, 68, 1217–1241, 2013.

Gariano, S. L., Melillo, M., Peruccacci, S., and Brunetti, M. T.: How much does the rainfall temporal resolution affect rainfall thresholds for landslide triggering?, Natural Hazards, https://doi.org/10.1007/s11069-019-03830-x, 2019.

Germann, U., Galli, G., Boscacci, M., and Bolliger, M.: Radar precipitation measurement in a mountainous region, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 132, 1669–1692, 2006.

Guzzetti, F., Peruccacci, S., Rossi, M., and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, Meteorology and Atmospheric Physics, 98, 239–267, https://doi.org/10.1007/s00703-007-0262-7, 2007.

Helton, J. C. and Davis, F. J.: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliability Engineering and System Safety, 81, 23–69, https://doi.org/10.1016/S0951-8320(03)00058-9, 2003.

Helton, J. C., Johnson, J. D., Sallaberry, C. J., and Storlie, C. B.: Survey of sampling-based methods for uncertainty and sensitivity analysis, Reliability Engineering and System Safety, 91, 1175–1209, https://doi.org/10.1016/j.ress.2005.11.017, 2006.

Håkanson, L.: The role of characteristic coefficients of variation in uncertainty and sensitivity analyses, with examples related to the structuring of lake eutrophication models, Ecological Modelling, 131, 1 – 20, https://doi.org/https://doi.org/10.1016/S0304-3800(00)00219-2, 2000.

Hofer, E.: The Uncertainty Analysis of Model Results, Springer International Publishing, https://doi.org/10.1007/978-3-319-76297-5, 2018.

Huebl, J. and Fiebiger, G.: Debris-flow mitigation measures, in: Debris-flow hazards and related phenomena, pp. 445–487, Springer, 2005.

Iadanza, C., Trigila, A., and Napolitano, F.: Identification and characterization of rainfall events responsible for triggering of debris flows and shallow landslides, Journal of Hydrology, 541, 230–245, https://doi.org/10.1016/j.jhydrol.2016.01.018, 2016.

Jakob, M., Owen, T., and Simpson, T.: A regional real-time debris-flow warning system for the District of North Vancouver, Canada, Landslides, 9, 165–178, 2012.

Lane, E. W.: Progress report on studies on the design of stable channels by the Bureau of Reclamation, in: Proceedings of the American Society of Civil Engineers, vol. 79, pp. 1–31, ASCE, 1953.

Marino, S., Hogue, I. B., Ray, C. J., and Kirschner, D. E.: A methodology for performing global uncertainty and sensitivity analysis in systems biology, https://doi.org/10.1016/j.jtbi.2008.04.011, 2008.

345 Marra, F.: Rainfall thresholds for landslide occurrence: systematic underestimation using coarse temporal resolution data, Natural Hazards, 95, 883–890, 2019.

Marra, F., Nikolopoulos, E. I., Creutin, J. D., and Borga, M.: Radar rainfall estimation for the identification of debris-flow occurrence thresholds, Journal of Hydrology, 519, 1607–1619, https://doi.org/10.1016/j.jhydrol.2014.09.039, 2014.

McKay, M. D., Beckman, R. J., and Conover, W. J.: A comparison of three methods for selecting values of input variables in the analysis of

350 output from a computer code, Technometrics, 42, 55–61, 2000.

Nikolopoulos, E. I., Crema, S., Marchi, L., Marra, F., Guzzetti, F., and Borga, M.: Impact of uncertainty in rainfall estimation on the identification of rainfall thresholds for debris flow occurrence, Geomorphology, 221, 286–297, 2014.

Pan, H.-L., Jiang, Y.-J., Wang, J., and Ou, G.-Q.: Rainfall threshold calculation for debris flow early warning in areas with scarcity of data, Natural Hazards and Earth System Sciences, 18, 1395–1409, 2018.

355 Peruccacci, S., Brunetti, M. T., Luciani, S., Vennari, C., and Guzzetti, F.: Lithological and seasonal control on rainfall thresholds for the possible initiation of landslides in central Italy, Geomorphology, 139, 79–90, 2012.

Pisoni, E., Albrecht, D., Mara, T. A., Rosati, R., Tarantola, S., and Thunis, P.: Application of uncertainty and sensitivity analysis to the air quality SHERPA modelling tool, Atmospheric environment, 183, 84–93, 2018.

Rosatti, G., Zugliani, D., Pirulli, M., and Martinengo, M.: A new method for evaluating stony debris flow rainfall thresholds: the Backward

360 Dynamical Approach, Heliyon, 5, https://doi.org/10.1016/j.heliyon.2019.e01994, 2019.

Rossi, M., Luciani, S., Valigi, D., Kirschbaum, D., Brunetti, M., Peruccacci, S., and Guzzetti, F.: Statistical approaches for the definition of landslide rainfall thresholds and their uncertainty using rain gauge and satellite data, Geomorphology, 285, 16–27, 2017.

Segoni, S., Piciullo, L., and Gariano, S. L.: A review of the recent literature on rainfall thresholds for landslide occurrence, https://doi.org/10.1007/s10346-018-0966-4, 2018.

365 Takahashi, T.: Mechanical characteristics of debris flow, Journal of the Hydraulics Division, 104, 1153–1169, 1978.

Takahashi, T.: Debris flow: mechanics, prediction and countermeasures, CRC press, 2014.