UNIVERSITY OF TRENTO

Doctoral School in Civil, Environmental and Mechanical Engineering

Curriculum 1. Civil and Environmental Engineering

Marta Martinengo

Improving some non-structural risk mitigation strategies in mountain regions: debris-flow rainfall thresholds, multi-hazard flooding scenarios and public awareness

2022 - Doctoral thesis

UNIVERSITY OF TRENTO - Italy
Department of Civil, Environmental
and Mechanical Engineering

Doctoral School in Civil, Environmental and Mechanical Engineering
Topic 1. Civil and Environmental Engineering - XXXIV cycle 2018/2021

Doctoral Thesis - September 2022

Marta Martinengo

Improving some non-structural risk mitigation strategies in mountain regions: debris-flowrainfall thresholds, multi-hazard flooding scenarios and public awareness

Supervisors
Prof. Giorgio Rosatti - University of Trento
Prof. Luigi Fraccarollo - University of Trento
Prof. Riccardo Rigon - University of Trento

Credits of the cover image: Provincia Autonoma di Trento
Debris flow of the Rotiano creek (Dimaro, Trento), the 29th of October 2018

Contents on this book are licensed under a Creative Common Attribution
Non Commercial - No Derivatives
4.0 International License, except for the parts already published by other publishers.

University of Trento
Doctoral School in Civil, Environmental and Mechanical Engineering
http://web.unitn.it/en/dricam
Via Mesiano 77, I-38123 Trento
Tel. +39 0461 282670 / 2611 - *dicamphd@unitn.it*

To my grandmother Teresa,
the strongest woman I have ever known.

Abstract

Hydrogeological hazards are quite diffuse rainfall-induced phenomena that affect mountain regions and can severely impact these territories, producing damages and sometimes casualties. Because of the potential great repercussions on both society and the economy, the mitigation of the hydrogeological risk, through its assessment and management, is crucial. In general, mitigation strategies attempt to reduce the hydrogeological risk to an acceptable residual level and they can be classified into structural (e.g., check dams, levees) and non-structural measures (e.g., early warning systems, land-use planning, public awareness). The structural measures are aimed at reducing the risk by modifying hazard magnitude and frequency through the employment of defence structures. Non-structural measures, on the other hand, are intended to mitigate the risk by reducing the damaging effects of an extreme event on people and assets. These latter measures are typical of the planning and are quite widespread since, in general, are more flexible, sustainable and cost-effective than structural ones.

This work focuses on enhancing some non-structural risk mitigation measures for mountain regions, such as: (i) debris-flow rainfall thresholds, as a part of a broader Early Warning System (EWS); (ii) multivariate rainfall scenarios with multi-hazard mapping purpose and (iii) public awareness.

With regards to rainfall thresholds, they are tools on which EWS for rainfall-induced natural phenomena are mainly based. Rainfall thresholds identify the rainfall limit conditions beyond which the probability to observe a rainfall-induced event is significant. However, there are some limitations to employing these tools in an operational EWS, mainly linked to their reliability and robustness. In this work, to face and try to overcome these limitations, with specific reference to rainfall thresholds for stony debris flows, an innovative calibration method, a suitable uncertainty analysis and a proper validation process are developed. First, a physical-based method to calibrate rainfall thresholds for stony debris flows is introduced. This method, named Backward Dynamical Approach (BDA), computes the rainfall conditions related to an observed debris flow by combining the forcing

and the dynamic of the phenomenon. The BDA is applied to a data set of a study area and a calibrated threshold is obtained. Secondly, the BDA robustness is tested by assessing the uncertainty in the calibrated threshold estimate. To this aim, a proper methodology composed of two Monte Carlo cascade simulations is developed and applied. In this way, the variability in the threshold estimate is assessed. Finally, a proper validation process comprised of two approaches is used to assess the calibrated threshold's reliability and possible forecast use. The first approach evaluates the reliability of the threshold by using the BDA to calculate the rainfall conditions related to already occurred debris flows. The second approach, on the other hand, takes into account all rainfall events, regardless of whether or not they are resulted in a debris flow, offering a comprehensive picture of threshold's reliability and paving the way for the application of a BDA-based threshold in a predictive framework. Overall, the results of these three analyses (i.e., calibration, uncertainty analysis and validation) suggest that the BDA method and the related calibrated threshold have satisfactory robustness and reliability. This finding sets the stage for the BDA approach to be used to calibrate stony debris-flow rainfall thresholds that can be employed in operational EWS. Moreover, the methods developed to compute the uncertainty analysis and the validation process are suitable to be applied to other calibration approaches and thresholds.

Regarding hazard mapping, it aims to identify the zones possibly affected by a hazard and is based on the definition of scenarios characterized by different occurrence probabilities and intensities of the phenomenon. When an area is potentially subjected to different dangerous phenomena, the hazard assessment is usually carried out examining each one separately. The so-called "multi-hazard" map is then obtained by gathering and overlaying the results of each hazard. This method fully ignores the possibility of many compound events occurring at the same time or in sequence, such as phenomena caused by statistically correlated drivers or hazards, as well as their interactions. With the goal of elaborating an actual multi-hazard map, this work aims to develop a proper methodology to construct multi-hazard scenarios considering, as an example study area, the confluence between a mountain stream susceptible to debris flows and a flood-prone valley river. Since precipitation is one of the main triggering factors for both debris flows and floods, the simultaneous occurrence of these phenomena is assessed, and the related multi-hazard scenarios construction is based on a rainfall multivariate analysis. The historical data of four reference rain gauges are elaborated and the related extreme values analysis is performed with the Simplified Metastatistical Extreme Value (SMEV) approach. Then, since

a correlation between data subsists, a four-variate copula is built, applying the pair copula construction (PCC) method, thus obtaining a vine copula, and the best-fit (conditional) bivariate copulas are estimated. Finally, the multivariate quantiles are computed considering both distinct definitions and several values of the multivariate return period. The results are multivariate scenarios with a varying occurrence probability and intensities of the compound event precipitation forcing under investigation. These scenarios are promising to be applied in hazard assessment to construct actual multi-hazard maps. Moreover, the results of this study highlight that neglecting the correlation between the drivers of the analyzed compound events can lead, in some cases, to an underestimation of their simultaneous occurrence probability. This underestimation can result in an unreliable hazard map of the investigated area. In addition, the methodological scheme developed can be applied to other types of compound events characterized by forcing that can result in simultaneous hazardous phenomena.

The final topic in this work is public awareness, which is briefly considered. Besides strictly technical mitigation strategies, a key element in hydrogeological risk management is public awareness and preparedness. In peacetime and emergency situations, a well-informed and prepared populace can indeed take preventative measures to decrease their vulnerability and exposure to hazards. The European project LIFE FRANCA (Flood Risk ANticipation and Communication in the Alps) fits in this context by promoting a hydrogeological risk prevention culture and aiming to increase the population safety degree. Among other project actions, the one considered in this work is focused on training and communication activities that aim to provide basic knowledge on several aspects of hydrogeological risk assessment and management. To this end, some courses and seminars have been held involving as many different types of stakeholders and speakers as possible. In this way, a multidisciplinary view on hydrogeological risk has been provided, and fruitful confrontations and constructive interactions have arisen.

Contents

Introduction **1**

I Stony debris flow rainfall thresholds **5**

1 A new method for evaluating stony debris flow rainfall thresholds: the Backward Dynamical Approach **7**

1.1 Introduction 7

1.2 Estimate of the rain relevant to a debris-flow event: the classical literature approach 9

1.2.1 A specific approach for the determination of t_{init}: the Critical Duration Method 10

1.2.2 The possible weak point of the classical approaches . . 13

1.3 Estimate of the rain relevant to a debris-flow event: the Backward Dynamical Approach 14

1.3.1 Basic assumptions for the BDA relation 16

1.3.2 Conceptual scheme of a stony debris-flow dynamics . . 17

1.3.3 Quantitative volumetric description of the debris-flow dynamics 19

1.3.4 The role of the rain volume in a debris-flow 25

1.4 Probability threshold for intensity-duration data: the frequentist method. 25

1.5 The rainfall thresholds for a sample study area 26

1.5.1 Available data 26

1.5.2 The CDM-based threshold 28

1.5.3 The BDA-based threshold 31

1.6 Discussion 32

1.6.1 The CDM-based threshold 33

1.6.2 The BDA-based threshold 34

1.6.3 Comparison between the CDM and the BDA-based thresholds . . . 35
1.6.4 Some issues regarding the forecast use of the BDA-based threshold . . . 36
1.7 Conclusions . . . 38

2 Uncertainty analysis of a rainfall threshold estimate for stony debris flow based on the Backward Dynamical Approach 41
2.1 Introduction . . . 41
2.2 The BDA-based threshold calibration . . . 43
2.3 Study area and data . . . 46
2.4 Method . . . 48
2.5 Results and discussion . . . 51
2.5.1 Variability of the event characteristics . . . 51
2.5.2 Correlation between the D and I absolute variability and some event features . . . 57
2.5.3 Variability of the threshold . . . 62
2.5.4 Reference values versus MC means . . . 63
2.5.5 Further elements of uncertainty . . . 66
2.6 Conclusions . . . 66

3 Validation and potential forecast use of a debris-flow rainfall threshold calibrated with the Backward Dynamical Approach 69
3.1 Introduction . . . 69
3.2 The Backward Dynamical Approach . . . 71
3.3 Study area and data . . . 73
3.3.1 Data uncertainties . . . 74
3.4 Method . . . 75
3.4.1 The BDA-based validation approach . . . 76
3.4.2 The potential debris flows validation approach . . . 77
3.5 Results . . . 80
3.5.1 BDA-based validation approach . . . 80
3.5.2 Potential debris flows validation approach . . . 83
3.6 Discussion . . . 86
3.6.1 Consistency of the approaches . . . 86
3.6.2 Reliability of the threshold . . . 88
3.6.3 Potential forecasting use of the threshold . . . 89
3.7 Conclusions . . . 91

II Multivariate rainfall analysis with multi-hazard mapping purpose 93

4 Multivariate rainfall analysis for multi-hazard assessment of combined debris flow-flood events in a mountain confluence 95
4.1 Introduction . . . 95
4.2 Basic theoretical framework . . . 98
4.2.1 Copulas . . . 98
4.2.2 Dependence measures . . . 103
4.2.3 Pair-copula construction and vine copulas . . . 105
4.2.4 Simplified Metastatistical Extreme Value . . . 108
4.2.5 Multivariate return period . . . 109
4.3 Study area and data . . . 112
4.4 Method . . . 113
4.4.1 Event and related observations definition . . . 116
4.4.2 Marginal distributions . . . 117
4.4.3 Dependence structure estimate, modeling and validation 120
4.4.4 Scenarios . . . 124
4.5 Results and discussion . . . 127
4.5.1 Observations and empirical dependence measures . . . 127
4.5.2 Fitted marginal distributions . . . 131
4.5.3 Dependence structure modelling . . . 137
4.5.4 Isohypersurfaces and scenarios . . . 152
4.5.5 Possible use of the obtained scenarios with a hazard mapping purpose . . . 163
4.6 Final remarks . . . 164

III Public awareness within the LIFE FRANCA project 167

5 The LIFE FRANCA European project 169
5.1 General framework and project overview . . . 169
5.2 LIFE FRANCA and public awareness . . . 171
5.3 Training and education action . . . 173
5.3.1 Method . . . 173
5.3.2 Main results . . . 174
5.4 Final considerations on the project . . . 175

Final considerations 177

Bibliography 178

Acknowledgements **203**

Introduction

Hydrogeological hazards, namely floods, landslides and debris flows, are climate-related phenomena connected to the interaction between meteorological events and the geological environment that can have severe consequences for the impacted areas, such as casualties and economic losses (e.g. Barredo, 2009, Canuti et al., 2001, Dowling and Santi, 2014, Jonkman and Kelman, 2005, Paliaga et al., 2020, Petrucci et al., 2019). Worldwide, between 1998 and 2017, hydrogeological disasters resulted in more than 160,000 casualties with another 2 billion people wounded, evacuated, or in need of emergency aid. In the same period, the direct economic losses recorded by the disaster-affected countries due to hydrogeological phenomena accounted for more than US$ 650 billion (CRED and UNISDR, 2018).

Worldwide, many areas may be affected by future hydrogeological phenomena. Considering for instance Italy, an overview of hydrogeological risk in the national territory is provided by the 2021 report on landslides and floods, realized by ISPRA (Istituto Superiore per la Protezione e la Ricerca Ambientale) (Trigila et al., 2021). The 18.4% of the national territory is classified as a very high or high hydrogeological hazard level with 1.3 million people potentially subjected to landslides and debris flows and 6.8 million people living in flood-prone areas. As highlighted in Trigila et al. (2021), the geological, morphological, meteo-climatic, and topographical characteristics of the territory, as well as its high urbanization, often developed in the absence of adequate land-use planning, cause the consistent presence of areas subjected to hydrogeological risk in Italy.

The consequences of past disasters and the huge number of people and assets potentially exposed and vulnerable to future phenomena point out the relevance of reducing hydrogeological risk, as also highlighted in the Sendai Framework for Disaster Risk Reduction 2015–2030 (UNDRR, 2015). Risk reduction can be achieved by mitigation strategies that aim to decrease it to an acceptable residual level and comprise a variety of actions, tools and policies to be planned, designed, developed and implemented before a hazard occurrence. Mitigation strategies are commonly subdivided into struc-

tural and non-structural measures (e.g. Dalezios, 2017, Godschalk et al., 1999, Kundzewicz, 2009, Starominski-Uehara, 2021). The structural measures aim to reduce risk by modifying the hazard in terms of its likelihood and/or magnitude so as to keep it away from potentially exposed and vulnerable elements (e.g. Simonovic, 2002). Examples of structural measures are check dams, levees, sediment catchment basins and slope stabilization. The non-structural measures, instead, intend to mitigate risk by reducing the exposure and vulnerability of people and assets potentially affected by a hazardous event. Early Warning Systems (EWS), land-use planning, public awareness-raising and emergency plans are examples of non-structural measures.

Nowadays, a long-term mitigation strategy is usually based on a coordinated and appropriate combination of measures of both types (e.g. Luz Barcellos et al., 2017) owing to the recognition that non-structural measures are an indispensable complement to structural ones (e.g. Barendrecht et al., 2020, Simonovic, 2002). Indeed, with the growing awareness that risk cannot be completely eliminated, the paradigm of risk mitigation has shifted from hazards removal to vulnerability and exposure reduction (e.g. Cardona et al., 2012, Mees et al., 2016, Mondino et al., 2020, Papagiannaki et al., 2019a), leading to a great diffusion of non-structural mitigation strategies (e.g. Kim et al., 2021, Kuller et al., 2021, Reghezza-Zitt and Rufat, 2019). This spread is also influenced by the fact that non-structural measures are typically more cost-effective (e.g. Dawson et al., 2011, Pesaro et al., 2018), adaptable and sustainable than structural ones (e.g. Kundzewicz, 2009).

Given their relevance and wide use, this thesis is focused on the improvement of some non-structural mitigation strategies, with particular reference to debris flows and floods risk mitigation in mountain regions. The considered measures are: (i) debris-flow rainfall thresholds, as a part of a more general Early Warning System (EWS); (ii) multivariate rainfall scenarios with multi-hazard mapping purpose and (iii) public awareness.

This thesis is structured in three parts, one for each considered non-structural mitigation strategy. The first part deals with stony debris-flow rainfall thresholds. First, a new physical-based method to calibrate rainfall thresholds for stony debris flows and its application to estimate a rainfall threshold for a study area are presented in Chapter 1. The topics of this chapter are covered in Rosatti et al. (2019). Second, the robustness of the developed calibration method, assessed through an uncertainty analysis of the threshold estimate, is described in Chapter 2. The analyzes presented in this chapter are published in Martinengo et al. (2021b). Finally, the reliability of the calibrated threshold, assessed performing a validation analysis,

and its potential forecasting use are described in Chapter 3. The analyzes covered in this chapter have been submitted for publication (Martinengo et al., 2022).

The second part of the thesis concerns a multivariate rainfall analysis with a multi-hazard mapping purpose. In Chapter 4, a copula-based model to jointly assess rainfall values at different locations is presented, and the possible use of the resulting rainfall scenarios with multi-hazard mapping purposes is described.

Finally, the third part of the thesis deals with public awareness. In Chapter 5, the European project LIFE FRANCA (Flood Risk ANticipation and Communication in the Alps) is briefly presented and the actions taken and the related results achieved are described.

Part I

Stony debris flow rainfall thresholds

Chapter 1

A new method for evaluating stony debris flow rainfall thresholds: the Backward Dynamical Approach

1.1 Introduction

In alpine regions, debris flows are rather common and widespread phenomena that produce considerable damages to houses and infrastructures. In the last decades, steps forward have been made in the field of mathematical and numerical modelling of these phenomena by using either a mono-phase or a two-phase approach (see e.g. Armanini et al., 2009, Brufau et al., 2000, Iverson, 1997, Liu et al., 2017, Pirulli and Sorbino, 2008, Pudasaini, 2012, Takahashi, 1978, among others) and even tools aimed at simplifying back-analysis of past events, evaluating forward scenarios of possible events and drawing up hazard maps (Chen et al., 2016, Iverson, 2014, Rosatti et al., 2018, among others) are now available. Nevertheless, a reliable forecast concerning the possibility that a storm may produce a debris flow is still hard, even just in a probabilistic framework. Attempts aiming at this goal have led to what is known in the literature as the *rainfall threshold* approach (Caine, 1980, Guzzetti et al., 2007). According to this approach, indicating with D the duration of a rainfall event pertaining to a debris flow and with I the relevant average rainfall intensity, the rainfall threshold is a function that splits the I-D domain in two fields: one located above the threshold and one located below. If a storm is foreseen to have a couple (I, D) falling in the upper field, it is expected that a debris flow may occur during that

storm. On the contrary, if the foreseen couple does not exceed the threshold, it is expected that the storm should not generate a debris-flow.

Two main components are necessary to set up a rainfall threshold starting from an historical series of rainfalls related to debris-flow events:

- a method for evaluating the amount of rain strictly relevant to each debris-flow event in term of cumulated volume of precipitation V, averaged intensity I and duration D (we stress here that only two of the previous variables are independent);
- a statistical approach that allows to identify a given occurrence probability threshold starting from a series of intensity-duration couples, namely (I_k, D_k) with $k = 1, \ldots, N$ and N is the total number of registered events in a given area.

In the literature, there are a few papers facing the rainfall threshold topic (see Sec. 1.2 and Sec. 1.4 for specific references). They differ (sometimes slightly) in the method used to evaluate the first or the second component of the procedure described above, but they share the same hydrological perspective, based on the assumption that the proper amount of water pertaining to a debris flow event can be estimated solely on the analysis of the rain hyetograph in a suitable time interval around the debris-flow occurrence time. Unfortunately, as stressed by some authors (Nikolopoulos et al., 2014, Staley et al., 2013), thresholds obtained so far with these approaches are unlikely low and many events with rain exceeding the threshold do not give rise to any debris-flow occurrence. Therefore, rainfall thresholds are seldom used as an effective tool in debris-flow warning systems.

The goal of the research underlying this paper is exploring new paths aimed at improving the reliability of rainfall thresholds. We started our work analyzing the possible weak point of the classical literature approach and it seemed to be the way the rain volume pertaining to a debris flow is determined. Then, we have looked for an alternative methodology for determining this volume involving, in some way, not only the forcing of the phenomenon (represented by the rain hyetograph), but also the dynamics of the debris flow. The novel approach we have developed starts from the knowledge of the volume occupied by the sediments deposited in an event (typically measured by regional agencies), then, the liquid volume (i.e. the rain) responsible of this mass movement is back reconstructed by using a simplified dynamical description of the phenomenon. Finally, the duration of the storm giving rise to the estimated volume (and the relevant average intensity as well) is identified from the analysis of the rain hyetograph. We have called this method the Backward Dynamical based Approach (BDA).

A proof of concept is presented in the paper by applying the BDA and a literature method to a sample study area.

The structure of the paper is the following. In Section 1.2, we give a synthetic overview of the classical literature approaches used to estimate the rain relevant to a debris-flow event, with particular focus on the Critical Duration Method (CDM), and we identify a possible intrinsic limit of these approaches. Section 1.3 is dedicated to the detailed description of the BDA theoretical framework while, in Section 1.4, a literature method for evaluating a threshold relation from an *I-D* series, namely the frequentist method, is briefly presented. A sample study area and the relevant rainfall thresholds obtained by using BDA and CDM are presented in Section 1.5. Discussion of the result is reported in Section 1.6. Conclusions end the paper.

1.2 Estimate of the rain relevant to a debris-flow event: the classical literature approach

In the literature, the rain relevant to a debris-flow event is assumed to be the volume of water poured into a basin from the early beginning of the storm system, indicated with t_{init} (time measured respect an arbitrary reference), up to the debris-flow triggering time (see e.g. Lazzari et al., 2013).

Assuming a uniform distribution of the precipitation over the basin, the volume V can be evaluated in the following way:

$$V = A_b \int_{t_{init}}^{t_{trig}} i(t)\, dt \tag{1.1}$$

where $i(t)$ is the measured hyetograph and A_b is the area of the basin. The rainfall duration D is then defined as the difference between two characteristic times, namely the triggering and the initial time:

$$D = t_{trig} - t_{init}. \tag{1.2}$$

Finally, the average intensity

$$I = \frac{1}{D} \int_{t_{init}}^{t_{trig}} i(t)\, dt \tag{1.3}$$

can be obtained dividing the volume V by the duration D and the area of the basin A_b:

$$I = \frac{V}{DA_b} \tag{1.4}$$

Therefore, considering a given hyetograph, the procedure to work out a couple (I, D) depends on the determination of the two characteristic times. Their practical determination, despite their simple definitions, is not easy at all. In fact, the triggering time is seldom a piece of data known with precision. In most cases, only the date of a debris-flow occurrence is available in an event report. In these situations, the triggering time is assumed the instant in which the storm reaches the maximum rainfall intensity, i.e. $t_{trig} \equiv t_{max}$ (Iadanza et al., 2016). Other approaches (see e.g. Lazzari et al., 2013) assume this time as the last rainy measurement of the day.

Even more complex is the determination of the beginning of the rainfall event. In the literature there are many methods developed for this purpose and can be divided into two categories. The first collects the methods in which t_{init} is estimated based on empirical and subjective rules, valid only for the specific sites for which they have been developed. The second refers to objective methods based on a probabilistic approach. Despite several works use the first way (e.g. Bel et al., 2017, Nikolopoulos et al., 2014, Zhuang et al., 2015), the methods belonging to the second category (e.g. Bonta, 2001, 2003, Bonta and Nayak, 2008, Iadanza et al., 2016) seem more attractive because of their generality and possibility to be applied to different study areas.

1.2.1 A specific approach for the determination of t_{init}: the Critical Duration Method

The Critical Duration Method (CDM) (Bonta and Rao, 1988, Restrepo-Posada and Eagleson, 1982) aims at identifying independent storm systems singling out a characteristic time, the Critical Duration (CD), such that if the time between two rainy periods is longer than this value, the two rainy periods can be considered belonging to independent storm systems. A synthetic summary of the statistical procedure for the evaluation of the CD is reported in the following Section. For a more detailed description, we refer the reader to the original papers mentioned above. A monthly value is commonly determined since the statistical characteristics of the storms change with the seasons.

Once the CD has been estimated, t_{init} is determined as the first dry instant before t_{trig}, whose time distance respect to the previous rainy instant is equal or longer than CD. An example of CDM application is reported in Figure 1.1, where the triggering time has been considered as the t_{max} of the event.

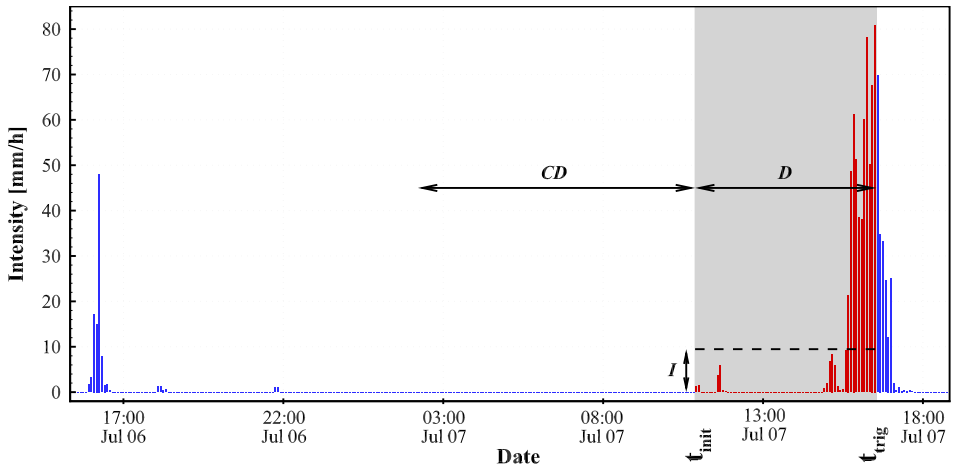

Figure 1.1: Example of determination the rain relevant to a real debris flow. Here, $t_{trig} \equiv t_{max}$, t_{init} is obtain by using the CDM, $D = t_{trig} - t_{init}$ and I is the event-averaged rain intensity.

1.2.1.1 A synthetic summary of the CDM

A time sequence of rainfall records consists of rainy periods separated by non-rainy intervals. When two rainy periods are separated by a time interval of the order of minutes (or possibly even some hours), they surely cannot be considered independent, but generated by the same meteorological system. On the contrary, when the two rainy periods are separated by a long dry interval, they can be considered caused by independent weather systems. The CDM identifies, by using a statistical approach, a characteristic time interval, called critical duration CD, that allows distinguishing independent rainstorms and it is defined as the minimum dry period, which separates two stochastically independent rainy periods.

The main hypothesis in the CDM, verified by Restrepo-Posada and Eagleson (1982), is that the rainfall event durations are much shorter than the interstorm (dry) periods. Under this hypothesis, as proposed by Bonta and Rao (1988), the occurrence of storms can be assimilated to a random Poisson process, whose interstorm period distribution can be approximated with an exponential function of this type:

$$f(t) = \alpha e^{-\alpha t} \tag{1.5}$$

where α is the reciprocal of the mean time between storms and t is the dry-period duration. It is useful to recall that the variation coefficient CV, given by the ratio of the standard deviation s to the expected value E, is

unitary. This feature can be used to single out the CD as the time beyond which the histogram of the dry periods can be well approximated with the exponential distribution. The procedure to get the CD is the following:

1. Subdivide a time axis in time steps long as the gauge sampling interval δt. Each interval is defined by:

$$[(n-1)\delta t, n\delta t], \quad n = 1, \dots N$$

   where the central value is $(n - 1/2)\delta t$. Consider the number of occurrence of dry periods whose durations fall inside each interval. Build the relevant histogram (see Fig. 1.2(a)). It is worth noting that some intervals may be empty.

2. Call "classes" the *non-empty* time intervals of the previous histogram, number them from 1 to K and associate to each class the central time of the relevant interval. Build the histogram of the occurrence as a function of the classes (see Fig. 1.2(b)). In this case, the histogram has a value associated to each class.

3. Consider the sequence given by:

$$CV_k = \frac{s_{[k,N]}}{E_{[k,N]}}, \quad k = 1, \dots K \tag{1.6}$$

   where $E_{[k,N]}$ and $s_{[k,N]}$ are, respectively, the expected value and the standard deviation for the dataset composed by the values included in the class interval $[k, N]$

4. Call $\widetilde{k}$, the first value of the sequence such that

$$CV_{\widetilde{k}} = \frac{s_{[\tilde{k},N]}}{E_{[\tilde{k},N]}} \leq 1 \tag{1.7}$$

5. The value of the critical duration is defined as the time corresponding to the $\tilde{k}$-th class, namely

$$CD = t_{\widetilde{k}} \tag{1.8}$$

An example of CD determination is shown in Fig. 1.2 where the bars represent the histogram of dry-periods while the dotted lines are the values of CV_k.

According to Iadanza et al. (2013), the minimum number of years of data necessary to correctly estimate CD is 6. In this study we have used 10 years of data.

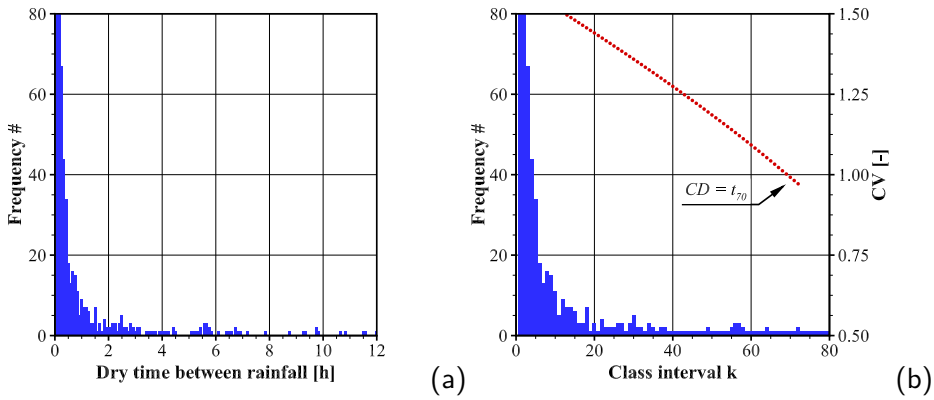

Figure 1.2: Histogram of dry period between rainfall associated with the time (a) and with the class used for the evaluation of CV (b). In the two plot are reported the same amount of data.

1.2.2 The possible weak point of the classical approaches

As we hinted at the beginning of the paper, the methods present in the literature produce unlikely low rainfall thresholds. The reasons of this behaviour may be multiple and connected both either to the first or to the second component of the procedure depicted in the Introduction. Nevertheless, we think that this feature is mainly due to a not entirely appropriate estimate of the rain relevant to debris-flow events or, according to eq. (1.1), a not entirely appropriate definition of the characteristic times.

Although the triggering time presents significant uncertainties that can affect the reliability of the threshold, we think the assumption that the initial time, defined as the early beginning of the storm system during which a debris-flow event occurs, is conceptually the main weak point of the classical approaches. In fact, this choice leads to include in the duration D even all the dry periods occurring *within* a given storm system, with a consequent average intensity I that can be non-significant or non-representative of the real intensity that caused a debris-flow (see Figure 1.1). Moreover, with the given definition of t_{init}, the water considered relevant to a debris-flow event is not only the water directly involved in the mass movement, but also the water that leads to the pre-condition for a debris-flow event, i.e. the saturation of the soil at least for a depth equal to the eroded layer.

The reason of this basic assumption can be due to a couple of factors. Firstly, the rainfall threshold approach and the given definition of t_{init} were originally proposed for generic hydrogeological phenomena, namely land-

slides and debris-flows together. For landslides, the water responsible of the phenomenon is probably all the water occurred before its triggering, while the water rained after that time plays a negligible role. Therefore, the assumption seems quite reasonable. On the contrary, in debris flows, the water volume flowing *after* the initiation of the flow plays an important role, since it determines the actual volume of the debris flow (as it will be explained in the next Section). Therefore, because of the differences of the two phenomena, an undistinguished application of the approach to landslides and to debris flows does not appear completely appropriate. The second factor, in a pure hydrological framework, is finding some features of the measured hyetograph that can be correlated to the debris flow itself to evaluate quantitatively the relevant rain. The hydrological beginning of a storm system, independently of the actual way it is determined, is therefore not only the natural choice but also perhaps even the only practical choice that can be done to characterize the starting of the rain relevant to a debris flow.

1.3 Estimate of the rain relevant to a debris-flow event: the Backward Dynamical Approach

Our perspective for estimating the (I, D) couple relevant to a debris-flow event is quite different from the perspective used so far in the literature. Hereafter we present a detailed description of the rigorous framework leading to the BDA, along with all the assumptions necessary to define it.

Let us indicate with V_{rain}^{DF}, the volume of rain directly embroiled in the mass movement. Assuming a uniform distribution of the precipitation over the basin, this volume can be expressed in similar way to eq. (1.1) as:

$$V_{rain}^{DF} = A_b \int_{t_1}^{t_2} i(t)\, dt \tag{1.9}$$

where t_1 and t_2 are two time limits related to the debris-flow duration and and A_b is the area of the basin. We assume that:

1. V_{rain}^{DF} can be back reconstructed from the knowledge of V_{dep}, i.e. from the volume occupied by the mixture that stops in a deposition fan during the event. The relation between the two volumes, that we call "BDA relation", can be expressed formally as

$$V_{rain}^{DF} = f(V_{dep}) \tag{1.10}$$

and can be obtained considering a simplified dynamical description of the debris-flow phenomenon, to be specified further on.

2. the time limits t_1 and t_2 are conceptually connected to t_{trig} and to the debris flow duration but, in general, not to t_{init}. Since the triggering time and the duration of a debris flow event is rarely available, following Iadanza et al. (2016), we assume $t_{trig} \equiv t_{\max}$ while, for the time interval, we assume that it is symmetrically distributed around $t_{\max}$, namely:

$$[t_1, t_2] = [t_{\max} - \Delta t, t_{\max} + \Delta t] \tag{1.11}$$

where Δt is an unknown value.

The duration and average intensity of an event can be determined by using the previous three equations in the following way:

i. considering eq. (1.10) and (1.11), eq. (1.9) can be rewritten as

$$f(V_{dep}) = A_b \int_{t_{\max}-\Delta t}^{t_{\max}+\Delta t} i(t)\, dt \tag{1.12}$$

where the only unknown is Δt; this equation can be solved by means of a trial and error method;

ii. the rainfall duration, that we will indicate with $\hat{D}$ in order to distinguish from the classical duration D, is equal to the length of the time interval used in the integral:

$$\hat{D} = 2\Delta t \tag{1.13}$$

iii. considering the previous two expressions, the average intensity

$$\hat{I} = \frac{1}{\hat{D}} \int_{t_{\max}-\Delta t}^{t_{\max}+\Delta t} i(t)\, dt$$

can be rewritten as:

$$\hat{I} = \frac{f(V_{dep})}{A_b\, \hat{D}} \tag{1.14}$$

Rainfall thresholds will be calculated using couples of values $(\hat{I}, \hat{D})$ instead of (I, D). Consequences of this new point of view on threshold evaluations are analysed further in the paper (see Section 1.6).

In order to make the previous procedure effective, it is now necessary to specify the BDA relation. This is the subject of the following Sections.

1.3.1 Basic assumptions for the BDA relation

The BDA relation is based on a series of assumptions and approximations that, even if apparently rough, are quite reasonable and can give, at least, the right order of magnitude of the rain estimate. We list here all the assumed hypotheses and some motivations of their reasonableness.

- We refer only to stony debris flows, i.e. flows in which silt or clay does not affect the overall behaviour of the mixture.
- The concentration of sediments in the bed c_b is constant everywhere. This assumption is generally accepted in debris-flow dynamics (Takahashi, 2007) .
- When a debris flow occurs, the soil is commonly completely saturated (e.g. Hungr et al., 2002). Therefore, we assume that the basin has been saturated by the rain foregoing the event, up to the level interested to the erosive stage. This condition is reasonable even if nobody has ever verified it in real cases. From this assumption and the previous ones, it follows that along the debris-flow path the pointwise water content in the terrain is everywhere $(1 - c_b)$.
- Since the characteristic time scale of a debris-flow event is much smaller than the time scale of the infiltration process, we assume that during a debris-flow occurrence, all the rainfall transforms into runoff.
- The volume of sediments surveyed in the field after an event is the major part of the sediment involved in the debris-flow. This can be accepted if the deposition fan is a piece of territory with average slope significantly smaller than the slope in the flowing part and the volume of small-size sediments carried away with the water is negligible. This situation commonly occurs if the basin has a well-defined deposition fan.
- The reach just upstream the deposition fan is characterized by an average slope $i_f = \tan\theta$, where θ is the angle that the bed forms with an horizontal reference direction, and by a length long enough

to allow the debris flow to be in uniform flow condition, i.e. with a volumetric solid concentration that, according to Takahashi (1978), for stony debris flows can be expressed as

$$c = \frac{\tan\theta}{\Delta\left(\tan\psi - \tan\theta\right)} \tag{1.15}$$

where $\Delta = (\rho^s - \rho^l)/\rho^l$ is the sediment relative submerged density, ρ^l and ρ^s are, respectively, the liquid and solid density while ψ is the internal friction angle. Considering that debris flows reach the uniform flow condition in rather short lengths, this assumption seems to be applicable in most cases.

1.3.2 Conceptual scheme of a stony debris-flow dynamics

Besides the previous assumptions, the BDA relation is based on a schematisation of a debris-flow dynamics. Obviously, real dynamics may be much more complex than the simple conceptual scheme herein reported. Nevertheless, we think that it is reasonably representative of a large set of real conditions.

For sake of simplicity, let us consider a long plane bed with unit width and constant slope i_f. An upper non-erodible transept followed by a lower erodible and saturated one characterizes the plane bed. At the end of the plane bed, there is a quasi-horizontal reach (see Fig. 1.3(a)).

The phenomena developing in the three transepts are:

Transept 1 : Runoff formation. Rain gives a significant contribution only in this transept.

Transept 2 : Erosion of the bed material and formation of a debris-flow characterized by a concentration given by eq. (1.15).

Transept 3 : Deposition of all the sediments with saturation water entrapment. The rest of the water flow away outside the deposition fan.

In Fig. 1.3(a) we have sketched the phenomena occurring in the three transepts. The picture can be considered also a superimposition of three snapshots taken at three subsequent times (Lagrangian description) in the case the three phenomena occur disjointedly. Actually, it is very likely that they occur simultaneously in different position of the flow field but, in order to make the plot clearer, we preferred to represent the disjointed case.

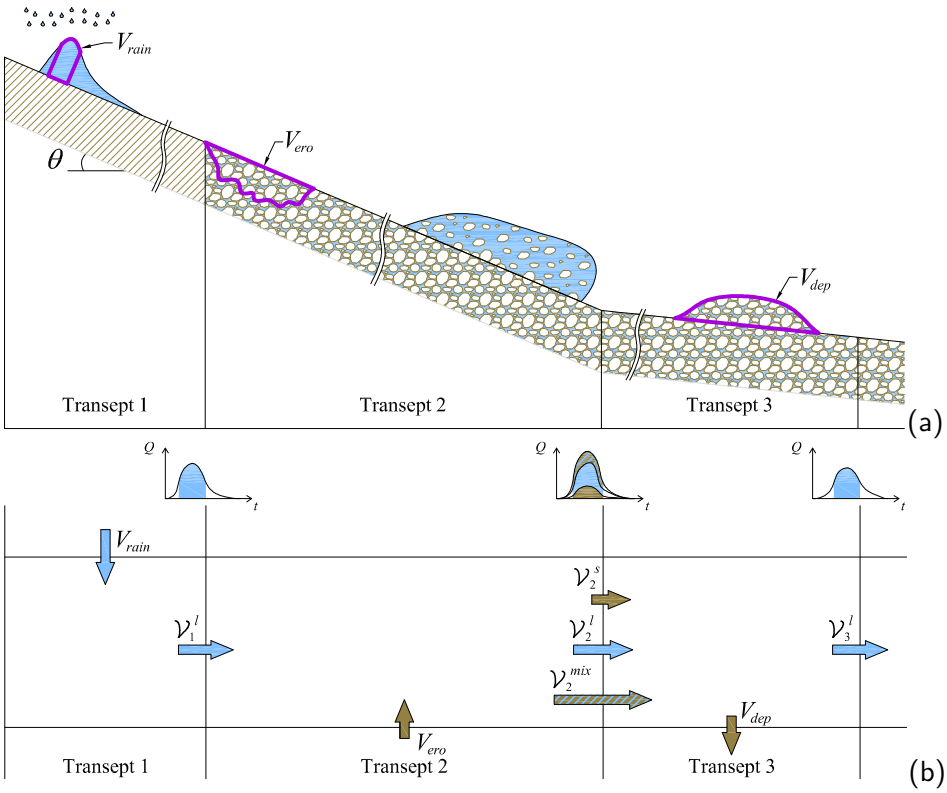

Figure 1.3: Sketch of the conceptual scheme of a debris-flow dynamics: (a) Lagrangian point of view; (b) Eulerian point of view. Meanings of the quantities can be found in the text.

1.3.3 Quantitative volumetric description of the debris-flow dynamics

In order to obtain a quantitative volumetric description of the conceptual debris-flow dynamics, it is useful turning to the one-dimensional Eulerian description of the flow sketched in Fig. 1.3(b). Here we report the detailed derivation of the relations valid in each transept, along with the assumptions necessary to obtain them. More complex relations can be obtained relaxing one or more hypothesis, but this generalization is left for a future work.

The debris-flow phenomenon can be described starting from the partial differential equations expressing the mass balances in a two-phase, depth-averaged framework in which there is no lag between the phases (see e.g. Armanini et al., 2009, Rosatti and Begnudelli, 2013). More specifically, the mass-balance equation of the liquid phase and the solid one can be written as:

$$\frac{\partial}{\partial t}\left[(1-c)h+(1-c_b)z_b\right]+\frac{\partial}{\partial x}\left[(1-c)uh\right]=S_l \tag{1.16a}$$

$$\frac{\partial}{\partial t}\left[ch+c_b z_b\right]+\frac{\partial}{\partial x}\left[cuh\right]=0 \tag{1.16b}$$

where h is the mixture depth, c is the solid volumetric concentration, z_b is the mobile-bed elevation, u the depth-averaged velocity and finally S_l is the source term for the liquid phase. Here, the constant densities of each phase has been simplified. Since no source for the solid phase is present, the relevant term is null. Momentum balance of the solid phase is also accounted for by mean of the uniform flow relation expressed by eq. (1.15).

The relations necessary for our approach are derived from the integration of the mass balance equations, namely eq. (1.16a) and (1.16b), along each transept in space, and throughout the duration of the debris flow in time. Rearranging some terms, these equations become:

$$\int_{t_k^{ini}}^{t_k^{end}} \int_{X_k^{ini}}^{X_k^{end}} \left(\frac{\partial}{\partial t} \left((1-c)h \right) + \frac{\partial}{\partial x} \left((1-c)uh \right) \right) dxdt = \int_{t_k^{ini}}^{t_k^{end}} \int_{X_k^{ini}}^{X_k^{end}} \left(-(1-c_b)\frac{\partial z_b}{\partial t} + S_l \right) dxdt \tag{1.17a}$$

$$\int_{t_k^{ini}}^{t_k^{end}} \int_{X_k^{ini}}^{X_k^{end}} \left(\frac{\partial}{\partial t} (ch) + \frac{\partial}{\partial x} (cuh) \right) dxdt = -\int_{t_k^{ini}}^{t_k^{end}} \int_{X_k^{ini}}^{X_k^{end}} c_b \frac{\partial z_b}{\partial t} dxdt \tag{1.17b}$$

where the k subscript refers to the k-th transept, X_k^{ini}, X_k^{end} are the x-coordinate respectively of the transept initial and ending point and t_k^{ini}, t_k^{end} are the initial and ending times of the debris-flow. Performing some formal integrations, they can be rewritten as:

$$\left[\int_{X_k^{ini}}^{X_k^{end}} (1-c)h\, dx \right]_{t_k^{ini}}^{t_k^{end}} + \left[\int_{t_k^{ini}}^{t_k^{end}} (1-c)uh\, dt \right]_{X_k^{ini}}^{X_k^{end}} = -(1-c_b) \left[\int_{X_k^{ini}}^{X_k^{end}} z_b\, dx \right]_{t_k^{ini}}^{t_k^{end}} + \int_{t_k^{ini}}^{t_k^{end}} \int_{X_k^{ini}}^{X_k^{end}} S_l\, dxdt \tag{1.18a}$$

$$\left[\int_{X_k^{ini}}^{X_k^{end}} ch\, dx \right]_{t_k^{ini}}^{t_k^{end}} + \left[\int_{t_k^{ini}}^{t_k^{end}} cuh\, dt \right]_{X_k^{ini}}^{X_k^{end}} = -c_b \left[\int_{X_k^{ini}}^{X_k^{end}} z_b\, dx \right]_{t_k^{ini}}^{t_k^{end}} \tag{1.18b}$$

in which we have used the notation that $[\mathcal{X}]_a^b$ represents the difference of the generic quantity $\mathcal{X}$ evaluated in b minus the same quantity evaluated in a.

The first term of each equation represents the time variation of respectively the liquid and the solid volume flowing inside the k-th transept. This term can be neglected under the assumption that $[(1-c)h]$, for the first equations, and $[ch]$, for the second one, are equal at the initial and ending times on the boundaries of each transept. This may not be true in some

cases, but we think that in the framework of a conceptual scheme, this approximation is acceptable.

The second term of each equation represents the difference of the fluxes integrated in time (namely, the volumes) that leave and enter the k-th transept in the given time interval. Indicating with:

$$\left[\int_{t_k^{ini}}^{t_k^{end}} (1-c)uh\,dt\right]_{X_k^{end}} = \mathcal{V}_k^l; \quad \left[\int_{t_k^{ini}}^{t_k^{end}} cuh\,dt\right]_{X_k^{end}} = \mathcal{V}_k^s$$

the outflow volumes of the liquid and the solid phase in the k-th transept, these second terms can be written, respectively, as:

$$\begin{cases} \mathcal{V}_k^l - \mathcal{V}_{k-1}^l \\ \mathcal{V}_k^s - \mathcal{V}_{k-1}^s \end{cases}$$

in which the inflow volumes of the k-th transept are expressed as outflow volumes of the $(k-1)$-th one.

The first term on the right hand side of each equation represents the volumes of liquid and solid released or stored in the bed because of the bed evolution. Since c_b is constant in time and space, they can be written as a function of the volume of bed variation in the transept, defined as:

$$\left[\int_{X_k^{ini}}^{X_k^{end}} z_b\,dx\right]_{t_k^{ini}}^{t_k^{end}} = V_k^{bed}$$

namely:

$$\begin{cases} (1-c_b)\,V_k^{bed} \\ c_b V_k^{bed} \end{cases}$$

The last term in the liquid mass-balance equation represents the source term for the liquid phase, i.e. the possible volume of rain. We indicate it as:

$$\int_{t_k^{ini}}^{t_k^{end}} \int_{X_k^{ini}}^{X_k^{end}} S_l\,dxdt = V_k^l$$

Considering the previous expressions, eq. (1.18a - 1.18b) can be rewritten in the following compact way:

$$\mathcal{V}_k^l - \mathcal{V}_{k-1}^l = -\,(1-c_b)\,V_k^{bed} + V_k^l \tag{1.19a}$$

$$\mathcal{V}_k^s - \mathcal{V}_{k-1}^s = -c_b V_k^{bed} \tag{1.19b}$$

Since the bed variation and the source term are not present in each transept, it is useful to write explicitly for each reach, the relevant equations according to the assumed conceptual scheme.

Transept $k = 1$: here we have no upstream input therefore $\mathcal{V}_0^{l,s} = 0$; we indicate with V_{rain} the liquid source V_1^l; finally, since no bed variation is present, $V_1^{bed} = 0$, then, the equations for this transept become:

$$\mathcal{V}_1^l = V_{rain} \tag{1.20a}$$
$$\mathcal{V}_1^s = 0 \tag{1.20b}$$

Transept $k = 2$: indicating with $-V_{ero} = V_2^{bed}$ the bed volume variation connected to the erosion (with $V_{ero} \geq 0$), we can write:

$$\mathcal{V}_2^l - \mathcal{V}_1^l = (1 - c_b)V_{ero} \tag{1.21a}$$
$$\mathcal{V}_2^s = c_b V_{ero} \tag{1.21b}$$

In the outflow section of this transept, according to the assumptions, the concentration is constant, and given by eq. (1.15). Indicating with $\hat{c}$ this concentration, the outflow volumes can be rewritten as:

$$\mathcal{V}_2^l = (1 - \hat{c}) \left[\int_{t_k^{ini}}^{t_k^{end}} uh \, dt \right]_{X_k^{end}} \quad ; \quad \mathcal{V}_2^s = \hat{c} \left[\int_{t_k^{ini}}^{t_k^{end}} uh \, dt \right]_{X_k^{end}}$$

We can express these volumes as a function of the outflow volume of the mixture

$$\mathcal{V}_2^{mix} = \mathcal{V}_2^l + \mathcal{V}_2^s = \left[\int_{t_k^{ini}}^{t_k^{end}} (uh) \, dt \right]_{X_k^{end}}$$

in the following way:

$$\mathcal{V}_2^l = (1 - \hat{c})\mathcal{V}_2^{mix} \tag{1.22a}$$
$$\mathcal{V}_2^s = \hat{c}\mathcal{V}_2^{mix} \tag{1.22b}$$

and finally eq. (1.21) can be reformulated in the following form that will be used further on:

$$(1-\hat{c})\mathcal{V}_2^{mix} - \mathcal{V}_1^l = (1-c_b)V_{ero} \tag{1.23a}$$

$$\hat{c}\mathcal{V}_2^{mix} = c_b V_{ero} \tag{1.23b}$$

Transept $k = 3$**:** here, V_{dep} denotes the bed volume variation connected to deposition, namely $V_{dep} = V_3^{bed}$ (with $V_{dep} \geq 0$). Moreover, since all the sediments stop in this reach according to the assumptions, $\mathcal{V}_3^s$ must be null. Therefore, the resulting equations are:

$$\mathcal{V}_3^l - \mathcal{V}_2^l = -(1-c_b)V_{dep} \tag{1.24a}$$

$$-\mathcal{V}_2^s = -c_b V_{dep} \tag{1.24b}$$

Combining the equations written for each transept, it is now possible to obtain the following useful relations.

BDA relation: substituting eq. (1.21b) in eq. (1.24b) it follows:

$$V_{ero} = V_{dep} \tag{1.25}$$

Moreover, we can derive $\mathcal{V}_2^{mix}$ from eq. (1.23b)

$$\mathcal{V}_2^{mix} = \frac{c_b}{\hat{c}} V_{ero} \tag{1.26}$$

and substitute it in eq. (1.23a):

$$(1-\hat{c})\frac{c_b}{c} V_{ero} - \mathcal{V}_1^l = (1-c_b)V_{ero}$$

By using eq. (1.25) and (1.20a) it follows:

$$\left[(1-\hat{c})\frac{c_b}{\hat{c}} - (1-c_b)\right] V_{dep} = V_{rain}$$

that can be simplified to the final form:

$$V_{rain} = \frac{c_b - \hat{c}}{\hat{c}} V_{dep} \tag{1.27}$$

This relation is the explicit expression of eq. (1.10) and, despite its simplicity, it represents a key element in our approach.

Volume of erosion: by using the previous relation and eq. (1.25), we obtain:

$$V_{ero} = \frac{\hat{c}}{c_b - \hat{c}} V_{rain} \tag{1.28}$$

which links the eroded volume reaching the deposition fan to the relevant volume of rain.

Amplification of the debris-flow volume respect the rain volume: in eq. (1.23a) we can substitute the expression of V_{ero} obtained from eq. (1.23b)

$$V_{ero} = \frac{\hat{c}}{c_b} \mathcal{V}_2^{mix} \tag{1.29}$$

and $\mathcal{V}_1^l$ from eq. (1.20a), obtaining:

$$(1 - \hat{c})\mathcal{V}_2^{mix} - V_{rain} = (1 - c_b)\frac{\hat{c}}{c_b}\mathcal{V}_2^{mix}$$

It is then possible to derive the relation

$$\mathcal{V}_2^{mix} = \frac{c_b}{c_b - \hat{c}} V_{rain} \tag{1.30}$$

that is nothing but the well-known volumetric amplification relation obtained by Takahashi (2007) starting from a quite different context: it expresses the volume of debris flow as a function of the volume of rain and of the equilibrium concentration in a reach with a given slope. It is worth noting that the present derivation allows to appreciate the numerous assumptions underlying this expression.

The liquid volume leaving the domain: by using eqs. (1.20a) and (1.21a) we obtain:

$$\mathcal{V}_3^l = V_{rain} + (1 - c_b)V_{ero} - (1 - c_b)V_{dep} \tag{1.31}$$

and using relation (1.25), it becomes:

$$\mathcal{V}_3^l = V_{rain} \tag{1.32}$$

This expresses the somewhat (a posteriori) obvious statement that the liquid volume leaving the domain is equal to the liquid volume entering upstream.

1.3.4 The role of the rain volume in a debris-flow

A by-product of the approach just presented is the explanation of the role of the rain volume in the simplified dynamics of a debris flow. In a Lagrangian framework, the rain volume firstly generates the hydrological flow, then induces a volume of erosion given by eq. (1.28), conveys downstream the sediments as a mixture whose volume is given by eq. (1.30) and finally, after having deposited the eroded volume, eq. (1.25), leaves the domain, eq. (1.32).

1.4 Probability threshold for intensity-duration data: the frequentist method.

It is widely accepted (see e.g. Brunetti et al., 2010, Nikolopoulos et al., 2014, Peruccacci et al., 2012) that the relation between rainfall intensity I and duration D is a power law of type:

$$I(D) = \alpha D^{-\beta} \tag{1.33}$$

where α and β are two constant parameters which are commonly estimated by using least square method (or other statistical approaches) starting from a significant set of couples (I_k, D_k). In this work, we assume that the same type of relation is valid also for the intensities and durations defined in the previous section, i.e.

$$\hat{I}(\hat{D}) = \hat{\alpha}\hat{D}^{-\hat{\beta}} \tag{1.34}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are two constant parameters similar to α and β, and estimated by using the BDA couple set $(\hat{I}_k, \hat{D}_k)$.

The approach we have chosen in this work to estimate the probability threshold is the frequentist method (see e.g. Peruccacci et al., 2012, among others), nevertheless other methods could be applied as well (see e.g. Berti et al., 2012, Peres and Cancelliere, 2014). For sake of completeness, we present here a summary of the methodology applied to the $(\hat{I}_k, \hat{D}_k)$ couples, but the procedure applies to (I_k, D_k) as well. We address the reader to the mentioned paper for more details.

According to this approach, the threshold is a curve that, in a log-log plot, is a straight line parallel to eq. (1.34) but with a lower value of the intercept such that the probability that the measured real event data exceeds the threshold value is a given value.

In order to obtain this threshold line, firstly the following difference set must be evaluated:

$$\delta(\hat{D}_k) = \log \hat{I}_k - \log \hat{I}(\hat{D}_k), \quad k = 1, \ldots, N \tag{1.35}$$

where $\hat{I}_k$ is the value of the k-th registered datum associated to the k-th duration $\hat{D}_k$, $\hat{I}(\hat{D}_k)$ is computed by means of eq. (1.34) and N is the total number of registered events. Then, the probability density function of this set is approximated by using a Kernel Density Estimation as proposed by Silverman (2018). Afterwards, this last function is sampled at regular intervals and the resulting set of values are used to estimate the parameters of a normal distribution:

$$f(\delta) = \frac{1}{2\pi\sigma} \exp\left(-\frac{(\delta - \mu)^2}{2\sigma^2}\right) \tag{1.36}$$

namely the mean μ and the variance σ. Finally, the previous distribution is integrated from $-\infty$ up to a threshold value δ_x such that the relevant probability of non-exceedance is equal to x. Commonly, this value is set equal to 5%. Therefore, the threshold curve can be written as:

$$\hat{I}_{5\%} = \hat{\alpha}_{5\%} \hat{D}^{-\hat{\beta}} \tag{1.37}$$

where $\hat{\alpha}_{5\%} = \hat{\alpha} - |\delta_{5\%}|$.

1.5 The rainfall thresholds for a sample study area

In order to verify the feasibility of the proposed method and to highlight the peculiarities of BDA approach respect to the classical one considered in this work, we have calculated the rainfall thresholds according to the procedures above described. The study area is Trentino-Alto Adige/Südtirol, located in the Alps, in the north of Italy (Figure 1.4).

1.5.1 Available data

Between 2006 and 2016, 161 debris flows were recorded by the regional agency. The average area of the catchments affected by the events is nearly $1.5\,\mathrm{km}^2$. Almost none of them has a precise indication about the triggering area while for 139 cases, a quantitative estimate of the deposited volume is available. It ranges from $100\,\mathrm{m}^3$ of the smallest events up to $50,000\,\mathrm{m}^3$ of the largest ones. In Fig. 1.5, the spatial distribution of the debris-flow events is reported on a geological map of the study area. It must be emphasized that all the events, even if occurring in geologically different areas, were characterized by a loose stony nature, and devoid of any significant presence of cohesive mud.

Figure 1.4: Location of Trentino-Alto Adige/Südtirol region (Italy), the sample study area.

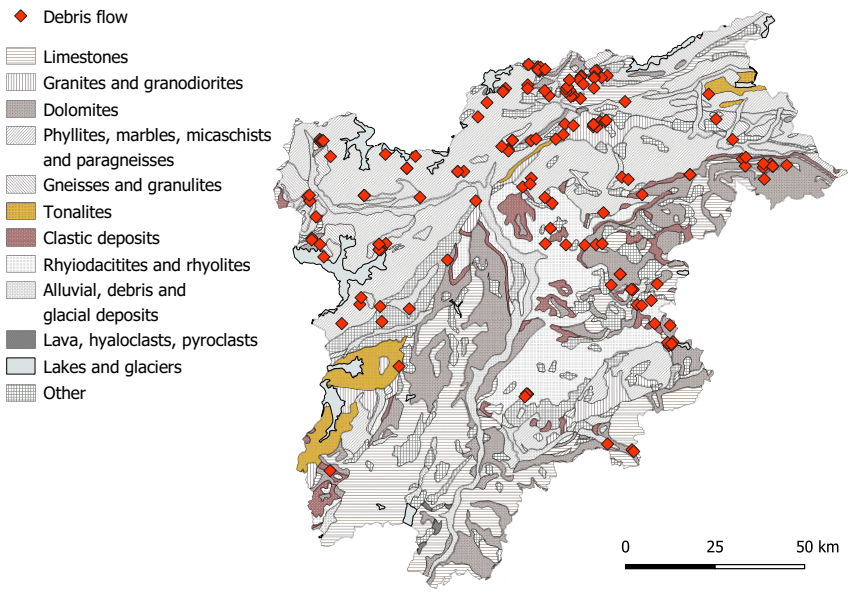

Figure 1.5: Simplified geological map of the study area with indication of the locations of debris-flow events.

A network of 195 rain gauges is available with an average spatial density of approximately $1/70\,\mathrm{km}^{-2}$ and an average altitude of $1,400\,\mathrm{m}$ asl. Nearly each rain gauge used in this study has a record frequency between 5 min and 10 min.

For the same period, records of a C-band Doppler weather radar are available. The radar is located in a central position of the region on the top of Macaion peak at $1,866\,\mathrm{m}$ asl (see Fig. 1.6), and it is effectively used to monitor an area within the range of 120 km. Precipitation is estimated converting the reflectivity Z into intensity of precipitation I (see e.g. Uijlenhoet, 2001). The radar output data are available over a square grid cell with resolution of 500 m, while the temporal resolution is between 5 min and 6 min.

Since generally, the cumulated rainfall depends on the altitude and rapidly decreases with the increase of the distance from the event area (Marra et al., 2016), in order to get reliable rain data relevant to a debris flow, a careful choice among the available rain gauges has been performed. For each debris flow event, the choice of the representative station was obtained by using an automatic procedure. From the knowledge of the coordinates of a debris-flow basin closure point, assumed located where the deposition starts, the horizontal distance between the closure point and each rain gauge was evaluated. Among all the available rain gauges with a distance less or equal to 5 km, the representative station was assumed the one with the smallest altitude difference respect the closure point. If no instrument matched the distance criteria, the debris-flow event has been considered unserviceable for the subsequent analysis.

Rain gauge data has been used only for the CD estimation, while radar data has been used to estimate the (I, D) and $(\hat{I}, \hat{D})$ couples. Nevertheless, because of the complex topography of the region, mountain beam shielding causes, in certain areas, a lack of measurement associated with a reported debris flow event (see Fig. 1.6). Therefore we have been preliminary filtered out all the events located in places where the radar signal is weakened more than 90%. Other sources of errors, such as signal attenuation in heavy rain or wet radome attenuation (see e.g. Marra et al., 2014), have not been taken into account.

1.5.2 The CDM-based threshold

For this site, 45 rain gauges, linked to the 161 debris-flow available events, have been used to determine the monthly CD reported in Table 1.1 and plotted in Fig. 1.7.

The CDM-based threshold has been obtained by using only 109 regis-

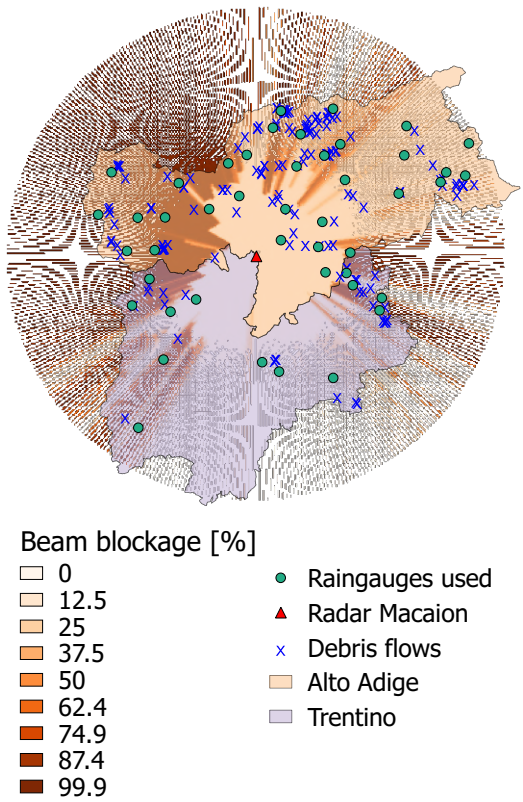

Figure 1.6: Percentage of signal attenuation due to beam blocking effect on the debris flows considered with radar beam elevation at $1°$.

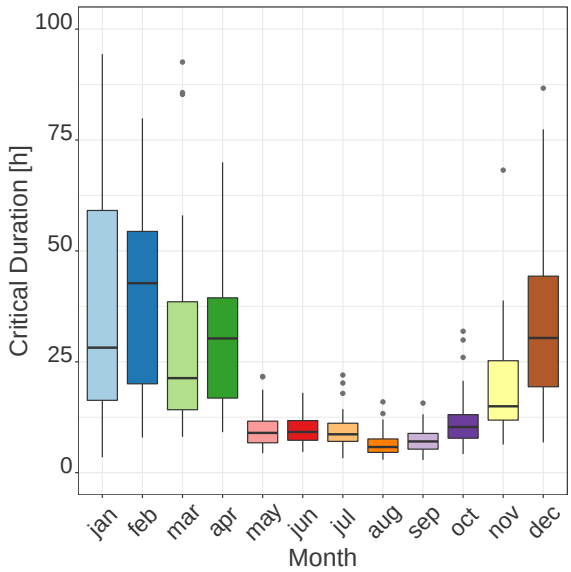

Figure 1.7: Box plot of the monthly CD for the study site.

month	1st quartile	median	3rd quartile
Jan	16.5	29.6	63.5
Feb	20.0	42.7	54.4
Mar	14.5	21.4	38.7
Apr	16.8	30.3	39.4
May	6.8	9.0	11.8
Jun	7.3	9.2	11.7
Jul	7.1	8.7	11.1
Aug	4.6	5.8	7.6
Sep	5.3	7.0	8.9
Oct	7.8	10.3	13.1
Nov	12.0	15.0	26.4
Dec	19.4	30.4	44.3

Table 1.1: Values of the monthly CD, expressed in hours, for the study case.

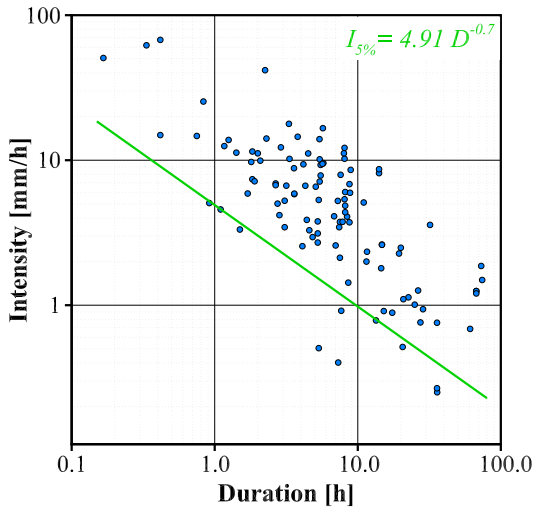

Figure 1.8: The CDM-based rainfall threshold for the study area.

tered events (out of the 161 registered) matching the 5 km distance and the radar signal criteria. The resulting threshold equation is:

$$I_{5\%} = 4.91 D^{-0.7} \tag{1.38}$$

and is plotted in Fig. 1.8 along with the relevant (I, D) event couples.

1.5.3 The BDA-based threshold

The BDA methodology requires the knowledge of the volume of deposited material (as described in Section 1.3). Due to this constrain, only 84 debris flows of the 109 registered events with unshielded radar data are serviceable. For each debris-flow case, the couple $(\hat{I}, \hat{D})$ was obtained using the following procedure:

1. the relevant basin was extracted from a DTM, setting the closure at the beginning of the deposition zone;

2. a representative bed slope i_f was considered as the average slope of the last 50 m of the basin network;

3. the debris-flow reference concentration $\hat{c}$ was evaluated by means of eq. (1.15), where we set $\Delta = 1.65$ and $\psi = 35\hat{\text{A}}°$;

4. the volume of rain that has caused the deposition was estimated by using eq. (1.27)

5. the duration $\hat{D}$ should have been evaluated by solving eq. (1.12); nevertheless, since intensity is actually a piece-wise constant function whose constancy interval is equal to the radar sampling interval δt, the equation cannot be solved exactly. Therefore, the procedure used to obtain $\hat{D}$ was the following:

   - consider the discrete hydrograph pertaining a debris flow event, place the origin of a discrete reference system in the interval where the intensity is maximum and label this interval as i_0;
   - consider the following sequence of discrete integrals of the hydrograph:

$$a_0 = i_0 \delta t$$
$$a_n = a_{n-1} + \begin{cases} i_{n/2+1/2}\,\delta t & if \quad n \text{ is odd} \\ i_{-n/2}\,\delta t & if \quad n \text{ is even} \end{cases}$$

     where $i_{\pm k}$ is the intensity of the $\pm k$-th interval located on the right and on the left of i_0 respectively;
   - consider the integer K as the first integer such that

$$V_{rain} \leq A_b a_K$$

   Then, since a_K is nothing but the approximation of the integral term of eq. (1.12), it follows that:

$$\hat{D} = (K+1)\,\delta t$$

6. the intensity $\hat{I}$ was obtained by using eq. (1.14)

The resulting BDA-based threshold equation, plotted in Fig. 1.9 along with the relevant $(\hat{I}, \hat{D})$ event couples, is:

$$\hat{I}_{5\%} = 6.2\hat{D}^{-0.67} \tag{1.39}$$

1.6 Discussion

In this Section, we discuss some aspects of the results obtained for the sample study area.

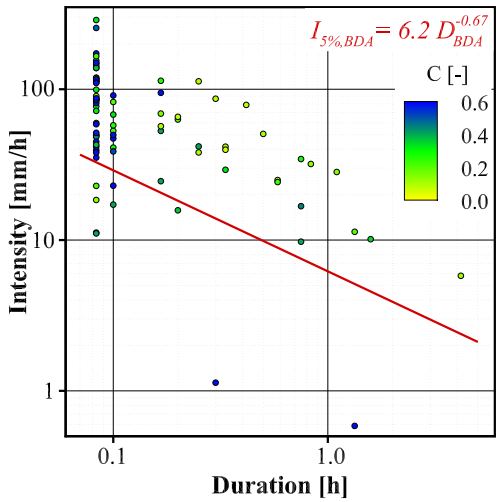

Figure 1.9: BDA-based rainfall threshold for the study area. The color scale indicates the equilibrium concentration, evaluated according to eq. (1.15), for each event.

1.6.1 The CDM-based threshold

The threshold obtained with this method, eq. (1.38), is similar to the ones obtained by Marra et al. (2014) and by Iadanza et al. (2016) who considered, for areas comparable to the one considered in this paper, rain data associated not only to debris flows but also to landslides. Moreover, it can be noted that in Fig. 1.8, even if data is distributed over almost three orders of magnitude, the largest number of values are approximately in the interval $[1\,\mathrm{h}, 10\,\mathrm{h}]$.

Regarding the monthly CDs, it is important to notice, Fig. 1.7, the dispersion of the summer months values is smaller than the dispersion of the other months. This means that for the summer period, the median value of CD is representative of the whole correspondent area. This is quite important because in this period the majority of the debris-flow events occur. Moreover, the summer CDs are shorter than in the rest of the year. This is due to the different structure of the summer storms with respect to the rainy systems of the other months. Similar trend is reported in Iadanza et al. (2016), where an area analogous to the one considered in this paper was investigated.

1.6.2 The BDA-based threshold

First of all, the assumption that the intensity $\hat{I}$ is a power law function of $\hat{D}$ (see sec. 1.4) seems to be confirmed by results, since data in Fig. 1.9 shows a clear linear trend in the log-log plot.

The data is distributed essentially over one scale of magnitude, ranging from 0.1 h to 1 h. This time scale is confirmed by both eyewitness and video testimonies (some of them can be found on the web) even if, a systematic and well documented study is still not available. However, this result allows a first assessment of the BDA relation reliability. On the other hand, the BDA relation is strictly connected to the Takahashi volumetric amplification relation, whose validity is widely accepted and assessed (see e.g. Rosatti et al., 2015). Therefore, the BDA relation appears to be a sufficiently robust estimator, provided that measured data is sufficiently accurate.

An accurate analysis of the result (still Fig. 1.9) showed several points characterized by a duration equal to the sampling interval of the radar (leftmost points) and this can be due to multiple reasons.

One of these can be the sampling interval is too long with respect to the time necessary to feed a debris-flow.

A second one is an underestimation of V_{dep}, whether for measurement errors or because only part of the debris-flow sediments stopped in the fan, while another significant part is not included in the measurements because it has been carried downstream by the flow. Therefore, the relevant V_{rain} is smaller than the actual value and consequently, the duration is shorter than the real one.

A third possible reason is an overestimation of c_{ref} due to a wrong estimation of a significant slope. This produces an underestimation of V_{rain}, as it can be deduced from eq. (1.27) and plotted in Fig. (1.10). It follows an underestimation of the event duration.

Last but not least, despite V_{dep} is essentially correct, the debris flow has not reached the hypothesized equilibrium condition because of lack of sediments, non-erodible zones etc. Therefore, also in this case c_{ref} is overestimated with respect to the actual value and the event duration is underestimated.

We are not able to quantify the errors in our data, or to verify their frequency distribution. Anyhow, we think that the last three reasons are probably the most diffuse and possibly mingled. In any case, we think that our sample is sufficiently reliable. Finally, we are not able even to predict the change in the threshold if more reliable data would be available.

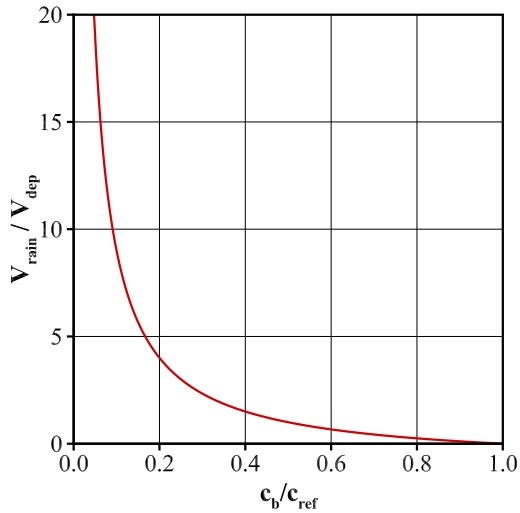

Figure 1.10: Trend of the ratio of the volume of rain over the deposited volume as a function of the ratio of the bed concentration over the reference concentration.

1.6.3 Comparison between the CDM and the BDA-based thresholds

A straightforward comparison between the two thresholds, namely eq. (1.38) and eq. (1.39), shows that the BDA-based threshold is, as expected, higher than the CDM-based one (see Fig. 1.11). Nevertheless, we cannot conclude that BDA is a better estimator, since we are comparing quite different approaches and quantities and therefore, some care must be paid in making the comparison.

Since it is not completely right extending the validity of each threshold outside the domain used for the interpolation, there is only a limited overlapping of the two domains. In this range, the relative difference with respect to the CDM data spans from 25% to 31%.

The characteristic time scale of the two approaches are quite different: $[1\,\mathrm{h}, 10\,\mathrm{h}]$ for CDM and $[0.1\,\mathrm{h}, 1\,\mathrm{h}]$ for BDA. Namely, BDA indicates as potential debris-flow inducing rainfalls, short or very short storm durations while CDM suggests quite longer durations. This is not surprising since, as we have already stressed, we are comparing different things, deriving from different points of view. In order to highlight this difference, it is useful to compare graphically (see Fig. 1.12), how the same hyetograph is considered by the two approaches. It is quite clear how the same event is characterized

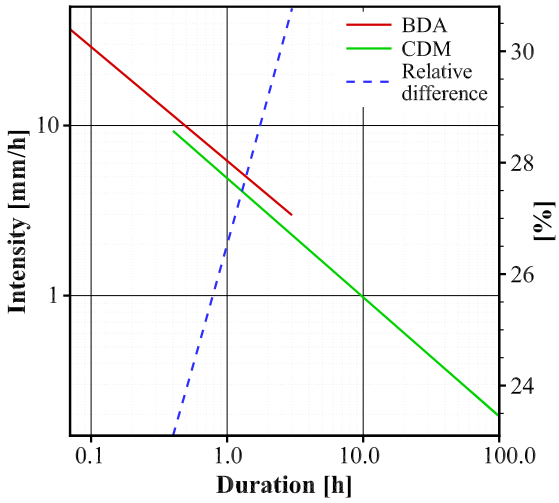

Figure 1.11: Comparison between the CDM- and the BDA-based rainfall thresholds for the study area and their relative difference respect the BDA values.

by different duration and average intensity, and why BDA provides shorter duration and larger intensities with respect to CDM.

1.6.4 Some issues regarding the forecast use of the BDA-based threshold

The forecast use of the BDA-based thresholds presents some potential issues.

Starting from a forecast hyetograph, while the peak intensity can be easily identified, the potential volume of rain that determines both the duration and the average intensity, is not known. A possible empirical choice, that can be deduced from the available data but that cannot be used in general, is to consider an average duration of $0.5\,\mathrm{h} \div 1.0\,\mathrm{h}$ centred around the peak.

It is not so uncommon in debris-flow events that the relevant hyetograph shows multiple peaks (see e.g. Rosatti et al., 2015). In these cases, it is even more complicated estimating a reference duration because the previous criterion could be largely inaccurate. Moreover, it is quite difficult to forecast if multiple peaks generate multiple events or a single long event.

A validation of the proposed reference duration choice, the development of more sophisticate approaches and the evaluation of the frequency of false positive respect the proposed BDA-based threshold are very interesting top-

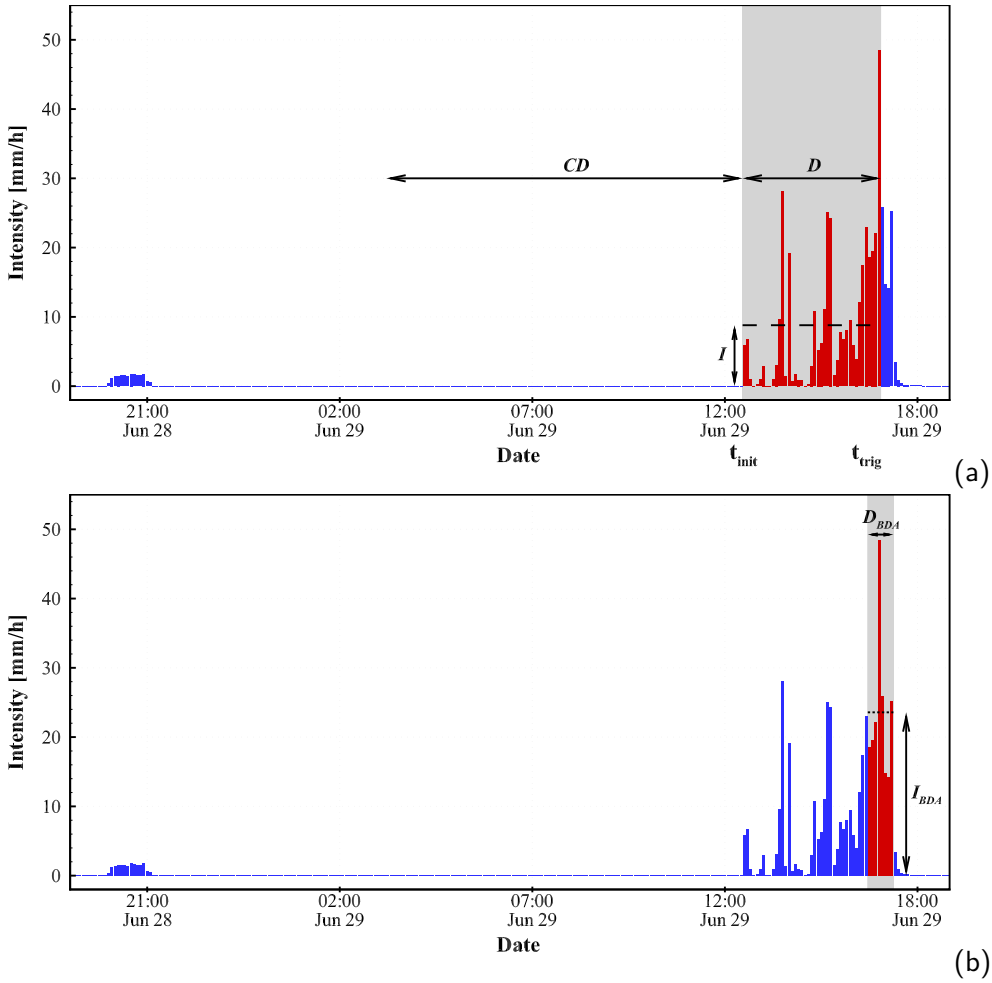

Figure 1.12: Portion of hyetograph relevant to a debris flow event according to (a) CDM, (b) BDA.

ics, but they are beyond the scope of this work and therefore are left for a future widening. Last, but not least, BDA focuses the attention only on one mandatory ingredient for a debris flow: the amount of water necessary to carry downstream the sediments. The other key element, the saturation of the basin, is completely disregarded in this approach and therefore, a more complete methodology considering both the ingredients is desirable.

1.7 Conclusions

The conclusions that can be drawn from this study are the following:

- The Backward Dynamical Approach presented in this paper appears to be a reasonable theoretical framework able to single out the amount of rain strictly pertaining a debris-flow event and, consequently, the related $(\hat{I}, \hat{D})$ couple. Other less simplified relations can be obtained relaxing one or more assumptions herein considered.

- Every method of extracting a threshold from a (I, D) couple set can be applied to the BDA couples as well.

- Analysis of a sample study area shows that the characteristic duration of the rain generating a debris-flows evaluated with BDA is one order of magnitude shorter than the characteristic duration given by CDM. This result appears to be consistent to several field observations, but a systematic validation is still missing because of lack of reliable data.

  Moreover the BDA-based threshold is, as expected, higher than the CDM-based one, provided that the same methodology is used to obtain the thresholds from the intensity-duration couples. Even if this result seems to overcome a bit the limit of the traditional approaches, it is not yet possible to state that this method performs better than the literature ones because a systematic validation is still not possible, once again, for lack of measured data.

- The use of the BDA-based thresholds for forecast purpose presents some potential issues. Moreover, the approach does not account for the reaching of any "triggering condition" (e.g. saturation of the basin, threshold condition for transport, abatement of possible superficial sediment cohesion, etc.) and for the effect of earthquakes on these conditions (see e.g. Shieh et al., 2009, Tang et al., 2009). A more complete, and presumably more reliable methodology for the rainfall threshold determination should consider, in some way, not only the

characteristics of the rain strictly pertaining to the debris flow but also the characteristics of the rain leading to the triggering conditions and at least some features of the sediments.

- To fully validate the proposed method, large samples of data from different areas should be considered. Our forthcoming work aims to achieve this goal. Unfortunately, the number of well-documented debris-flow events are rather few, not just because debris flows are infrequent, but also because so far the detailed survey of the deposited volumes was extremely difficult and not considered so important by the public agencies in charge of debris-flow data collection. Nevertheless, in the last years application of drone technology to land surveying has made it possible to simplify the measurement of debris-flow deposits.

We are confident that in some years the number and quality of available data should help to validate both the reliability of the proposed method and its possible future enhancements.

Chapter 2

Uncertainty analysis of a rainfall threshold estimate for stony debris flow based on the Backward Dynamical Approach

2.1 Introduction

In mountain regions, rainfall-induced natural phenomena, as shallow landslides and debris flows, are relatively frequent events that have a significant impact on the territory in which they occur, causing damages and, in some cases, casualties (Cánovas et al., 2016, Dowling and Santi, 2014, Fuchs et al., 2013). The risk management of these phenomena is crucial to reduce their effects on the territory and it is based on both active and passive mitigation strategies. 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.

The early warning systems for these phenomena are mainly based on rainfall thresholds (Chien-Yuan et al., 2005, Segoni et al., 2018), namely rainfall conditions beyond which the occurrence probability of a rainfall-induced event is considered significant. In this framework, most rainfall thresholds are power-law relations expressing the rainfall event cumulated or intensity as a function of the event duration (Segoni et al., 2018). A considerable literature deals with this topic (e.g. Caine, 1980, Guzzetti et al., 2008, Iadanza et al., 2016, Jakob et al., 2012a, Marra et al., 2014, 2016, Pan et al., 2018, Staley et al., 2013, Winter et al., 2010, Zhou and Tang, 2014).

In some studies rainfall thresholds concern a wide typology of phenomena (Segoni et al., 2018), other works focus on both shallow landslides and debris flows (e.g. Baum and Godt, 2010, Cepeda et al., 2010), other on shallow landslides (e.g. Frattini et al., 2009, Giannecchini, 2005) and, finally, some studies are specifically conceived for debris flow (e.g. Giannecchini et al., 2016, Li et al., 2016, Nikolopoulos et al., 2014).

Power-law thresholds can be derived in the following way. Given a historical dataset of rainfall-induced events, the rainfall associated with each event is determined and described in terms of the couple of synthetic quantities employed in the threshold (e.g. rainfall event cumulated - event duration). Classically, these quantities are defined only on the basis of a hyetograph analysis (Segoni et al., 2018), without considering the characteristics of the rainfall-induced phenomenon. In a log-log plane, the resulting set of couples becomes a cloud of points and the power-law function is a straight line. Starting from these couples set, the threshold is determined by locating the straight line in the log-log plane using one of the several estimate strategies available in the literature, e.g. manual methods, statistical approaches, probabilistic procedures (Guzzetti et al., 2007, Segoni et al., 2018). The result is the calibrated rainfall threshold.

One of the critical issues of the calibration is the uncertainty related to both data and models parameters (Gariano et al., 2020). Here with the term "model", we indicate generically a single equation or a set of operations that, given some input data and model parameters, provide an output. In the case of the rainfall threshold, the uncertainties derive mainly from direct data error measurements (e.g. in rainfall), from the non-unique definition of the models parameters (e.g. distance within which to select the rain gauge to define the event precipitation) and from the strategy used to calibrate the threshold. The result is an uncertainty framework that can significantly impact the threshold estimate.

Some studies have already investigated the uncertainty in threshold determination, focusing on some aspects that can affect the hyetograph or the event synthetic quantities used in the threshold. For instance, Nikolopoulos et al. (2014) has analysed the consequences of the spatial variability of the precipitation while Marra (2019) and Gariano et al. (2020) have investigated the effects of the rainfall temporal resolution. Moreover, the uncertainty arising from the choice of the reference rain gauge and the differences between the radar and the rain gauge measurements have been examined in Rossi et al. (2017). Besides, the effect of the uncertainty in triggering rainfall estimate has been investigated in Peres et al. (2018) while Abraham et al. (2020) has analysed the consequences of the scale of analysis, the rain

gauge selection and how the intensity is quantified.

Rosatti et al. (2019) has introduced an innovative method to calibrate an intensity-duration rainfall threshold for stony debris flow, a particular type of debris flow, frequent in some mountain areas as in the Alps, in which the presence of silt and/or clay in the mixture is negligible and the internal stresses are mainly caused by the collision among the particles (e.g. Bernard et al., 2019, Stancanelli et al., 2015, Takahashi, 2009). The new method, called Backward Dynamical Approach (BDA), starts from the knowledge of the volume of sediments deposited after an event and, thanks to a schematic description of the stony debris-flow dynamic, it is able to identify, in the related hyetograph, the rainfall event volume, intensity and duration strictly pertaining to the debris-flow event. Hence, the BDA differs from the classical literature approaches since the synthetic quantities describing the rainfall events are defined involving not only the forcing (i.e. the hyetograph) but also the dynamic of the rainfall-induced event.

This work focuses on the uncertainty deriving from data and parameters inherent to the BDA, leaving out the uncertainty related to the hyetograph, already investigated in the literature. In particular, the aim is to perform an uncertainty analysis on the threshold calibration to check the robustness of the BDA. To reach this goal, among the different strategies and methods available in the literature (e.g. Coleman and Steele, 2018, Helton et al., 2006, Hofer, 2018), we have chosen the Monte Carlo (MC) approach. With this tool, we have developed a proper methodology composed of two MC cascade simulations and we have applied it to a dataset concerning a specific study area. Detailed analysis of intermediate and final results have also been performed to better understand the uncertainty analysis outcomes.

The paper structure is the following. A brief description of the BDA method is presented in Sect. 2.2. The study area and data are described in Sect. 2.3. The method used to assess the uncertainty propagation in the BDA-based threshold calibration is described in Sect. 2.4. The obtained results are presented and discussed in Sect. 2.5. Conclusions end the paper.

2.2 The BDA-based threshold calibration

As mentioned in the Introduction, the BDA determines the rainfall event intensity and duration, namely the couple (I, D), associated with a stony debris flow, by using not only the hyetograph but also information concerning the occurred debris flow.

The BDA starts from the knowledge of the deposited volume V_{dep} occupied by the sediments after a debris-flow event. Thanks to a simplified

global volumetric description of the debris-flow dynamic (Fig. 2.1), the rainfall volume pertaining to the debris flow V_r^{DF}, defined as the volume of water necessary to convey downstream V_{dep} as a mixture, can be express as:

$$V_r^{DF} = \frac{c_b - c}{c} V_{dep} \tag{2.1}$$

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.

The expression of Takahashi (1978) is valid in permanent and uniform conditions and it can be used as reference concentration:

$$c = \min\left(\frac{i_f}{\Delta\left(\tan\psi - i_f\right)},\, 0.9\, c_b\right) \tag{2.2}$$

where i_f is the bed slope, 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. Δ is constant and assumed equal to 1.65 (e.g. Prancevic and Lamb, 2015). According to the assumptions of the BDA, the reference concentration is evaluated considering the bed slope in the last portion of the debris-flow channel, just upstream of the deposition area. This means that the information concerning the triggering conditions and the detailed evolution of the debris flow in the upper part of the basin are not considered.

The rainfall volume pertaining to the debris flow can also be expressed as the product of the rainfall volume per unit area E and the event basin area A_b:

$$V_r^{DF} = E A_b \tag{2.3}$$

from which, a backward dynamical expression for the rainfall volume per unit area can be obtained by equating (2.1) with (2.3):

$$E = \frac{1}{A_b} \frac{c_b - c}{c} V_{dep} \tag{2.4}$$

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

$$E = \int_{t_1}^{t_2} i(t)\, dt \tag{2.5}$$

where $i(t)$ is the measured rainfall intensity and t_1 and t_2 are the unknown start and end times related to the debris-flow duration. In the absence of

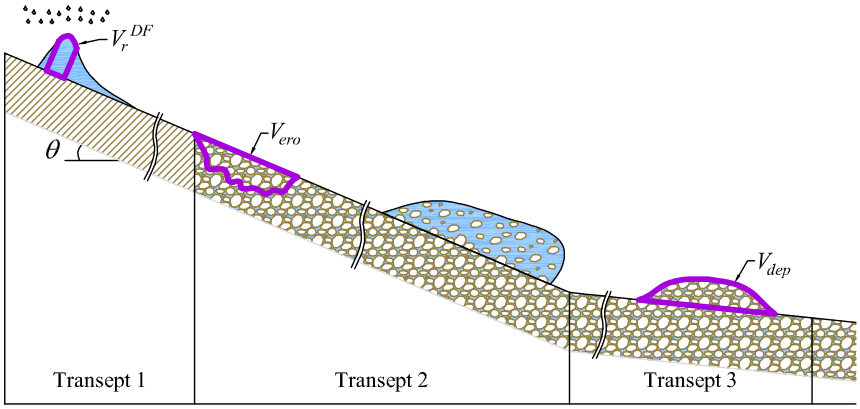

Figure 2.1: Conceptual Lagrangian volumetric description of debris-flow dynamic from Rosatti et al. (2019). The scheme is divided into three transepts: transept 1 is characterised by the run-off formation; the bed material erosion and the achievement of equilibrium conditions occur in transept 2; transept 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 $\theta = \arctan(i_f)$ is the inclination angle of the bed with respect to a reference horizontal direction.

detailed data of the event, these times are expressed as:

$$\begin{cases} t_1 = t_{max} - \Delta t_1 \\ t_2 = t_{max} + \Delta t_2 \end{cases} \tag{2.6}$$

where t_{max} is the instant of maximum intensity during the event and Δt_1 and Δt_2 are unknown intervals. These intervals can be obtained equating the right hand side terms of Eq. (2.5) and (2.4):

$$\int_{t_{max}-\Delta t_1}^{t_{max}+\Delta t_2} i(t)\, dt = \frac{1}{A_b} \frac{c_b - c}{c} V_{dep} \tag{2.7}$$

Because of the measurement technique, $i(t)$ is a piecewise constant function on time intervals δt, namely $i(k)$. Consequently, the reference times becomes: $t = k\delta t$, $t_{max} = M\delta t$, $\Delta t_1 = n_1 \delta t$ and $\Delta t_2 = n_2 \delta t$ where M is the number of time intervals that identifies the peak and now n_1 and n_2 are unknown integers. In addition, the integral in Eq. (2.7) must be rewritten in discrete form (namely a summation) and the previous equation cannot be satisfied exactly.

An approximation algorithm, able to determine in a univocal way the unknowns, can be introduced: starting from zero and increasing of one unit

alternatively n_1 and n_2, the first couple $\hat{n}_1, \hat{n}_2$, for which the condition

$$\sum_{k=M-\hat{n}_1}^{M+\hat{n}_2} i(k)\,\delta t \geq \frac{1}{A_b}\frac{c_b - c}{c} V_{dep} \tag{2.8}$$

is satisfied, is the searched couple. If a zero-intensity interval is reached, the sum stops being symmetrical with respect to M and only either n_1 or n_2 is increased until the previous relation is satisfied.

Finally, the duration D and the average intensity I can be expressed as:

$$D = \Delta t_1 + \Delta t_2 = (\hat{n}_1 + \hat{n}_2)\,\delta t \tag{2.9}$$

$$I = \frac{\sum_{k=M-\hat{n}_1}^{M+\hat{n}_2} i(k)\,\delta t}{D} \tag{2.10}$$

Once the (I, D) couple is computed for each event of the available dataset, the rainfall threshold is estimated by using the frequentist method (e.g. Brunetti et al., 2010, Peruccacci et al., 2012). According to this method, the (I, D) couples are plotted in a log-log ID plane and a straight line fitting these points is determined. The slope and the intercept of this straight line are the logarithms of the coefficient of the following power law:

$$I = \hat{a} D^{-b} \tag{2.11}$$

The rainfall threshold is then obtained translating vertically the straight line in the log-log ID plane so that the non-exceedance probability of the dataset events (namely the occurrence probability of debris flows related to (I, D) points located below the threshold) is equal to a given value. The final expression is:

$$I = a D^{-b} \tag{2.12}$$

in which $a < \hat{a}$.

For more details on the BDA and the frequentist method, we refer the reader to the above-mentioned references.

2.3 Study area and data

The study area and data used in this analysis are the same as those used in Rosatti et al. (2019). In particular, the study area is the Trentino-Alto Adige/Südtirol region, in the north east of the Italian Alps (Fig. 2.2(a)). The region covers 13607 km^2, has an altitude range between 40 and 3900

m a.s.l. with mean about 1600 m a.s.l. (Fig. 2.2(b)) and a climate characterised mostly by a continental regime (Bisci et al., 2004, Nikolopoulos et al., 2014).

The regional agencies between 2006 and 2016 have reported 161 debris flows (Fig. 2.2(b)) but only 139 events present the survey of the deposits, whose volumes range between 100 m^3 and 50000 m^3. In every event, sediments are characterised by the absence or, at least, the negligible presence of silt and clay thus resulting in stony debris flows.

The rainfall data related to these events derives from a radar located in a central position with respect to the region, on the Mt. Macaion at 1866 m a.s.l. (Fig. 2.2(b)). A C-band Doppler weather radar measures the reflectivity Z over an area of 120 km of radius and the rainfall is computed converting Z into precipitation intensity I (e.g. Uijlenhoet, 2001). Since radar data in mountain regions are typically affected by the beam shielding (Germann et al., 2006), which can cause errors in the measurements, the debris-flow events located in an area with a weakening of the signal greater than 90% have been excluded from the dataset. Overall, the debris flows suitable for the analysis were 84 and are highlighted in Fig. 2.2(b) with circles.

Additional data required for the BDA, namely i_f, A_b, ψ and $i(t)$, was defined for each event in the following way. The basin outlet was located downstream of a segment with a sufficiently constant slope just upstream of the deposition area and the upstream basin area was determined. Then, i_f was calculated as the mean slope of the last 50 m of the torrent upstream of the outlet point. Besides, due to the scarcity of sediments information, ψ was assumed to be equal to 35° for all the events. The hyetograph $i(t)$ was computed at each instant averaging over the respective basin area the radar intensities. In this way, both the spatial and temporal variability of the rainfall were taken into account.

Starting from this data and setting the non-exceedance probability equal to 5%, Rosatti et al. (2019) obtained the following threshold:

$$I = 6.2D^{-0.67} \tag{2.13}$$

From now on, the quantities involved in the calibration performed by Rosatti et al. (2019) will be considered as reference values and they will be indicated with a subscript r.

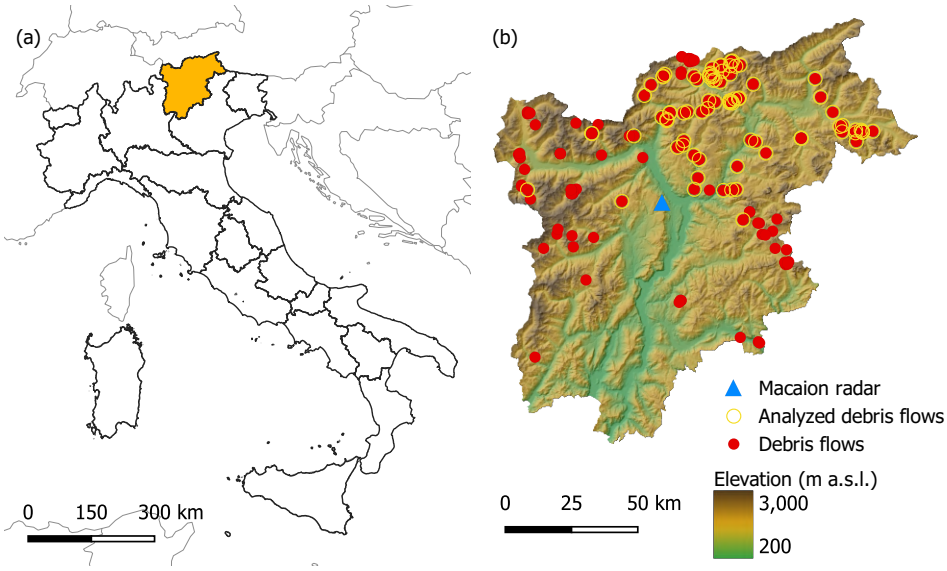

Figure 2.2: (a) Location of Trentino-Alto Adige/Südtirol region (Italy) and (b) the Macaion radar and debris-flow events: red dots show all debris flows while yellow circles highlight the suitable ones for the study.

2.4 Method

As described in Sect. 2.2, the BDA-based threshold calibration starts from the definition of the following input parameters and data for each considered event: i_f, A_b, V_{dep}, ψ and $i(t)$. Subsequently, based on these values, what we call the "event characteristics" are computed for each analysed debris flow: first c (Eq. (2.2)) and E (Eq. (2.4)), and then D (Eq. (2.9)) and I (Eq. (2.10)). Finally, the (I, D) couples of the events are used to calibrate the threshold, namely to quantify the threshold coefficients a and b of Eq. (2.12).

Coherently to the estimate procedure, the uncertainty analysis of the BDA-based threshold calibration is divided into three parts (Fig. 2.3). First, the uncertainty characterisation of the input parameters and data is determined (Fig. 2.3(a)). Then, for each debris flow, the uncertainty analysis of the event characteristics is performed with an MC simulation, starting from the above-defined uncertain quantities (Fig. 2.3(b)). Finally, a further MC simulation is carried out to perform the uncertainty analysis of the threshold, using as input the (I, D) couples of the events obtained from the first MC simulation (Fig. 2.3(c)). In this way, the impact of the uncertain parameters and data on the threshold is quantified. All the

analyses are performed using the R software (R Core Team, 2013).

Regarding the uncertainty characterisation, as explained in the Introduction, the focus of this study is on the uncertainty in the physical and morphological parameters and data used in the BDA to describe ,in a simplified way, the debris-flow dynamic. Therefore, in this analysis, the variables considered are i_f, A_b, V_{dep} and ψ. According to their estimate, described in Sect. 2.3, these variables are mainly affected by epistemic uncertainty due to measurement and estimate errors and lack of information (Oberkampf et al., 2004). The characterization of the uncertainty in the variable, namely the probability distribution function (pdf) of their values, has to be defined both in term of distribution type and statistical quantities (e.g. mean and variation coefficient CV) (Fig. 2.3(a)). Lacking certain data concerning the pdfs, according to Marino et al. (2008), all the variables are assumed to be uniformly distributed and, for each event, the means of the distributions are set equal to the corresponding reference values. Regarding the deviations from the means, ψ is the only variable whose variability is constrained by a validity range: for stony debris flow, according to Lane (1953) and Blijenberg (1995), ψ can vary between 32° and 38°. Assuming 35° as the mean of the ψ distribution, the variability range (32°, 38°) can be obtained by imposing CV equals to about 5%. The uncertainty in V_{dep} can not be accurately estimated since the survey methodology, and the related measurement errors, used by regional agencies, is not univocal (Marchi et al., 2019). However, Brardinoni et al. (2012) has proposed, for a similar study area, a relative error of 10% in the estimate of V_{dep}, namely a corresponding CV equals to about 5%. Therefore, we assume that this uncertainty value is valid for this analysis. Moreover, the uncertainty in i_f and A_b is hardly quantifiable given their computation method. For these reasons and homogeneity, the degree of uncertainty of ψ and V_{dep} is considered suitable also for i_f and A_b. The resulting uncertainty characterization is summarized in Table 2.1.

The procedure used to assess the propagation of the uncertainty in i_f, A_b, V_{dep} and ψ on the event characteristics (i.e. D, I, c and E), related to each debris flow, is schematized in Fig. 2.3(b) and it is composed of two main steps. First, the input samples, namely the ordered sets of variable values in the form (i_f, A_b, V_{dep}, ψ), must be obtained. These samples are generated by using the Latin Hypercube Sampling (LHS) (Fig. 2.3(b.1)), introduced by McKay et al. (2000). This method produces N samples starting with a division of each variable uncertainty range into N disjoint intervals of equal probability. Then, one value is randomly selected within every interval, thus obtaining N values for each variable. These values are then

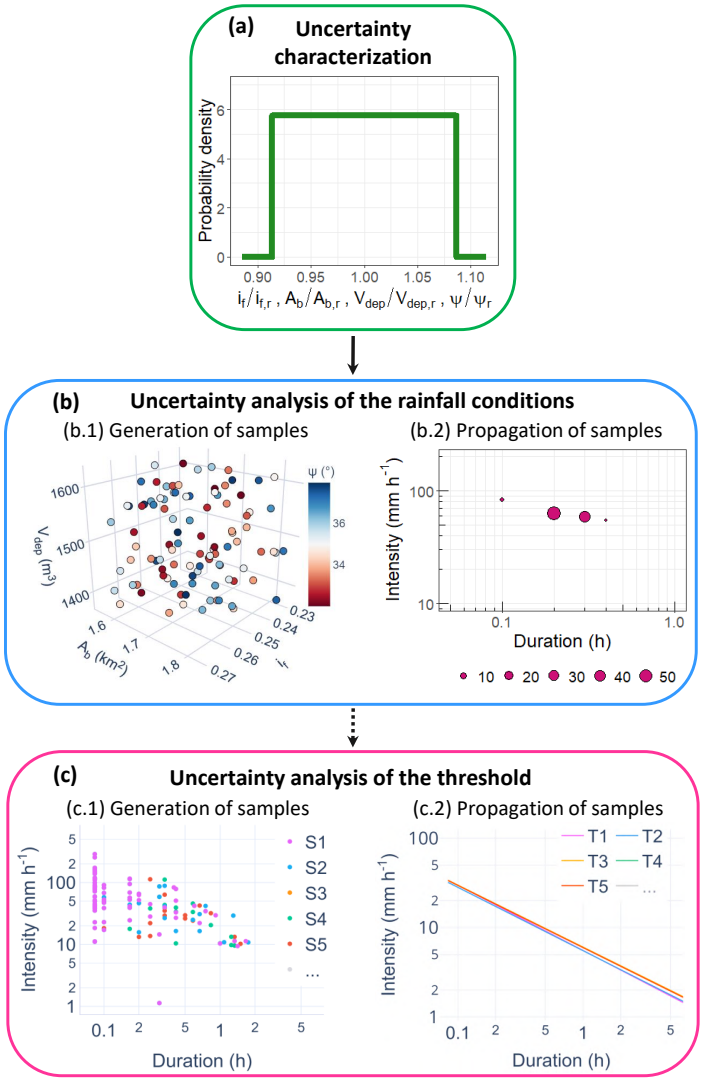

Figure 2.3: Scheme of the uncertainty analysis performed with two cascade MC simulations. (a) Uncertainty characterization of the parameters and data: the non-dimensional form of the uncertain parameters and data, obtained by dividing the variables by the reference values, are assumed to be uniformly distributed. (b) First MC simulation to compute the uncertainty analysis of the event characteristics for each debris flow: (b.1) samples generation performing the Latin Hypercube Sampling (LHS) and (b.2) propagation of samples to compute the event characteristics. The dots size in the log-log ID plane indicates the absolute frequency of obtaining the (I, D) couples. (c) Second MC simulation to perform the uncertainty analysis of the threshold: (c.1) random samples S generation (one of the previous obtained (I, D) couples for each event) and (c.2) propagation of samples to estimate the thresholds T.

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

Table 2.1: Probability distributions of the uncertain variables for each event. $i_{f,r}$, $A_{b,r}$ and $V_{dep,r}$ are the event reference values of the average slope, the basin area and the deposited sediments respectively.

arranged in the LHS matrix, composed of N rows and k columns, where k is the number of the variables (four, in the specific case). In each column, the N values relevant to a single variable are inserted in random order (Helton et al., 2006). Each row of this matrix gives one of the N variable samples. According to Marino et al. (2008), to ensure 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.
Second, the event characteristics are obtained starting from each input sample, resulting in 100 (I, D) couples (Fig. 2.3(b.2)), together with the related c and E values, for each event. Therefore, the overall total of (I, D) couples obtained is $100 \cdot 84 = 8400$, where 84 is the number of considered debris flows.

The uncertainties propagation in the threshold estimate is then quantified with a further MC procedure. In this case, a sample is generated selecting randomly one of the possible 100 (I, D) couples for each event, resulting from the previous MC simulation (Fig. 2.3(c.1)). Hence, one sample consists of 84 (I, D) couples. Following this procedure, 5000 samples are created and used to estimate as many thresholds (Fig. 2.3(c.2)), namely 5000 (a, b) couples.

2.5 Results and discussion

2.5.1 Variability of the event characteristics

As described in Sect. 2.4, the outputs of the first MC simulation applied to the dataset are 100 possible event characteristics (i.e. D, I, c and E) for each debris flow. The relative variability of all these outputs is quantified

through the computation of the CV of each event characteristics distribution. This allows providing a complete inspection and interpretation of all outputs. Then, the absolute variability is quantified through the computation of the variability range given by the difference between the minimum and the maximum values of the variable and it is evaluated only for the D and I distributions. This analysis allows highlighting the variability of the (I, D) couples in the ID plane for each event.

The CV, by definition, is a standardized measure of dispersion and allows comparing the relative variability of the results independently of their measurement units and of their means (e.g. Abdi, 2010, Håkanson, 2000). For this reason, the CV is chosen as the statistical quantity for comparing the relative variability between both the same characteristic of different events and different characteristics of the same event. The CVs of the distributions of D, I, E and c for each event are shown in Table 2.2. In the following, the trends and the differences in the CVs are highlighted and justified on the basis of some event aspects:

- the D distributions have the largest and most variable CVs with respect to all the other event characteristics: 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 30.0%. The reason for this behaviour is connected to the fact that I, by definition, is an average and therefore the effects of the variables uncertainty are smoothed by the averaging. 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. (2.2), to the value of $i_{f,r}$. As shown in Fig. 2.4, 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 greater than about 0.4 as c is always equal to $0.9\,c_b$ independently from the values of the i_f samples. However, even in the worst conditions in terms of variability, CV_c is small and reaches a maximum value of 14.3%;

- the distributions of the volumes per unit area have 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). This suggests that the relative variability in I and D does not depend only on the relative variability in the needed rainfall volume per unit area but also on how the available rainfall volume is distributed into the hyetograph time intervals.

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 MC simulations have provided always the same (I, D) couple. For these events, the propagation of the variables uncertainty does not affect the (I, D) couple estimation resulting in $CV_D = CV_I = 0\%$. This type of result is related to two conditions:

- regardless of the uncertainty of the variable, the concentration is always equal to $0.9\,c_b$. In this case, also CV_c is equal to zero (e.g. events 1, 13 and 25) and the variation in E is only due to the propagation of A_b and V_{dep} uncertainty (Eq. (2.4)), namely $CV_E \simeq 7.1\%$. For these 14 events, such a small variation in E results in 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. (2.8) is satisfied, in all the 100 simulations, considering always the same hyetograph time intervals. 34 events fall into this condition.

For the 19 events that belong to the second category, the uncertainty in the variables results in the computation of more (I, D) couples that, however, are relatively close to the mean values: the I and D distributions are characterised by a standard deviation much smaller than the mean.
The third category includes 17 events for which the uncertainty in the variables implies high values of CV_D. This means that, for these events, the number of time intervals needed to satisfy the condition of Eq. (2.8) varies greatly with respect to the mean one: the variables uncertainty has a relative great impact on the computation of D. Moreover, the highest values of CV_D highlight the presence of extremes in the D distribution, namely of

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

Table 2.2: Coefficients of variation of the event characteristics related to each debris flow 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.

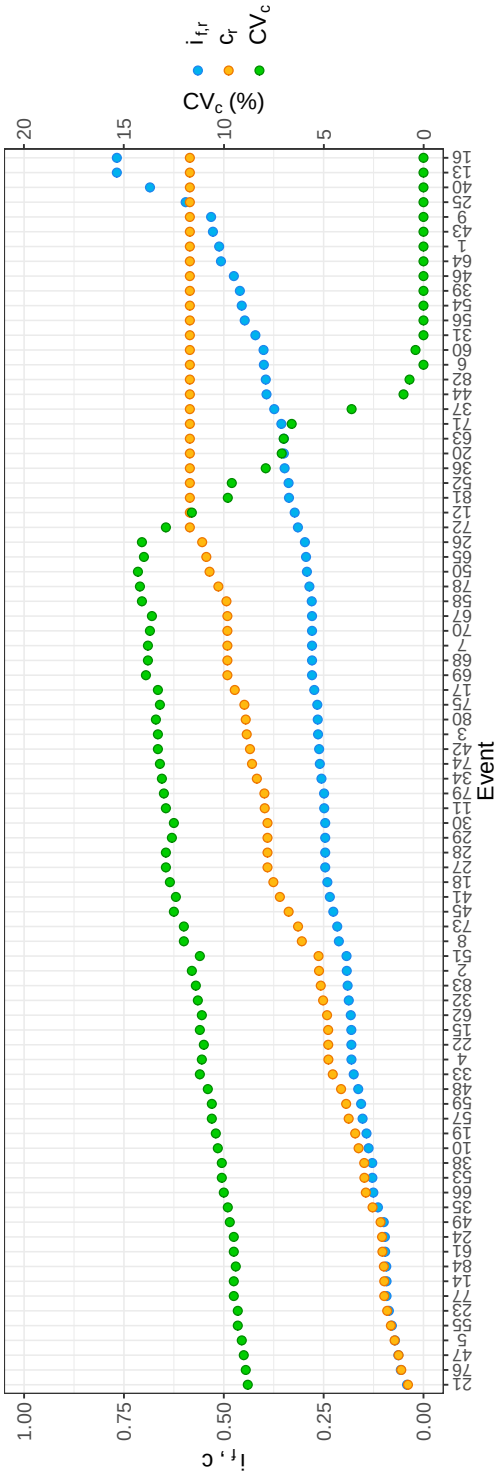

Figure 2.4: Reference slope $i_{f,r}$, reference concentration c_r and CV_c of each event. The events have been sorted with increasing $i_{f,r}$.

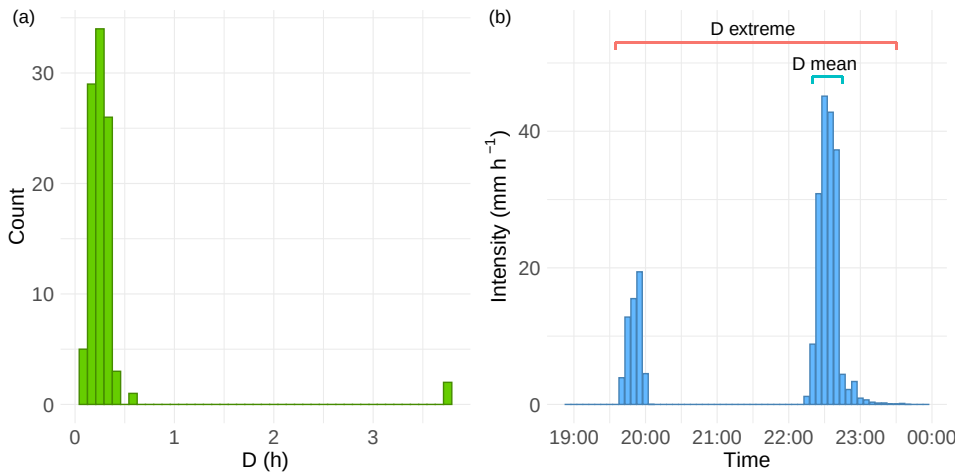

Figure 2.5: (a) Duration distribution histogram and (b) hyetograph of the event 58. The histogram shows the presence of an extreme isolated from the mass of the distribution. This extreme ($D = 3.75$ h) is shown in the hyetograph and compared with the mean ($D = 0.32$ h).

values of D very distant from the mean. Indeed, CV is very sensitive to the extremes (e.g. Arachchige et al., 2020, Chau et al., 2005), mainly if they are located in the right-hand tail of the distribution (Bendel et al., 1989). For instance, the effect of the extremes on the CV_D is evident in event 58: the D distribution of this event has an extreme much greater than the mean (Fig. 2.5(a)). This is due to the presence of zero-intensity temporal instants in the middle of the hyetograph that must be considered (for two out of a hundred samples) to reach the highest values of E (Fig. 2.5(b)). This results in a high value of the standard deviation with respect to the mean, namely a high value of the CV_D. It is worth noting that the effects of this condition on I are smaller thanks to the mean carried out to obtain this event characteristic.

Regarding the absolute variability, the (I, D) couples variability ranges, in the ID plane, allow us to get an idea of how variable an event as a whole is and to presume how this variability may affect the threshold estimate. Consistently with the relative variability, the events with $CV_D = CV_I = 0\%$ have also zero-length absolute variability ranges. The non-zero ranges are shown in Fig. 2.6. As evident, the length of the ranges varies greatly depending on the event. In term of intensity, the maximum length is 66.21 mm h^{-1} and it is reached with the event 67 while, for the duration, the event 38 is characterised by the maximum length that is equal to 5.25 h. Besides, the length of the range for D is less than 1 h in all but 8 events

while for I it is less than 20 mm h^{-1} in all but 7 events. Moreover, in most cases, the mean is located neither vertically nor horizontally in the middle of the variability ranges, namely the D and I distributions are asymmetrical. To quantify their asymmetry, the related skewness SK_D and SK_I are computed for each event and shown in Fig. 2.7. The events with zero variability are characterised by $SK_D = 0$ and $SK_I = 0$. Moreover, in most cases, SK_D is positive while SK_I is negative: the longest tail of the distributions of D and I tends to be located on the right and the left of the mean respectively. This suggests that, given an event, the majority of the D values are characterized by duration shorter than the mean and the greatest contribution to the absolute variability is given by the longest durations (i.e. by the D distribution right extremes) as in event 38. Consistently, comparing Fig. 2.7 and Table 2.2, the events with the highest positive SK_D are the events with the highest CV_D (e.g. events 58 and 67). Instead, given an event, the concentration of the intensity values is greater towards the highest values and the smallest intensities (i.e. the I distribution left extremes) mostly contribute to the absolute variability (e.g. event 20 and 73). However, as said before, the I extremes have a slight impact on the absolute variability of I thanks to the mean procedure necessary for its computation that reduces the interval ranges.

2.5.2 Correlation between the D and I absolute variability and some event features

Despite the specificity of each considered event, it's possible to identify some event features that are correlated with the D and I absolute variability. It is worth noting that, in general, correlation does not imply causation (Wiedermann and Von Eye, 2016) but it is a starting point to understand if causality between the variables can be established.

We define E_{av} as the rainfall volume per unit area available in the "main part of the hyetograph", namely the integral of the rainfall intensity on the smallest time interval comprising the peak and included between two instants with null intensities. We can then introduce the ratio E_r/E_{av}. As shown in Fig. 2.8, the absolute variability of D and E_r/E_{av} are positive correlated. A small value of E_r/E_{av} means that the main part of the hyetograph is amply able to provide E_r (i.e. to satisfy the condition of Eq. (2.8) in the reference conditions). This tends to avoid having to consider null intervals to achieve the values of E resulting from the MC simulation, namely to avoid D extremes. The opposite situation occurs if the ratio takes high values.

Regarding the intensity, we define I_{max} as the hyetograph maximum

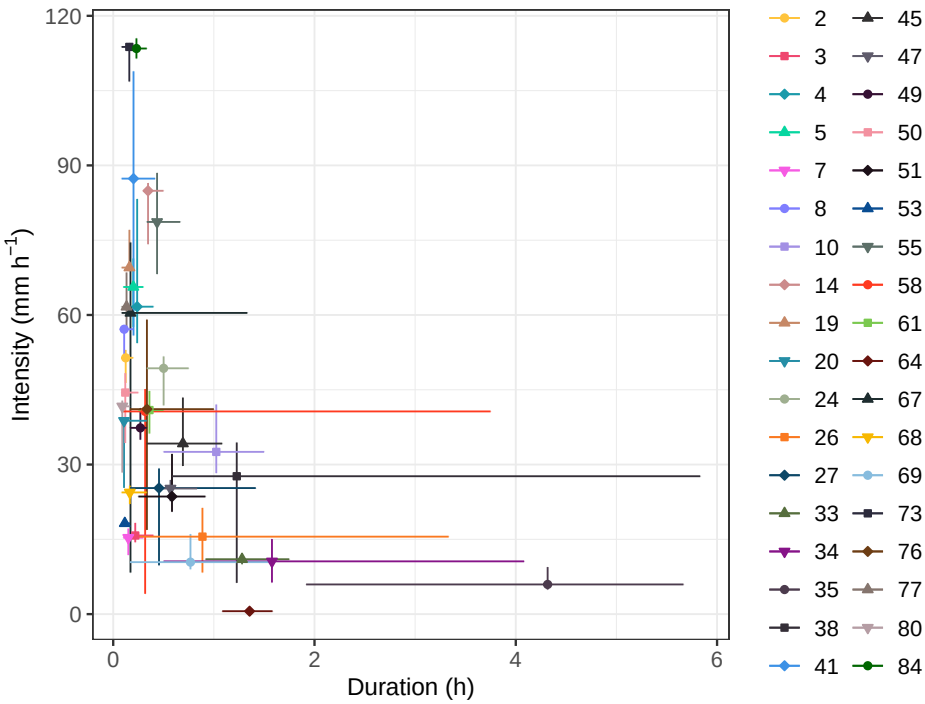

Figure 2.6: Absolute variability in the (I, D) couples. The symbols are the mean values and the horizontal and vertical lines are respectively the duration and intensity variability ranges. To make the graph clearer, the events with ranges equal to zero have not been represented and the linear scale is used for both axis.

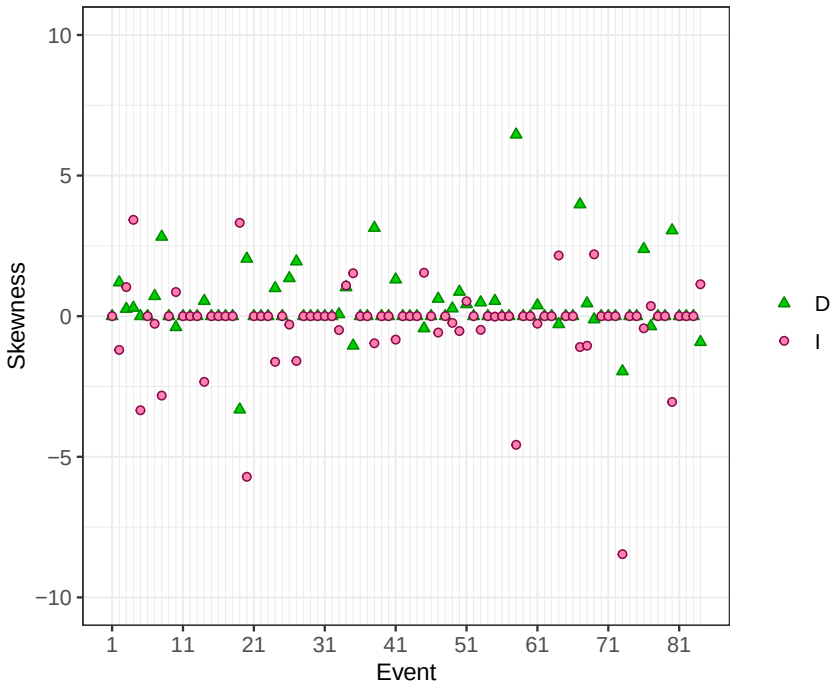

Figure 2.7: Skewness of the distributions of D and I for each event.

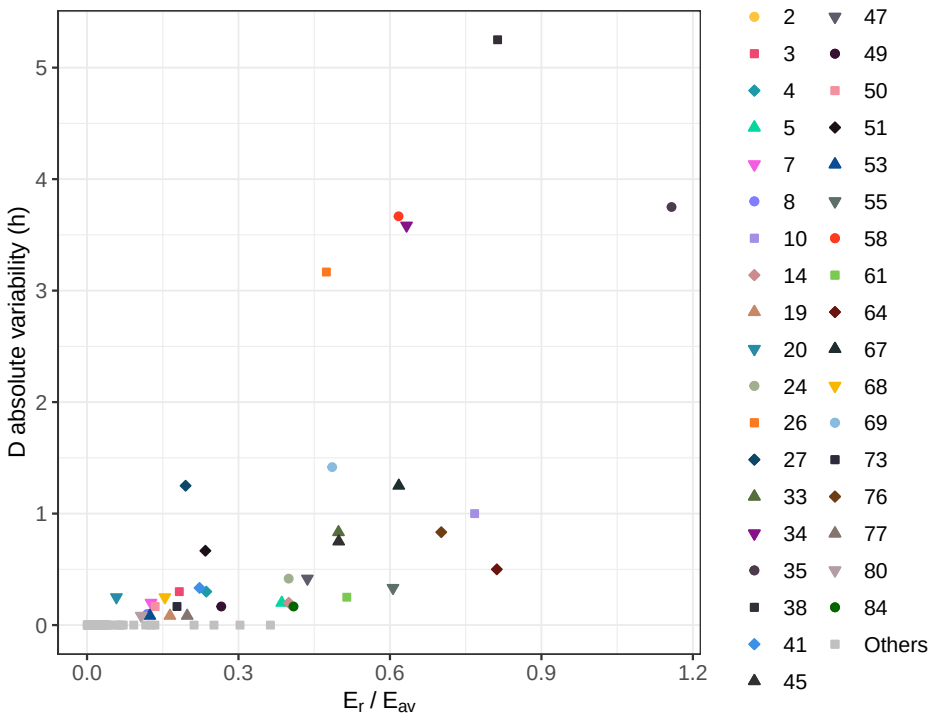

Figure 2.8: Positive correlation between the absolute variability of D and E_r/E_{av}, where E_{av} is the rainfall volume per unit area available in the main part of the hyetograph. Spearman correlation coefficient equals to 0.82 ($p < 2.2 \times 10^{-16}$). To make the graph clearer, the events with absolute variability of D equals to zero are represented with the same symbol.

intensity and I_{mean} as the mean intensity of the main part of the hyetograph for each event. The ratio I_{max}/I_{mean} provides a quantitative measure of the shape of the event hyetograph or, equivalently, of how impulsive the event is. As shown in Fig. 2.9, a positive correlation subsists between the non-zero absolute variability of I and I_{max}/I_{mean}. If the shape of the hyetograph around the peak is flat, and the ratio I_{max}/I_{mean} is low, the variability of I, connected to the variability of D, is small since the average procedure, necessary to compute I, involves similar intensities intervals. The opposite occurs when the event is impulsive and the ratio is high. This consideration is valid only for events with non-zero absolute variability in I and D and tends to explain why some events with high variability in D have small variability in I (e.g. event 26).

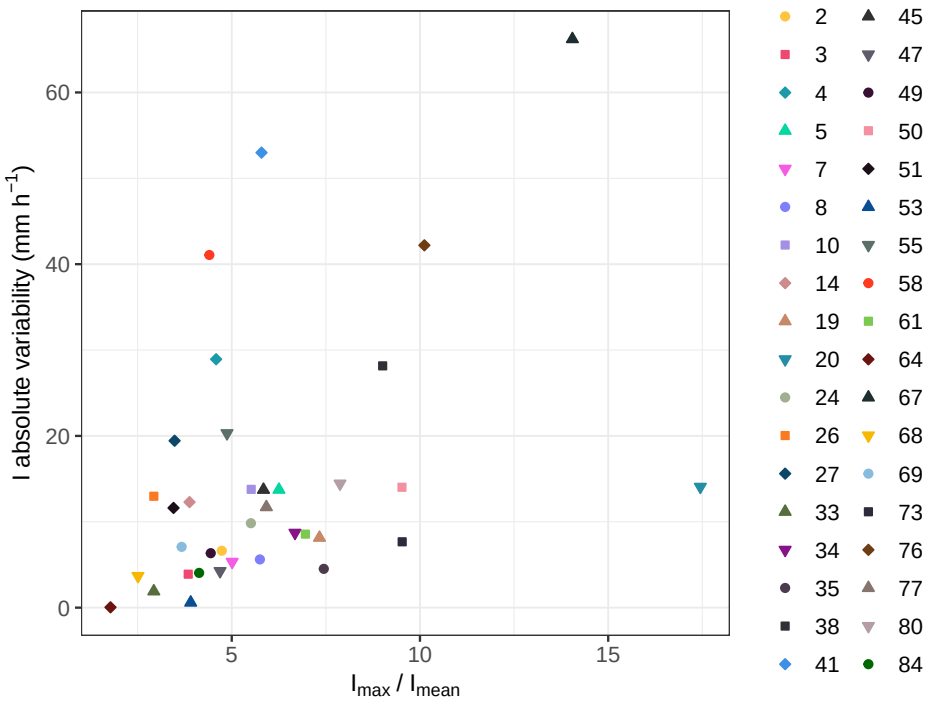

Figure 2.9: Positive correlation between the non-null absolute variability of I and I_{max}/I_{mean}. I_{max} is the maximum intensity and I_{mean} is the mean intensity of the main part of the hyetograph. Spearman correlation coefficient equals to 0.48 ($p = 0.0035$).

Coefficient	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

Table 2.3: Mean, standard deviation, variation coefficient CV and mean 95% confidence interval CI of the coefficients a and b of Eq. (2.12), computed performing the second MC simulation.

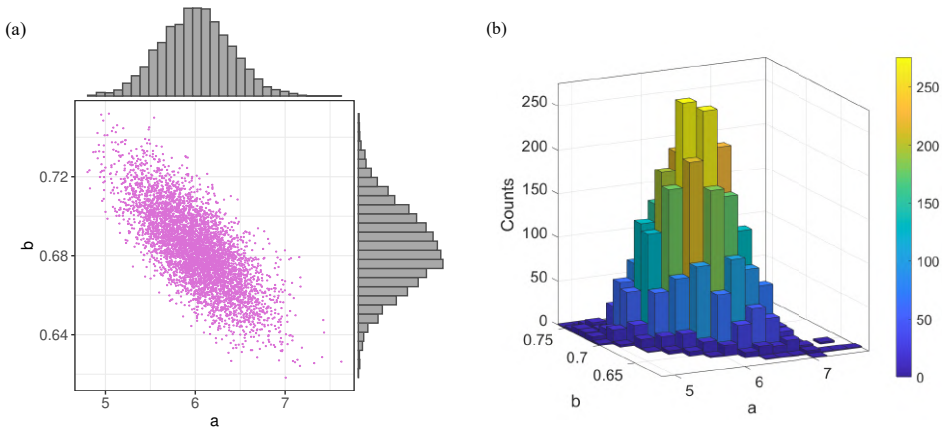

Figure 2.10: Values a and b of Eq. (2.12) obtained performing 5000 MC simulations: (a) scatter-plot and histogram and (b) $3D$ histogram

2.5.3 Variability of the threshold

The result of the second MC simulation is 5000 (a, b) couples. The main statistical quantities of their distributions are given in Table 2.3. The relative variability is quantified through the CV that is equal to 6.46% for a and 2.91% for b. The low spread nature of a and b, highlighted by the small CV values, is also evident in the scatter plot and in the $3D$ histogram, respectively shown in Fig. 2.10(a) and 2.10(b). In addition, to analyse the absolute variability of the I-D threshold relation, the intensity values for each (a, b) couple are calculated for D values spanning from five minutes to six hours with a five minutes time step. In this way, for each duration, we obtain an intensity distribution 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 absolute variability. The result is shown in Fig. 2.11. According to the substantially symmetrical distributions of a and b (Fig. 2.10(a)), the threshold computed with the mean values of a and b (Table 2.3) is essentially equidistant from the lower and upper bounds. The variability bandwidth decreases monotonically by increasing the duration

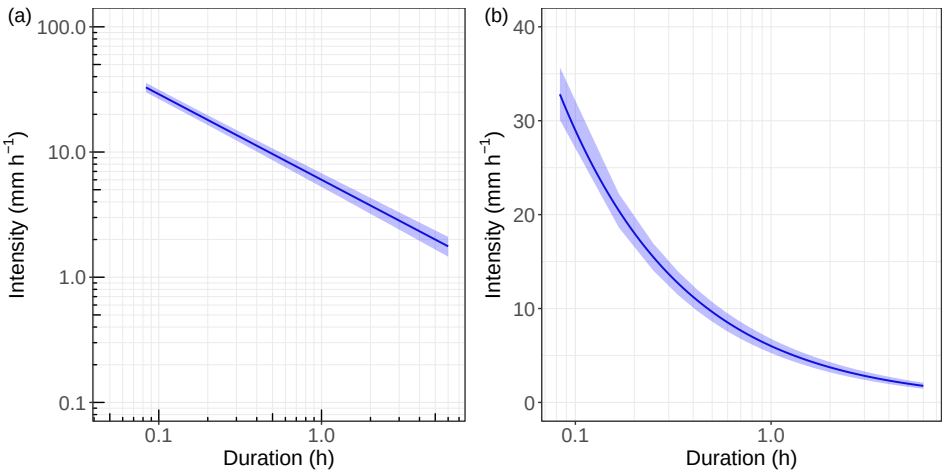

Figure 2.11: (a) Log-log and (b) semi-log plot of the threshold absolute variability. The blue line is the rainfall threshold obtained using the mean value of a and b (Table 2.3). The shaded area represents the threshold absolute variability whose upper and lower bounds have been computed considering the 2.5 and 97.5 percentiles of the intensity distributions for fixed durations.

and varies between 5.61 mm h^{-1} and 0.64 mm h^{-1}.

Hence, both the relative and the absolute variability highlight that the effect of the uncertainty in the variables on the threshold estimate is small. This is mainly due to the zero variability in the D and I distributions of 48 events out of 84: 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.

2.5.4 Reference values versus MC means

Finally, a comparison between the results of the first and second MC simulation and the reference values is carried out. In particular, we compare:

- the means of the D and I distributions (Fig. 2.6) to the corresponding reference ones, for each event;
- the mean threshold (i.e. threshold computed with the mean values of a and b) and the threshold absolute variability bounds (Fig. 2.11) to the reference threshold (i.e. Eq. (2.13)).

As regards D and I, according to Marra (2019), the bias of the duration

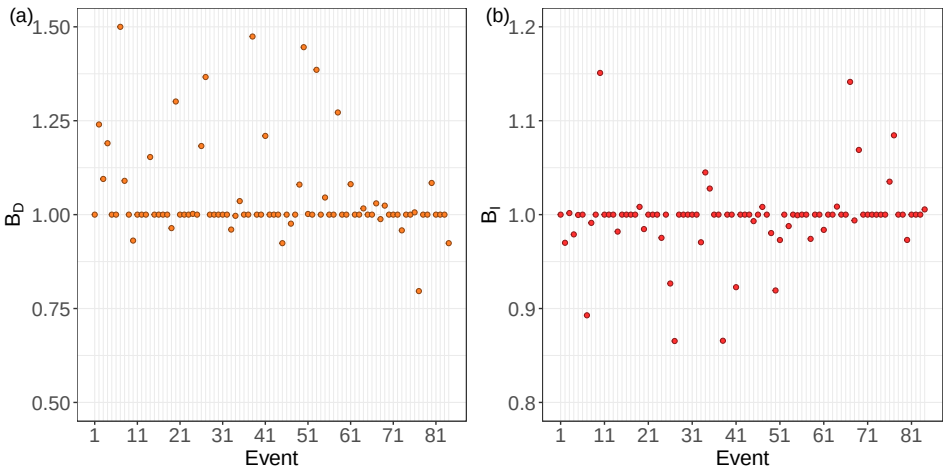

Figure 2.12: Bias of (a) duration B_D and (b) intensity B_I between the mean values of the D and I distributions obtained performing the first MC simulation and the corresponding reference values for each event.

B_D and the intensity B_I are computed for each event as:

$$B_D = \frac{D_m}{D_r} \qquad B_I = \frac{I_m}{I_r} \tag{2.14}$$

where the subscripts m represent the mean of the MC D and I distributions. The result is shown in Fig. 2.12: B_D deviates between 0.8 and 1.5 (Fig. 2.12(a)) while B_I between 0.86 and 1.15 (Fig. 2.12(b)). Consistently with the variability analysis described in Sect. 2.5.1, most events (48) are characterised by $B_D = B_I = 1$. This means that for these zero-variability events, the reference duration and intensity are exactly the MC mean values of I and D, namely the only MC (I, D) couple. Moreover, most of the remaining events have $B_D > 1$ and $B_I < 1$. This signifies that the MC (I, D) mean couples tend to be located lower and more to the right than the reference ones in the log-log ID plane.

Regarding the threshold, the differences between the MC intensities $I_{MC,k}$, where k stands for mean, upper bound and lower bound, and the reference threshold ones $I_{t,r}$ are carried out for the same durations used to define the absolute variability of the threshold:

$$\begin{aligned} &\mathrm{Diff}\,(k, D) = I_{MC,k}\,(D) - I_{t,r}\,(D)\,, \\ &k = \text{mean, upper bound, lower bound} \end{aligned}$$

The result is shown in Fig. 2.13(a). For almost all durations, the intensities of the mean threshold are slightly lower than the reference threshold ones:

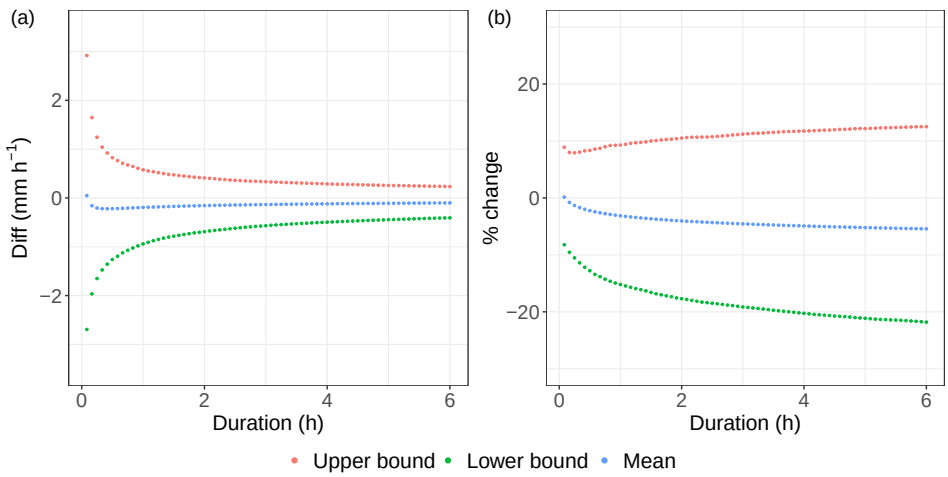

Figure 2.13: (a) Difference between the MC intensities (upper bound, lower bound and mean) and the reference threshold ones as a function of the duration; (b) percentage change of the MC intensities respect the reference threshold ones, as a function of the duration.

a positive difference occurs only for the first time interval. Consistently with the obtained a and b mean values (Table 2.3) 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 and lower bounds are respectively always higher and lower than the reference threshold.

Subsequently, the percentage changes, defined as:

$$\%\ \text{change}\,(k, D) = \left(\frac{\text{Diff}\,(k, D)}{I_{t,r}\,(D)}\right) \cdot 100,$$
$$k = \text{mean, upper bound, lower bound}$$

are computed to figure out how much the second MC outcomes deviate relatively from the reference threshold. The percentage changes are plotted in Figure 2.13(b): 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.

It can therefore be generally stated that the outcomes of the uncertainty analyses, both (I, D) couples and threshold estimate, are consistent with the reference ones. Coherently with the previous analysis, also in this comparison, the duration is the quantity with the highest bias values. However, the mean threshold and the reference one are very close, pointing out the small effects of the differences between D_m and D_r on the threshold computation.

2.5.5 Further elements of uncertainty

In the calibration of the BDA-based threshold and in the assumptions of the developed method used to assess the uncertainty, it is possible to identify some elements that may introduce further uncertainty, beyond that considered in this analysis, in the calculation of the event characteristics and, consequently, in the estimate of the threshold. Firstly, the variability ranges and the probability distributions of the parameters and data, namely the uncertainty characterisation of the variables, are uncertain. Secondly, the equations of the BDA may be uncertain since they are based on some simplifications and hypothesis. Finally, the radar data may be affected by uncertainty due to other sources of error, beyond the beam shielding one (considered in this analysis), such as signal attenuation in heavy rain or wet radome attenuation (Marra et al., 2014). Nevertheless, at the present state of the research, it is not possible to assess the impact of these uncertainties on the event characteristics estimate and further study is required.

2.6 Conclusions

This study has aimed to assess the effects of the uncertainty in the physical and morphological parameters and data on the BDA-based threshold calibration to evaluate the method robustness. To that end, a suitable methodology composed of two MC cascade simulations has been developed and applied to a specific study area and dataset. The first MC simulation has allowed examining the uncertainty propagation in the event characteristics estimate. The results have highlighted that most of the events (i.e 48 events out of 84) are characterised by zero variability in the estimation of the (I, D) couples while the duration and the intensity related to the remaining events are affected by variability, that can be low or high depending on the event. Overall, the duration has found to be the most variable outcome in relative term while I, thanks to the average procedure, has a lower relative variability. In absolute term, the variability of the (I, D) couples differs greatly between the events and the D and I distributions tend to be skewed to the right and left respectively. Moreover, considering the mean values of the events with non-zero variability (36 events out of 84), the uncertainty in the variables tends to provide slightly longer durations and slightly smaller intensities with respect to the reference ones. Notwithstanding, the second MC simulation has shown that the threshold computation is affected by small variability. The low dispersion of the threshold coefficients is mainly due to the 48 events with zero variability. As a result, the BDA method, applied to the considered dataset, can be described as robust since it provides

a calibrated threshold low sensitive to the considered uncertainty in the parameters and data. This is also highlighted from the consistency between the uncertainty analysis mean threshold and the reference one.

Overall, the results of this analysis can be useful to calibrate a BDA-based threshold for a different study area since the investigation has highlighted the main elements that could undermine the BDA robustness. In particular, given a debris flow and the related rainfall event, it was noted that some event features are correlated with the variability of D and I. The percentage of the needed rainfall volume and the available one in the main part of the hyetograph is positive correlated with the absolute variability of D. Moreover, the shape of the main part of the hyetograph, described by the ratio between the maximum and the mean intensity, is positive correlated with the non-null absolute variability of I. Therefore, given an event, these trends can be used to presume the possible variability in the estimate of D and I, without carrying out a specific uncertainty analysis. In other words, if an event is characterised by (i) low availability of rainfall volume in the main part of the hyetograph with respect to the needed one and (ii) a peak intensity much greater than the mean one, variations in the parameters and data is likely to result in high variability in D and I estimate. The presence of many events of this type could undermine the BDA robustness. Therefore, in these cases, it is advisable to put care in the estimate of the parameters and data. Besides, given an event, further elements likely affecting the estimate of event characteristics have been highlighted in this study: (i) the variability ranges and the probability distributions of the parameters and data, (ii) the equations constituting the BDA model and (iii) radar data. These elements can be affected by uncertainty and impact the event characteristics estimate. The uncertainty analysis performed in this study does not provide quantitative information on these impacts. Further analysis will assess how these three elements affect the (I, D) couple estimate and, consequently, the threshold calibration.

Moreover, the developed method, composed of two cascade MC simulations, can be applied to assess the uncertainty related to other threshold calibration approaches whose event characteristics estimate is based not only on the hyetograph but also on other variables (e.g. the one proposed by Zhang et al. (2020)). Indeed, the developed method allows considering the entire range of uncertainty of the variables and, therefore, avoiding the analysis by scenarios, quite widespread in the literature for the uncertainty analysis of rainfall thresholds (e.g. Nikolopoulos et al., 2014, Peres et al., 2018). Analysing by scenarios may not be suitable if the uncertain parameters have a continuous range of variability. Indeed, a low number of input

values combinations may not provide an overall assessment of the variability of the outputs.

Finally, it is worth noting that the results of this analysis are not useful to check the forecast capability of the threshold. Indeed, the variability in the threshold estimate due to the uncertainty of the inputs is not related to its forecast effectiveness but only to its robustness. The threshold forecast capability can be proved only by performing a proper validation analysis, essential to make this tool operational. Since the calibration method applied to the specific study area is proved to be robust, further analysis will assess the forecast capability of the threshold, developing an appropriate validation method.

Chapter 3

Validation and potential forecast use of a debris-flow rainfall threshold calibrated with the Backward Dynamical Approach

3.1 Introduction

Debris flows are usually rainfall-induced phenomena that may occur in mountain regions and have serious repercussions on the society and economy of affected territories (e.g. Dowling and Santi, 2014, Jakob et al., 2012c, Jalayer et al., 2018). For this reason, the reduction of the debris-flows risk to an acceptable residual level is crucial and it can be pursued through both structural and non-structural mitigation strategies. In this context, an early warning system (EWS) is a non-structural strategy that aims to reduce vulnerability by allowing the implementation of civil protection activities (e.g. evacuation) before debris flow occurrence. Since rainfall is one of the main triggering factors for debris flows occurrence, EWS for these phenomena are usually based on rainfall thresholds (e.g. Pan et al., 2018, Ponziani et al., 2020). Indeed, rainfall thresholds are mostly power-law relations that define the critical rainfall conditions above which it is likely to observe debris flow (e.g. Caine, 1980, Cannon et al., 2011, Giannecchini et al., 2016, Huang et al., 2019, Jakob et al., 2012b, Nikolopoulos et al., 2014, Staley et al., 2013). Generally, the rainfall physical quantities employed to

describe a rainfall event and used in the computation of the thresholds are its duration and its average intensity or cumulative precipitation (Segoni et al., 2018) and the threshold is estimated (i.e. calibrated) by applying empirical or physical approaches (Nikolopoulos et al., 2014).

The reliability of a rainfall threshold is crucial for its operational use and, to assess it, a proper validation procedure is required (Gariano et al., 2015). The validation aims to verify the capability of a calibrated threshold to classify some physical quantities related to rainfall events that resulted (or not) in debris flow and can be performed considering different approaches. The validation approaches mainly differ in terms of rainfall events considered, namely only rainfall events that resulted in a hazardous phenomenon (e.g. Gioia et al., 2015) or all rainfall events recorded during a reference period (e.g. Brunetti et al., 2018, Gariano et al., 2015), and strategies and parameters used to quantify the threshold performance (e.g. receiver operating characteristic curve, skill scores).

However, as pointed out by Segoni et al. (2018), the threshold reliability is rarely evaluated after the calibration and this strongly limits its operational use (Leonarduzzi et al., 2017). Furthermore, in many works in which the validation is performed, the reliability is estimated using the same dataset employed for the threshold calibration and/or a very small number of events (Segoni et al., 2018). This may lead to a no trustworthy performance estimate.

In Rosatti et al. (2019) a stony debris flow rainfall threshold has been calibrated, for a study area, with the Backward Dynamical Approach (BDA). The BDA is a physical-based method to estimate stony debris flow rainfall thresholds in which the rainfall event duration D and average intensity I related to an occurred debris flow are computed by combining the forcing and the dynamic of the phenomenon. In particular, the rainfall volume and the corresponding rainfall conditions strictly related to a debris flow are estimated based on the volume of the sediments, surveyed after the event occurrence, through a simplified description of the phenomenon dynamic. The robustness of the BDA method in calibrating a rainfall threshold has been analyzed in Martinengo et al. (2021b).

This work aims to assess the reliability of the rainfall threshold calibrated in Rosatti et al. (2019). To this aim, we carry out a validation process composed of two different approaches. The first approach is coherent to the calibration procedure and considers only rainfall events that resulted in debris flows. This approach aims to check the reliability of the threshold in the identification of the $I - D$ plane area where the probability of debris-flow occurrence is significant. The second validation approach is introduced to

perform a more comprehensive and more robust threshold reliability assessment. In this approach, all rainfall events that occurred during a reference period are taken into account (Brunetti et al., 2018, Gariano et al., 2015) and a proper method to compute the related rainfall conditions is developed. This approach aims to assess the reliability of the threshold in terms of its ability to correctly classify the rainfall events that resulted or not in a debris flow. Both approaches are applied considering a proper dataset related to a study area, contained in the one used for calibration, and a reference period not used previously.

The paper is organised as follow. In Section 3.2 we briefly describe the BDA method to provide an overview of the assumptions on which the threshold calibration is based in Rosatti et al. (2019). The considered study area and the available rainfall and debris-flow data are described in Section 3.3 and the two validation approaches adopted are presented in Section 3.4. The obtained results are presented in Section 3.5 while in Section 3.6 we discuss them focusing on the resulting threshold reliability, the consistency between the two validation approaches and the potential forecasting use of the threshold. Finally, the key findings and the possible further analysis are summarized in Section 3.7.

3.2 The Backward Dynamical Approach

The Backward Dynamical Approach (BDA) is the method for calibrating stony debris flow rainfall threshold introduced by Rosatti et al. (2019). It allows you to estimate the rainfall duration D and average intensity I strictly related to a surveyed debris flow deposit, considering both the hyetograph and the dynamic of the occurred event.

According to this approach, the rainfall volume V_r^{DF} related to a debris flow is defined as the water quantity required to carry downstream, as a mixture, the surveyed deposited volume V_{dep}. Following this definition and describing the debris-flow dynamic with a simplified global volumetric approach (see Rosatti et al., 2019, for details), V_r^{DF} can be written as:

$$V_r^{DF} = \frac{c_b - c}{c} V_{dep} \tag{3.1}$$

in which c_b and c are respectively the maximum (volumetric) bed concentration (Takahashi, 2014) and the solid reference concentration of the considered debris flow. In the BDA method, c is computed according to the expression proposed in Takahashi (1978):

$$c = \min\left(\frac{i_f}{\Delta\left(\tan\psi - i_f\right)},\ 0.9\,c_b\right) \tag{3.2}$$

in which i_f is the bed slope in a significant reach (see sec. 3.4.1 for a detailed definition), Δ is the relative submerged density, equal to $(\rho_s - \rho_w)/\rho_w$ where ρ_s and ρ_w are, respectively, the solid and the water density and, finally, ψ is the sediments dynamic friction angle.

A quantity derived from (3.1) is E, defined as the rainfall volume per unit area of the basin where the debris flow occurred:

$$E = \frac{1}{A_b}\frac{c_b - c}{c}V_{dep} \tag{3.3}$$

This quantity can also be identified within the hyetograph of the associated phenomenon by integrating the measured rainfall intensity $i(t)$ over a suitable time interval:

$$E = \int_{t_1}^{t_2} i(t)\,dt \tag{3.4}$$

where the extremes of integration, t_1 and t_2, are assumed to be the start and end times of the debris-flow duration. In the absence of detailed information about these times, the BDA method assumes that they can be expressed as a function of the event peak intensity instant t_{max} in the following way:

$$\begin{cases} t_1 = t_{max} - \Delta t_1 \\ t_2 = t_{max} + \Delta t_2 \end{cases} \tag{3.5}$$

where Δt_1 and Δt_2 are unknown intervals that can be determined by equating the right-hand sides of equations (3.3) and (3.4). Nevertheless, since the rainfall intensity $i(t)$ is a piece-wise constant function of the measurement time interval δt, the equality cannot be satisfied exactly. Therefore, setting $\Delta t_1 = n_1\delta t$ and $\Delta t_2 = n_2\delta t$, these intervals are calculated increasing n_1 and n_2 until:

$$\int_{t_{max}-n_1\delta t}^{t_{max}+n_2\delta t} i(t)\,dt \geq \frac{1}{A_b}\frac{c_b - c}{c}V_{dep} \tag{3.6}$$

The reader is referred to Rosatti et al. (2019) and Martinengo et al. (2021b) for more details on the estimate of Δt_1 and Δt_2. Once Δt_1 and Δt_2 are computed, the duration D and the average intensity I are defined as:

$$D = \Delta t_1 + \Delta t_2 \tag{3.7}$$

$$I = \frac{E}{D} \tag{3.8}$$

Finally, the $I - D$ threshold can be estimated considering all the (I, D) couples related to the available debris-flow data and applying the frequentist method (e.g. Brunetti et al., 2010, Peruccacci et al., 2012). For more details on threshold calibration, we refer again the reader to Rosatti et al. (2019) and Martinengo et al. (2021b).

3.3 Study area and data

In the calibration work (Rosatti et al., 2019), and the related uncertainty analysis (Martinengo et al., 2021b), the study area considered is the Trentino-Alto Adige/Südtirol region (Italy) (Figure 3.1(a)) while the debris flows occurred between 2006 and 2016 in this area are considered as the referring dataset. Based on these data and assuming a 5% non-exceedance probability, the following BDA-based threshold is calibrated for the analyzed region:

$$I = 6.2D^{-0.67} \tag{3.9}$$

In this work, the study area used for the threshold validation is a part of the previous study area, namely the Autonomous Province of Bolzano, (Figure 3.1(a)) while the reference period runs from May to September 2017, a period that is not included in the calibration one. During this period, the regional agency has reported 108 debris flows occurred in the study area (Figure 3.1(b)). As highlighted from Figure 3.2(a), most of them occurred in August while no events were recorded in May and September. In particular, more than a third of the debris flows (i.e. 38) occurred on the same day, namely on the 5th of August, due to a very intense storm that mainly affected the Alta Val Pusteria, the easternmost part of the study area (notice the cluster of events in this zone).

The deposited volume of the surveyed events ranges between 30 m^3 and 100000 m^3 with a prevalence (third quartile) of volumes smaller than 8000 m^3 (Figure 3.2(b)).

As for the precipitation, the available rainfall data derives from a C-band Doppler weather radar located in the study area, over the Mt Macaion at 1866 m a.s.l. (Figure 3.1(b)). This radar provides the reflectivity Z for an area of 240 km of diameter, over a grid with 500 m pixel size, with a time step of 5 minutes. The precipitation is then obtained by converting the

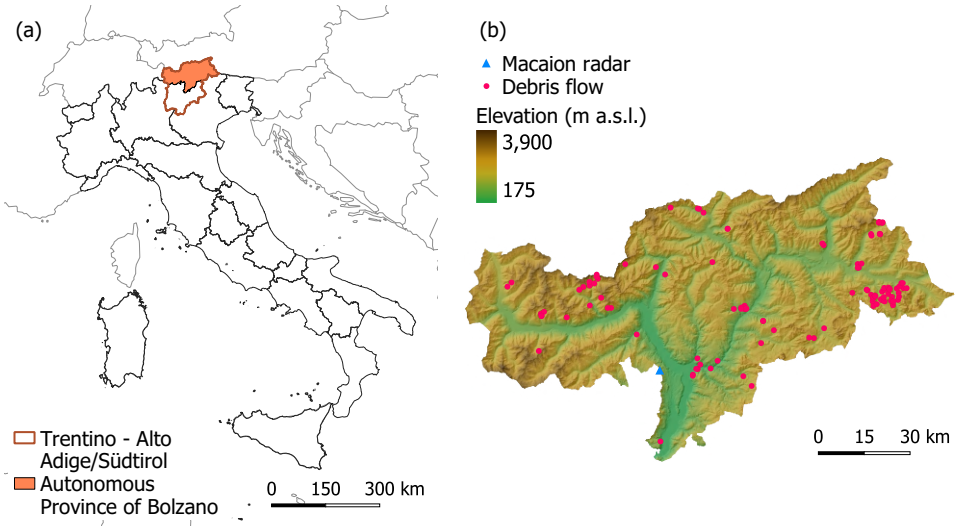

Figure 3.1: (a) Location of Trentino-Alto Adige/Südtirol region (Italy) and Autonomous Province of Bolzano; (b) position, within the study area, of the Macaion radar and the reported debris-flow events for the period May-September 2017.

reflectivity in rainfall intensity with the Marshall-Palmer relationship (e.g. Marshall and Palmer, 1948, Uijlenhoet, 2001).

3.3.1 Data uncertainties

The sources of uncertainty that may affect the data used in this work are the same as those highlighted in Martinengo et al. (2021b). The most significant are the ones connected to the radar rainfall data, the BDA parameters, and the regional agency's debris-flow events report.

Regarding the radar data, these can be affected by different sources of error such as beam shielding, attenuation of the signal in heavy precipitation and attenuation due to wet radome (e.g. Marra et al., 2014). These errors can result in unreliable estimates of the rainfall variables used in the validation. Nevertheless, it is worth noting that, in general, the main aim of a rainfall threshold is its use in a real-time operational EWS. In this context, it is reasonable that the raw radar rainfall data are considered due to the lack of time and data for the errors correction (Marra et al., 2014).

Moreover, uncertainty in data and parameters affects the BDA approach. As highlighted in Martinengo et al. (2021b), this uncertainty, in most cases, results in negligible effects on the estimate of the (I, D) couples computed with the BDA. Therefore, in this work, the impact of data uncertainty on the validation processes is neglected.

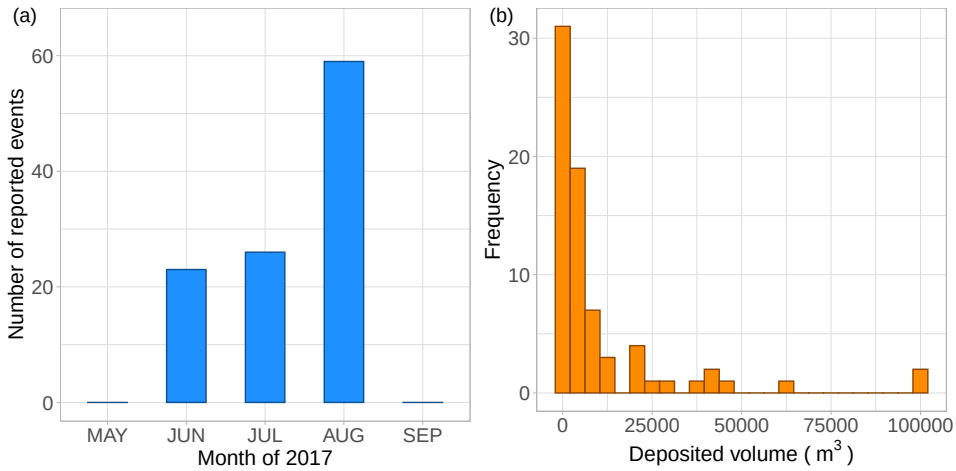

Figure 3.2: (a) Number of debris flows reported for each month of the period May-September 2017; (b) histogram of deposited volume values surveyed after the events.

Finally, the debris-flow event reports may be characterised by a lack of information (e.g. a debris flow occurred but did not report) or errors. As stressed in Gariano et al. (2015), the epistemic uncertainty due to the degree of completeness and correctness of the report of a debris-flow event can significantly affect the assessment of the threshold's reliability. However, this uncertainty is hardly quantifiable and, for this reason, it is neglected in this analysis.

3.4 Method

To test the reliability of the calibrated threshold (Eq. (3.9)), we perform a validation process based on two different approaches. The first one consists in applying the same method used in the calibration, namely the BDA, to compute the (I, D) couples related to rainfall events that resulted in debris flows in a reference period not included in the calibration one. In this work, we name this approach the *BDA-based validation approach*. Then, to obtain a complete overview and a more robust result on the threshold reliability, we use a second approach that considers all rainfall events that occurred within a reference period (e.g. Brunetti et al., 2018, Gariano et al., 2015, Peres and Cancelliere, 2014, Segoni et al., 2014). In this work, we name this second approach the *potential debris flows validation approach*

3.4.1 The BDA-based validation approach

This approach consists in the following steps:

- consider a set of debris flows, not included in the determination of the threshold defined by Eq. (3.9), and evaluate the relevant (I, D) couples by using the BDA method;
- place the obtained couples in the $I - D$ plane and classify them according to whether or not they exceed the threshold;
- choose an appropriate skill score based on the previous classification and quantify the reliability of the threshold.

For the estimate of the $(I - D)$ couple, we followed the BDA procedure (Rosatti et al., 2019): the basin outlet is placed just upstream of the deposition zone and downstream of a creek segment characterised by a sufficiently constant slope; in this way, A_b is computed and i_f is calculated as the average slope of the last 50 m of stream segment located upstream of the outlet. Then, c is computed by using Eq. (3.2) and E by using Eq. (3.3) in which V_{dep} is related to the considered debris flow. For what concerns the rainfall event, the hyetograph $i(t)$ pertaining to the analyzed debris flow is defined as the average radar intensities, calculated instant by instant, over the basin area. Finally, E is obtained by solving Eq. (3.6) and the relevant (I, D) couple is calculated with Eqs. (3.7) and (3.8).

As for the classification, the threshold acts as a binary classifier of the rainfall events (see e.g. Brunetti et al., 2018, Gariano et al., 2015, Staley et al., 2013). Accordingly, each rainfall event is categorised as:

- True Positive (TP), if the related (I, D) couple is located over the threshold;
- False Negative (FN), if the related (I, D) couple is located under the threshold.

These combinations are summarized in Table 3.1.

Finally, the Probability Of Detection POD (e.g. Doswell III et al., 1990, Gariano et al., 2015, Kharin and Zwiers, 2003) is chosen as a suitable skill score to assess the threshold performance. The POD (also called Hit Rate) is defined, in this case, as the ratio between the number of rainfall events exceeding the threshold and the total number of observed debris flows:

$$POD = \frac{TP}{TP + FN} \tag{3.10}$$

POD varies between 0 and 1 and its optimal value is equal to 1.

	Debris flow occurred
Threshold exceeded	True Positive (TP)
Threshold not exceeded	False Negative (FN)

Table 3.1: Two possible combinations in which the rainfall events are classified in the BDA-based validation approach.

3.4.2 The potential debris flows validation approach

As evident from Sections 3.2 and 3.4.1, in the $I - D$ plane, a (I, D) couple refers to the rainfall intensity and duration strictly relevant to a debris flow event, i.e. only the portion of the rainfall event that is responsible for the occurrence. As a result, it appears that representing in this plane a rainfall event that did not result in a debris flow and, consequently, verifying the threshold's ability to classify (in some way) this type of event is impossible. To overcome this limit, we introduce the concept of "potential debris flows" associated with a rainfall event, which allows us to represent a given rainfall event in the $I - D$ plane, and a criterion to classify this representation as above or below the threshold.

First of all, given a catchment, the hyetograph of the considered period is divided into independent rainfall events using the critical duration method (CDM) (Restrepo-Posada and Eagleson, 1982) and considering the monthly values of critical durations calculated in Rosatti et al. (2019).

Then, given the hyetograph of an independent rainfall event, according to the definition of intensity and duration of a debris flow given in Section 3.2, a set of potential debris flows can be identified starting from t_{max} and fixing a set of durations. Since $i(t)$ is a piece-wise constant function of δt, the set is finite. Moreover, since the duration of most debris flows in the study area is less than an hour (see Rosatti et al., 2019) and $\delta t = 5$ min, we define the set of potential debris flows associated with a given rainfall event as the set of couples (I_k, D_k) with $k = 1, \ldots, 12$, where $D_k = k\delta t$ and I_k is the relevant intensity obtained according to a discrete expression of Eq. (3.4). Labelling the time interval exhibiting the maximum intensity with 0 and numbering the intervals so that positive values indicate instants after the peak while negative values are associated with instants before the peak (Figure 3.3) and indicating with i_k the intensity associated with the kth interval, we employ the following recursive formula to obtain the intensities of the set:

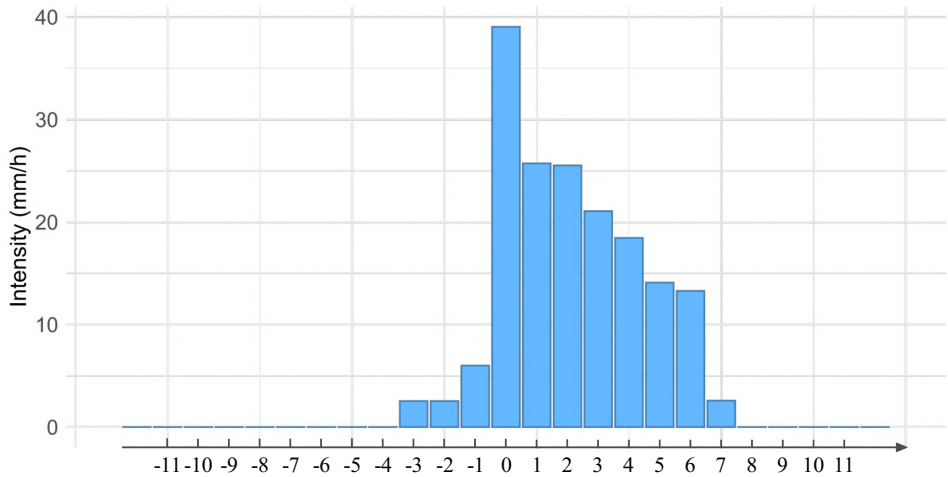

Figure 3.3: Example of an independent rainfall event with the discrete reference system and the related enumeration of the intensity intervals. For this example, the order of choice of intervals is: $[0, 1, 2, 3, 4, 5, 6, -1, 7, -2, -3, 8]$

$$\begin{cases} I_1 = i_0 \delta t \\ \mathbf{K} = [-1, 1] \end{cases}$$
$$\begin{cases} I_n = \dfrac{(n-1)\, I_{n-1} + \max\left[i_{\mathbf{K}(1)}, i_{\mathbf{K}(2)}\right] \delta t}{n} \\ loc = maxloc\left[i_{\mathbf{K}(1)}, i_{\mathbf{K}(2)}\right] \\ \mathbf{K}(loc) = \mathbf{K}(loc) + sign(\mathbf{K}(loc)) \end{cases} \qquad n \in [2, \ldots, 12] \tag{3.11}$$

where the function *maxloc* gives the position of the maximum value (namely, 1 or 2) in the relevant vector, while the function *sign* gives the value ± 1, depending on the sign of the argument. Moreover $\mathbf{K}(\cdot)$ indicates the elements in position $(\cdot)$ of the vector $\mathbf{K}$. On this method, any independent rainfall event can be represented in the $I - D$ plane by means of the twelve points defining the relevant set of potential debris flows (Figure 3.4).

The location of the relevant set with regard to the threshold can be used as a criterion to classify the event as a whole as over or under the threshold. We consider 12 classification methods based on the number of couples in the set located under the threshold, and we labelled each method with this number. For example, the classification method 1 considers a rainfall event as under the threshold if at least one of the relevant (I_k, D_k) couples is located under the threshold in the $I - D$ plane.

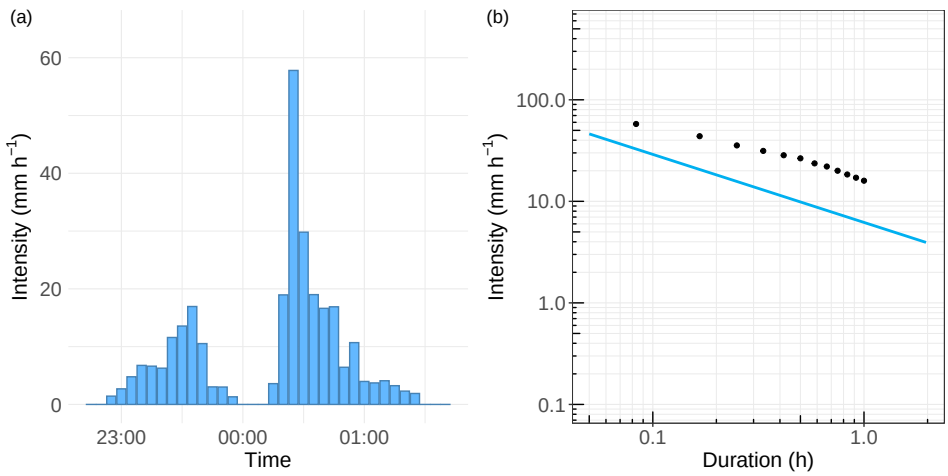

Figure 3.4: Example of an independent rainfall event described according to the potential debris flow validation approach. (a) The hyetogrpah of the considered event and (b) the 12 (I_k, D_k) couples that represent the set of potential debris flows related to the rainfall event.

According to a given classification method, a rainfall event that resulted in a debris flow is categorized as:

- True Positive (TP), if the event, as a whole, is over the threshold;
- False Positive (FP), if the event, as a whole, is under the threshold;

Instead, a rainfall event that is not related to an observed debris flow is categorized as:

- False Positive (FP), if the rainfall event, as a whole, is classified over the threshold;
- True Negative (TN), if the rainfall event, as a whole, is classified under the threshold;

These four possible combinations result in a classical contingency table (e.g. Fagerland et al., 2017, Wilks, 2011), reported in Table 3.2, for each classification method.

Once the contingencies are calculated, the reliability of the threshold is estimated with proper skill scores for each classification method. In addition to the POD (Eq. (3.10)), we introduce:

- Probability Of False Detection $POFD$ (also called False Alarm Rate) (e.g. Barnes et al., 2009, Jolliffe and Stephenson, 2012, Peres and

	Debris flow occurred	Debris flow not occurred
Threshold exceeded	True Positive (TP)	False Positive (FP)
Threshold not exceeded	False Negative (FN)	True Negative (TN)

Table 3.2: Contingency table highlighting the four possible combinations in which the rainfall events are classified according to the potential debris flows approach.

Cancelliere, 2021, Wilks, 2011) that quantifies the capability of the threshold in classifying the rainfall events that did not result in a debris flow:

$$POFD = \frac{FP}{FP + TN} \tag{3.12}$$

$POFD$ varies between 0 and 1 and its optimal value is equal to 0.

- True Skill Statistic TSS (also called Hanssen & Kuipers discriminant (Hanssen and Kuipers, 1965) or Peirce Skill Score (Peirce, 1884)) that combines correct and wrong classifications considering all the contingencies (e.g. Hirschberg et al., 2021, Leonarduzzi et al., 2017, Peres and Cancelliere, 2021):

$$TSS = \frac{TP}{TP + FN} - \frac{FP}{FP + TN} = POD - POFD \tag{3.13}$$

TSS varies between -1 and 1 and its optimal value is equal to 1.

3.5 Results

3.5.1 BDA-based validation approach

For the application to the study case, we consider as the set of events only those debris flows described in Section 3.3 that are suitable for analysis using the BDA approach, i.e., only those for which the following information is available: (i) the location of the deposition zone, (ii) V_{dep} that is assumed it must be greater than 100 m^3 for the analysis to make sense and (iii) the rainfall data when the event occurred. The resulting set is composed of 45 of the 108 reported events. Figure 3.5 shows the spatial distribution of the considered events along with the relevant deposited volume sizes.

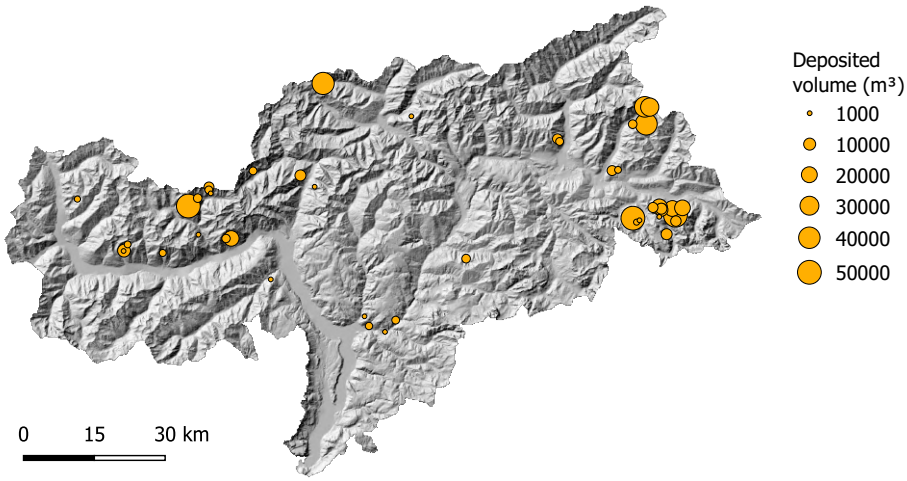

Figure 3.5: Location of the debris flows used for the BDA-based validation approach. The dot size is a function of the value of the surveyed deposited volume.

Table 3.3 summarizes the main morphological and dynamical-related quantities of the set in terms of the first and third quartile and median value. As evident, the analyzed debris flows occurred in small catchments having an area A_b mainly lower than 3.4 km^2. The bed slopes i_f, estimated just upstream of the deposition areas, range between 0.17 and 0.28 for 50% of the considered debris flows while the reference concentrations c are predominantly less than 0.39. Finally, the required rainfall volume V_r^{DF} is the quantity characterised by the widest range of variability, mainly due to the wide variability range of V_{dep} of the considered debris flows (Figure 3.5).

The (I, D) couples related to the analyzed debris flows are shown in Figure 3.6. Coherently with the calibration analysis, most of the rainfall events are characterised by $D < 1$ h (Rosatti et al., 2019). This further confirms the plausibility of the predefined durations considered in the potential debris flow validation approach. Overall, the maximum values of D and I are respectively equal to about 5 h and 118 mm/h while the minimum values are 5 minutes and about 7 mm/h. Regarding the positions of the (I, D) couples with respect to the calibrated threshold (Eq. (3.9)) in the $I - D$ plane, all but two rainfall conditions are located over the threshold. Therefore, in terms of rainfall events classification and skill score, this validation process results in $TP = 43$, $FN = 2$ and $POD = 0.96$.

	A_b (km^2)	i_f	c	V_r^{DF} (m^3)
First quartile	0.60	0.17	0.20	2008
Median	1.63	0.21	0.26	7633
Third quartile	3.40	0.28	0.39	18248

Table 3.3: First quartile, median and third quartile of the morphological and dynamical-related quantities involved in the BDA method implementation. A_b is the basin area, i_f is the bed slope, c is the reference concentration and V_r^{DF} is the rainfall volume strictly pertaining to the considered debris flow.

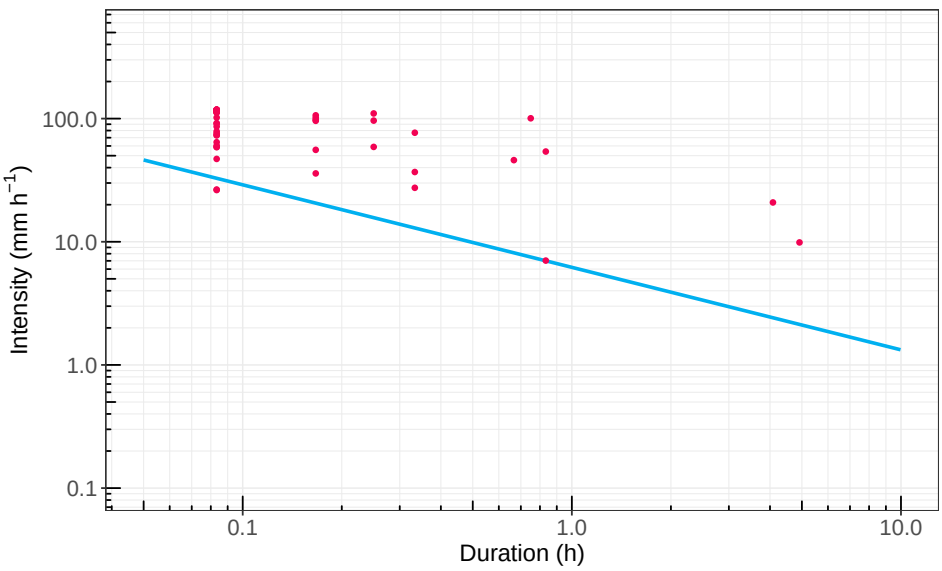

Figure 3.6: (I, D) couples resulting from the application of the BDA method to the 45 suitable debris flows event. The blue line is the threshold calibrated in Rosatti et al. (2019) end expressed by Eq. (3.9).

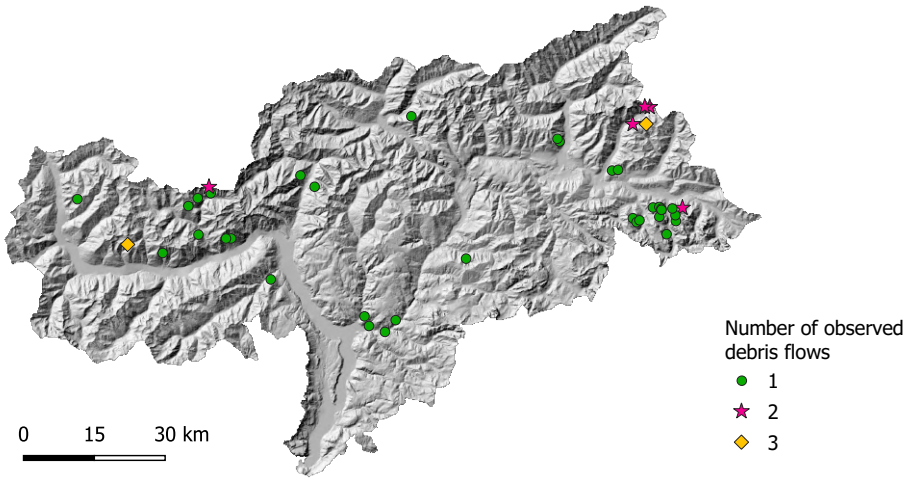

Figure 3.7: Location of the catchments considered for the validation process based on the potential debris flows validation approach along with the number of debris flows observed in each basin.

3.5.2 Potential debris flows validation approach

The starting data, in terms of catchments and debris flows, to be used in this approach are the same as those used in the BDA-based validation approach. However, some of these catchments (and the related debris flows) are excluded from this analysis since the relevant rainfall data are not available for the whole reference period. Instead, regarding the observed debris flows, all the events that affected the selected catchments are considered, thus going to include some events excluded in the BDA-based validation approach due to the lack of the deposit volume value. Indeed, knowing whether or not a debris flow occurred is sufficient for this analysis. Overall, the total number of catchments considered is 40 in which an overall of 49 debris flows was observed (Figure 3.7). The potential debris flows set is evaluated for each independent rainfall event, as explained in Section 3.4.2, and can be positioned with respect to the threshold in the $I - D$ plane in the following way:

- the whole set is over the threshold (Figure 3.8(a));
- the set intersect the threshold (Figure 3.8(b)). We name the events characterized by this kind of set as *intermediate* rainfall events;
- the whole set is below the threshold (Figure 3.8(c)).

Overall, the resulting number of sets for each position is respectively 355, 248 and 1948.

Regarding the intermediate rainfall events, Figure 3.9(a) shows the percentage of these events that have the couple (I_k, D_k) under the threshold, as a function of k: the (I_1, D_1) couple is below the threshold in almost 90% of the events while the (I_2, D_2) couple below the threshold in about 50% of cases. For the other couples, the percentage is around 25 - 30. Moreover, Figure 3.9(b) shows the relative number of intermediate rainfall events that present a set with k elements below the threshold as a function of k. As evident, 30% of the intermediate events have a set with only one rainfall condition under the threshold, 16% show a set with two elements; sets with more than two elements are less than 9% each.

Regarding the contingencies, Figure 3.10 shows the number of TP, FP, FN and TN for each analyzed catchment and classification method. Depending on the basin, the number of TP varies from 0 to 3 (Figure 3.10(a)) while the number of FN is equal to 0 or 1 (Figure 3.10(c)). In addition, the number of TP and FN are not affected by the classification method for all but one catchment. This means that only one observed debris flow is related to an intermediate rainfall event. Regarding the number of FP (Figure 3.10(b)) and TN (Figure 3.10(d)), these quantities are characterised by more variability than TP and FN both between catchments and method of classification. Depending on the basin, the minimum number of FP ranges between 2 and 14 while the maximum varies from 4 to 24. The ranges of variability for TN are between 35 and 63 and between 47 and 69 for the minimum and maximum number respectively, according to the catchment. In addition, given a basin, the number of FP increases moving from classification method 1 to classification method 12 while the opposite occurs for the number of TN. Regarding the results at the study area scale, the contingencies values (computed by adding the values obtained in each catchment) and the related skill scores are reported in Table 3.4, for each classification method. Consistently with Figure 3.10, the overall number of both TP and FN assumes only two different values and the change occurs when the minimum number of (I_k, D_k) couples required under the threshold switches from 7 to 8. Coherently, also POD takes just two distinct values, with maximum equals 0.96, constant for classification method from 8 to 12. Instead, according to the results at the basin scale, the total number of both FP and TN takes different values based on classification methods. This results in a variable value of $POFD$ which ranges between 0.12 and 0.22. The greatest relative variation in the number of FP and TN and in the value of $POFD$ occurs when switching from classification method 1 to 2 due to

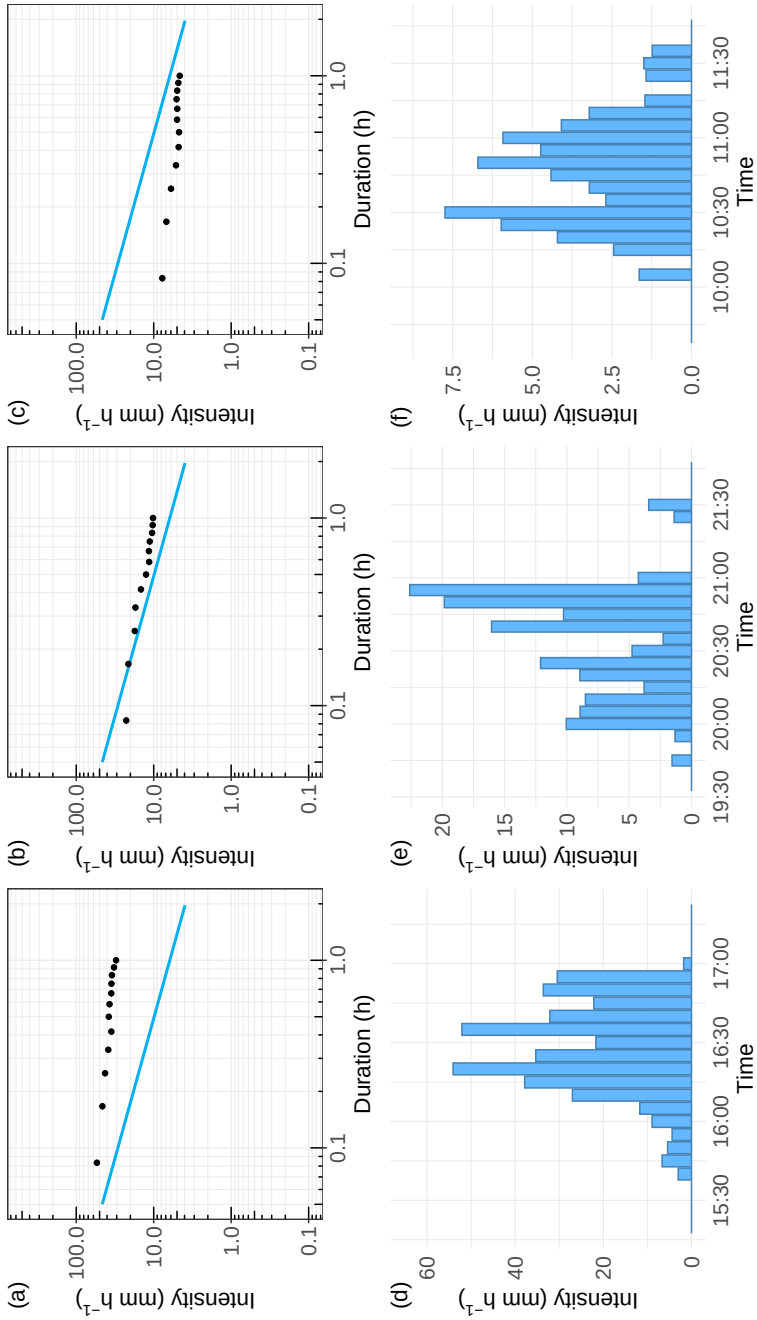

Figure 3.8: (a), (b) and (c): examples of possible positions of the rainfall event sets of potential debris flows with respect to the calibrated threshold (Eq. (3.9)) in the log-log $I - D$ plane; (d), (e) and (f) hyetographs generating the corresponding set.

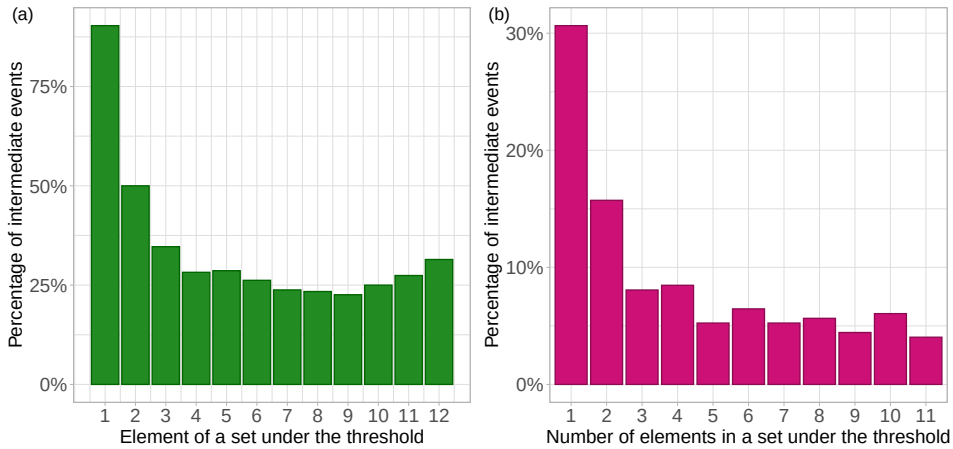

Figure 3.9: Bar graphs showing the percentage of intermediate rainfall events that (a) have the (I_k, D_k) couple under the threshold, plotted as a function of k; (b) have a set with k elements under the threshold, plotted as a function of k.

the presence of many intermediate events having only one (I_k, D_k) couple under the threshold (Figure 3.9). Finally, TSS ranges from 0.74 to 0.82, its maximum value is reached with classification method 1 and its variability is mainly due to the variations in $POFD$ rather than those in POD.

3.6 Discussion

3.6.1 Consistency of the approaches

Although the two validation analyses carried out in this work are based on quite different approaches, consistency in the classification of the rainfall events related to the observed debris flows can be highlighted:

- The two FN obtained in the back-analysis validation approach (Figure 3.6) correspond to the two rainfall events classified always as FN in the potential debris flow validation approach (Figure 3.10). This means, for both the events, not only the rainfall conditions identified by applying the BDA are under the threshold, but also the whole (I_k, D_k) sets are.

- The only intermediate rainfall event associated with a debris flow, resulting from the potential debris flow validation approach, matches to the rainfall event described by the (I, D) couple located about on the threshold in the BDA-based approach (Figure 3.6).

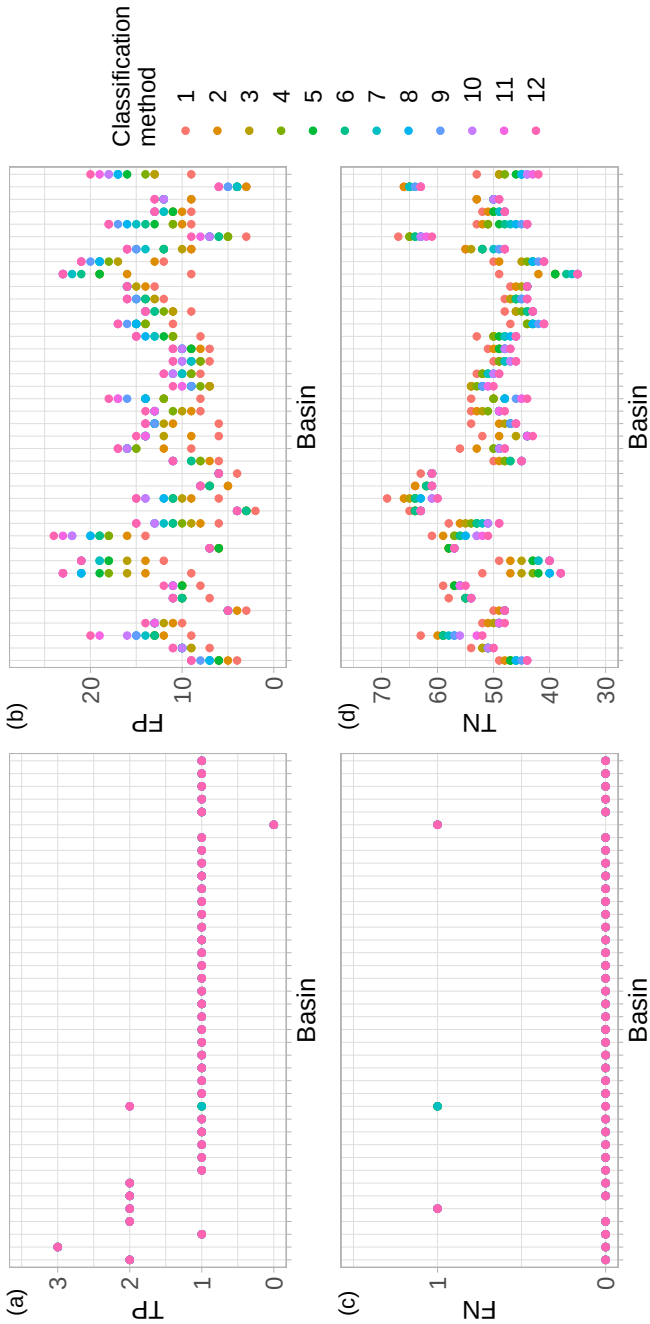

Figure 3.10: Contingencies for each basin and classification method.

Classification method	TP	FP	FN	TN	POD	POFD	TSS
1	46	309	3	2193	0.94	0.12	0.82
2	46	385	3	2117	0.94	0.15	0.78
3	46	424	3	2078	0.94	0.17	0.77
4	46	444	3	2058	0.94	0.18	0.76
5	46	465	3	2037	0.94	0.19	0.75
6	46	478	3	2024	0.94	0.19	0.75
7	46	494	3	2008	0.94	0.20	0.74
8	47	506	2	1996	0.96	0.20	0.76
9	47	520	2	1982	0.96	0.21	0.75
10	47	531	2	1971	0.96	0.21	0.75
11	47	546	2	1956	0.96	0.22	0.74
12	47	556	2	1946	0.96	0.22	0.74

Table 3.4: Contingencies and skill scores obtained at the study area scale for each classification method.

This consistency demonstrates the coherence between the results of the two validation approaches.

3.6.2 Reliability of the threshold

Both the validation approaches provide an assessment and quantification of the reliability of the threshold. The BDA-based validation approach emphasizes the threshold's satisfactory reliability. Indeed, a value of POD equals 0.96 highlights that the threshold identifies adequately the region of the $I - D$ plane where the rainfall conditions are likely to result in a debris flow and the low number of FN indicates that the threshold is not too high. On the other hand, the analysis does not provide an estimate of false positives, since it considers only rainfall events related to observed debris flows, and, therefore, it is not possible to verify whether the threshold is too low.

Because it can classify both rainfall events that result in a debris flow and those that do not, the potential debris flows validation approach provides a more complete description of the threshold reliability than the previous method, allowing some considerations on the threshold both at the basin and study area scale. Considering the values of the contingencies obtained in each catchment (Figure 3.10), we can say that some basins present a relatively large number of FP (Figure 3.10(b)) suggesting that the threshold is too low for these catchments. Conversely, for the three cases characterized

by a number of FN equal to 1 (two basins for all classification methods while one for some of them) (Figure 3.10(c)) the threshold seems to be too high, as it was not able to correctly classify (always or in some cases) the rainfall events associated with debris flow events that actually happened. However, the limited number of observed events in each catchment prevents these results from being sufficiently robust.

On the contrary, at the study area scale, the number of events is considered sufficient to quantify the actual reliability of the threshold. As evident from Table 3.4, the threshold performance varies according to the classification method due to the presence of the intermediate events, which represents 10% of the analyzed rainfall events. In particular, the classification of the intermediate rainfall events affects the threshold performance mainly in terms of $POFD$ and TSS. In any case, the values of all three skill scores (i.e. POD, $POFD$ and TSS) can be deemed satisfactory, regardless of the classification method considered. In addition, the resulting skill scores values are consistent with those obtained in other studies of rainfall thresholds validation based on contingencies and skill scores computation (e.g. Gariano et al., 2015, Jordanova et al., 2020, Piciullo et al., 2017).

Hence, the results of the potential debris flows validation approach confirm the results of the BDA-based one regarding the satisfactory capability of the threshold in classifying the rainfall events that resulted in debris flows. In addition to this, the second validation approach points out that the threshold has also satisfactory reliability in classifying the rainfall events that did not result in debris flows.

It is worth noting that the POD values, obtained in both the BDA-based and the potential debris flow validation approach, are consistent with the non-exceedance probability level of the threshold. Indeed, in Rosatti et al. (2019), the threshold is calibrated applying the frequentist method and imposing a non-exceedance probability equal to 5%. This means that the threshold is estimated by accepting that there is a 5% of probability to experience a debris flow related to rainfall conditions below the threshold (Brunetti et al., 2010). Hence, in both validation approaches, the relative frequencies of occurred events not correctly classified by the threshold (i.e. $1 - POD$) are coherent with the non-exceedance probability set in calibration. This consistency further confirms the reliability of the threshold.

3.6.3 Potential forecasting use of the threshold

The validation based on the potential debris flow approach has set the stage for the use of the calibrated threshold in a predictive framework. Indeed, the possibility to describe an event in terms of potential debris flows has

eliminated the need to know ex-ante the debris flow occurrence. This means that this method, used in this work only in a back-analysis context, can be applied also in a predictive framework to compare the (I_k, D_k) set associated to a rainfall event with the calibrated threshold.

In this context, the skill scores can be used to optimize the prediction capability. Usually, this optimization is performed considering several thresholds with variable non-exceedance probability to select the *optimal* one (e.g. Gariano et al., 2015, Piciullo et al., 2017), namely the one that maximises the chosen skill score or combination of skill scores. In general, this choice is not obvious since it depends on the main objective of the analysis. For instance, if the main purpose is to minimize the loss of credibility in the forecast of debris flows occurrence, namely, to minimize FN, the skill score to maximize is POD. Within this approach, given the threshold, optimization can be used to select the most performing classification method. According to Leonarduzzi and Molnar (2020) and Peres and Cancelliere (2021), it is advisable to optimize the TSS value since it is an indicator of the overall forecast ability of the threshold. By using this criterion, classification method 1 turns out to be the *optimal* one in predictive terms, i.e. the best TSS value is obtained when all intermediate rainfall events are classified as under the threshold. Combining this result with the fact that most of the intermediate rainfall events have, at least, the (I_1, D_1) couple under the threshold (Figure 3.9), it may be inferred that one of the key characteristics of a rainfall event to result in a debris flow is having at least the maximum rainfall intensity exceeding the threshold. However, because the number of FP is not equal to zero even with classification method 1 (Table 3.4), this characteristic cannot be considered sufficient to observe a debris flow.

Finally, the (I_k, D_k) set can be interpreted in the BDA perspective: given a catchment and a rainfall event, each (I_k, D_k) couple can be seen as the water volume made available to the catchment for a potential debris flow. Then, by inverting Eq. (3.1), we can estimate the relevant solid volume we can expect at the outlet of the considered catchment. An example of this interpretation is provided in Figure 3.11 in which, for each (I_k, D_k) couple of the considered rainfall event, the available rainfall volume and the related potential solid volume are quantified.

Clearly, exceeding the threshold is not sufficient to "guarantee" that the debris-flow occurs since other site-specific predisposing and triggering factors, such as the presence of loose material and soil moisture (e.g. Mostbauer et al., 2018, Zhao et al., 2021) must be present. Nevertheless, the information on the possible occurrence and the estimate of the size of the debris flow can be useful in an EWS.

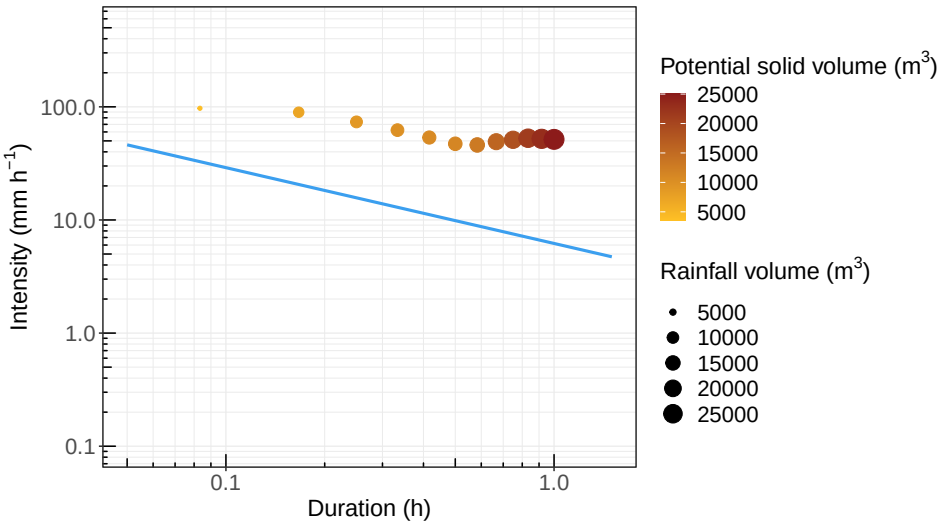

Figure 3.11: (I_k, D_k) couples of the potential debris flow validation approach interpret in the BDA perspective. For each (I_k, D_k) the related rainfall and potential solid volumes are computed according to the BDA method.

3.7 Conclusions

In this work, two different approaches, namely the BDA-based and the potential debris flows validation approach, are introduced to perform the validation of the physically-based threshold calibrated in Rosatti et al. (2019). The threshold reliability is then quantified by applying these approaches to a proper dataset.

The results of the BDA-based validation approach reveal that the threshold is satisfactorily capable of identifying the $I - D$ plane region where rainfall conditions are likely to result in a debris flow. The potential debris flow validation approach confirms this result and, in addition, highlights the threshold's reliability in classifying the rainfall events that did not result in debris flows.

Moreover, the introduction of the potential debris flow concept allows us to put the threshold in a predictive framework by letting to place a rainfall event in the $I - D$ plane regardless of whether or not it is associated with a (already occurred) debris flow. In this context, classification method 1 results to be the *optimal* one in terms of forecast and this highlights the potential key role of the maximum intensity for a debris flow occurrence. Furthermore, given a catchment and a rainfall event, the interpretation of the potential debris flow set in the BDA perspective allows us to provide

an estimate of the solid volume we may expect at the basin outlet for each couple of the set. Although this estimate is rough, it might be helpful in an EWS to hypothesize the order of magnitude of what could happen if the threshold is exceeded.

Hence, overall, the results of this work show that the BDA method, combined with the potential debris flow approach, is promising to provide a reliable threshold usable in a forecasting framework. Nevertheless, to further verify this aspect, it is advisable to calibrate, by applying the BDA method, and validate, by using the process proposed in this work, other thresholds for different study areas. In addition, these further studies should be based on accurate data, collected specifically for this type of analysis, to reduce uncertainty as much as possible. In addition, careful consideration should be given to the choice of the spatial scale for calibration and validation of the threshold. In Rosatti et al. (2019) and this work, the study area's scale was chosen primarily based on administrative borders. Nevertheless, in further studies, it would be advisable to develop some robust criteria for identifying a region that is homogeneous in terms of features that affect the BDA-based threshold, such as precipitation regime, type of predominant storms, morphology, lithology and debris flow characteristics.

Part II

Multivariate rainfall analysis with multi-hazard mapping purpose

Chapter 4

Multivariate rainfall analysis for multi-hazard assessment of combined debris flow-flood events in a mountain confluence

4.1 Introduction

In mountain regions, hydrogeological hazards are quite widespread usually rainfall-induced phenomena that can have severe consequences for the impacted areas, such as damages and casualties (e.g. Barredo, 2009, Dowling and Santi, 2014, Jonkman and Kelman, 2005, Petrucci et al., 2019). As a result, the development and improvement of mitigation strategies aimed to reduce the hydrogeological risk are crucial.

Hazard mapping is an example of a mitigation strategy in this context, as it identifies the zones of a region susceptible to a hydrogeological hazard, allowing for more suitable land use planning and management, as well as the creation of emergency response plans (Dottori et al., 2016, Mudashiru et al., 2021). Usually, hydrogeological hazard maps are constructed based on scenarios (Teng et al., 2017, van Westen and Greiving, 2017) characterized by varying occurrence probability and related magnitude of the examined phenomenon or, in the lack of appropriate historical data on the phenomenon, the corresponding main triggering factor, namely rainfall (Serinaldi, 2009).

Many areas are potentially subjected to several hydrogeological hazards

(De Angeli et al., 2022, Liu et al., 2016) and their assessment with mapping purpose is usually performed assuming all hazards as independent (Gallina et al., 2016, Gill and Malamud, 2016, Javidan et al., 2021, Ming et al., 2022, van Westen and Greiving, 2017). However, extreme rainfall may affect different locations at the same time. This could result in several hydrogeological phenomena in close proximity in time and space (Dave et al., 2021, Petrucci and Polemio, 2003, Revellino et al., 2019) and their interactions may greatly magnify their single impacts on the investigated area (Aceto et al., 2017, Bevacqua et al., 2017, Hao et al., 2018, Liu et al., 2016, Tilloy et al., 2020). For instance, the confluence of a mountain stream susceptible to debris flows and a valley river prone to floods is an area potentially subjected to coupled multi-hazard. The simultaneous occurrence of a debris flow and a flood, mainly due to the co-occurrence of heavy rainfalls in the relevant catchments, can lead to potential interactions between the phenomena at the confluence. Debris flow can affect the river flow and the river flood can influence the deposition zone of the debris flow, resulting in compound effects that cannot be assessed by analyzing each phenomenon separately.

In this context, a multi-hazard approach is required to assess the possible co-occurrence of hydrogeological phenomena and construct scenarios to evaluate the related compound effects. The multi-hazard approach has to be based on a multivariate frequency analysis that jointly considers rainfall extremes at relevant locations for the investigated phenomena of the considered area (Le et al., 2019, Thibaud et al., 2013), taking into account their dependence structure.

In literature, several approaches exist to assess the spatial dependence between rainfall extremes, including hierarchical Bayesian models (e.g. Cooley and Sain, 2010, Cooley et al., 2007, Dyrrdal et al., 2015, Schliep et al., 2010, Sharkey and Winter, 2019), max-stable models (e.g. Coles, 1993, Le et al., 2018, Stephenson et al., 2016, Thibaud et al., 2013) and copula-based models (Han et al., 2020, Zhu et al., 2020). In general, unlike the first two types of models, copula-based models allow to deal with the assessment of the dependence structure between the investigated variables, by means of copula, separately and independently from their univariate analysis (Brunner et al., 2019, Genest and Favre, 2007). Furthermore, copula-based models do not impose any restrictions on the distributions used to characterize the variables' univariate behavior (Huang et al., 2018). Thanks also to these advantages, copula-based models have been widely used for multivariate analyses in the hydrological and water management fields in recent years (e.g. Bevacqua et al., 2017, Brunner et al., 2019, De Michele et al., 2005,

Gao et al., 2021, Graler et al., 2013, Hao and Singh, 2016, Jane et al., 2020, Sadegh et al., 2018, Tavakol et al., 2020, Volpi and Fiori, 2012). Nevertheless, only a few studies applied a copula-based approach to construct a multivariate model to assess rainfall extremes at different locations and, to the best of the authors' knowledge, only bivariate models have been developed. However, jointly considering rainfall extremes at all relevant locations for the investigated phenomena allows you to model the spatial variability of both their values and dependence structure, as well as estimate overall scenarios.

In order to simulate the dependence structure between rainfall values at more than two locations, a high-dimensional copula is required. Constructing high-dimensional copulas is widely acknowledged to be a difficult problem (Aas et al., 2009). In literature, there is a wide range of well-investigated bivariate copulas (e.g. Joe, 1997, Nelsen, 2006) while the number of copulas with a dimension greater than two are rather limited (e.g. Aas et al., 2009, Brechmann and Schepsmeier, 2013, Graler et al., 2013). Moreover, high-dimensional copulas lack flexibility needed to model, if present, an heterogeneous dependence structure between the investigated variables (Brechmann and Schepsmeier, 2013). To overcame these limitations, Joe (1996) introduced the pair-copula construction (PCC). PCC is a building method that allows multivariate copulas to be constructed using just (conditional) bivariate copulas. The result is very flexible multivariate copulas able to model arbitrary complex dependence structures (Aas et al., 2009, Bedford and Cooke, 2002, Czado, 2010).

As regards the univariate extremes rainfall analysis, the classical extreme value theory (EVT) is usually applied. EVT aims to characterize the extreme behavior of the investigated process describing it through an asymptotic probability distribution (e.g. Serinaldi et al., 2020). This is possible under the assumptions that (i) the rainfall values are independent and identically distributed and (ii) the number of realization of the investigated process is large enough (i.e. tending to infinity) in each considered year (e.g. Davison and Smith, 1990, Marani and Ignaccolo, 2015). To meet the independence assumption, commonly rainfall values are selected applying the annual maxima approach or the peak-over-threshold approach (e.g. Coles et al., 2001, Volpi et al., 2019). As a result, most of the observations (and related information) of the investigated process are discarded from the analysis (Serinaldi et al., 2020, Zorzetto et al., 2016). Moreover, the number of yearly independent realization of the investigated process from which the extremes are selected is necessarily limited, preventing the fulfilment of the asymptotic assumption (Marra et al., 2019). As an alternative to the clas-

sical EVT, the Metastistical Extreme Value (MEV) framework, introduced by Marani and Ignaccolo (2015), and the related Simplified Metastatistical Extreme Value (SMEV) formulation, presented in Marra et al. (2019), relax the asymptotic assumption and permit us to make advantage of all information provided by the observations of the investigated process (Zorzetto et al., 2016). Assessment of the performance of the MEV and SMEV formulations with respect to the classical EVT can be found in Formetta et al. (2022), Marani and Ignaccolo (2015), Marra et al. (2018, 2019, 2020, 2022), Miniussi and Marani (2020), Schellander et al. (2019), Zorzetto et al. (2016). Although their adequacy and robustness have been demonstrated, both MEV and SMEV formulations have not yet been applied in a multivariate framework.

This work aims to develop a multivariate copula-based model to jointly assess rainfall values at different locations with a multi-hazard analysis purpose. To this end, a confluence between a creek susceptible to debris flows and a flood-prone river is considered as study area. Some rain gauges relevant to the phenomena under investigation are selected and a strategy to identify synchronous events in all locations is introduced. The univariate extremes rainfall analysis at each rain gauge is performed by applying the SMEV formulation while the PCC is adopted to obtain the multivariate copula describing the dependence structure between the investigated variables. The multivariate model thus obtained is used to estimate rainfall scenarios related to different definitions and values of the multivariate return period. Finally, the possible use of the obtained scenarios for hazard mapping purposes is discussed.

4.2 Basic theoretical framework

4.2.1 Copulas

A copula C of d-dimension is a multivariate distribution defined on $\mathbb{I}^d = [0,1]^d$ with uniform marginal distributions (e.g. Czado et al., 2012, Salvadori et al., 2011). A joint cumulative distribution function F of d random variables $X_1, \ldots, X_d$, with marginal distributions $F_1(x_1), \ldots, F_2(x_2)$, is defined as (e.g. Durante and Sempi, 2016, Nelsen, 2006):

$$F(x_1, \ldots, x_d) = P\{X_1 < x_1, \ldots, X_d < x_d\} \tag{4.1}$$

According to the Sklar's theorem (Sklar, 1959), there exists a connection between F and the related C in the form:

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \tag{4.2}$$

for all $(x_1, \ldots, x_d) \in \mathbb{R}^d$. If then $F_1(x_1), \ldots, F_2(x_2)$ are continuous, C is unique and can be written as:

$$C(u_1, \ldots, u_d) = F\left(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\right) \tag{4.3}$$

in which $F_k(x_k) = u_k$ with $k = 1, \ldots, d$, namely $U_k = F_k(X_k)$ with U_k uniformly distributed on $\mathbb{I}$ (e.g. Angus, 1994, Nelsen et al., 2001), and $(u_1, \ldots, u_d) \in \mathbb{I}^d$. The related copula density is defined as:

$$c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \ldots \partial u_d} \tag{4.4}$$

By differentiating Eq. (4.2), it's also possible to obtain the joint probability density function (PDF):

$$f(x_1, \ldots, x_d) = c(u_1, \ldots, u_d) \cdot f_1(x_1) \cdot \ldots \cdot f_d(x_d) \tag{4.5}$$

in which $f_1(x_1), \ldots, f_d(x_d)$ are the marginal PDF.

Moreover, it is useful to introduce the concept of empirical copula. To this end, first it is necessary to define the empirical marginal distributions. Given a random sample $X_1^k, \ldots, X_d^k$, with $k = 1, \ldots, s$, from the random variables $X_1, \ldots, X_d$, the empirical marginal CDF are given by:

$$F_{ni}(x_i) = \frac{1}{s+1} \sum_{k=1}^{s} 1_{[X_i^k \leq x_i]} \tag{4.6}$$

in which $1_{[\ldots]}$ is the indicator function of the set $[\ldots]$ and the subscript n indicates that the function is empirical. The empirical copula is hence defined as:

$$C_n(u_1, \ldots, u_d) = \frac{1}{s} \sum_{k=1}^{s} 1_{[U_1^k \leq u_1, \ldots, U_d^k \leq u_d]} \tag{4.7}$$

where $U_i^k = F_{ni}\left(X_i^k\right)$ for all $k = 1, \ldots, s$ and $(u_1, \ldots, u_d) \in \mathbb{I}^d$ (e.g. Tsukahara, 2005).

In this context, it is worth considering two fundamental copulas (e.g. Durante and Sempi, 2016, Genest and Favre, 2007):

- the *comonotonicity* copula that describes a perfect positive dependence (as following defined in Sect. 4.2.2) between variables $X_1, \ldots, X_d$ and is defined as:

$$M(u_1, \ldots, u_d) = \min(u_1, \ldots, u_d) \tag{4.8}$$

- the *independence* copula that is the copula corresponding to stochastically independent random variable $X_1, \ldots, X_d$ and is defined as:

$$\Pi\left(u_1, \ldots, u_d\right) = \prod_{j=1}^{d} u_j \tag{4.9}$$

In addition, considering $F\left(x_1, \ldots, x_d\right)$ as a random variable, it possible to introduce the *Kendall* function of $X_1, \ldots, X_d$ defined as:

$$K_C\left(t\right) = P\left\{F\left(x_1, \ldots, x_d\right) \leq t\right\} = P\left\{C(u_1, \ldots, u_d) \leq t\right\} \tag{4.10}$$

in which $t \in \mathbb{I}$ (e.g. Nelsen et al., 2001, 2003). Hence, the Kendall function is nothing but the univariate distribution of C (Joe, 2014). For example, considering the comonotonicity copula, the related Kendall function is (e.g. Garcin et al., 2018):

$$K_M\left(t\right) = t \tag{4.11}$$

while the d-dimensional independence copula one is:

$$K_\Pi\left(t\right) = t + t\sum_{j=1}^{d-1} \frac{\left(-\log t\right)^j}{j!} \tag{4.12}$$

According to Genest and Rivest (1993), given a random sample $X_1^k, \ldots, X_d^k$, with $k = 1, \ldots, s$, from the random variables $X_1, \ldots, X_d$ the empirical version of the Kendall function is estimated introducing:

$$Q_i = \frac{1}{s-1}\sum_{i \neq j} 1_{\left[X_1^j \leq X_1^i, \ldots, X_d^j \leq X_d^i\right]} \tag{4.13}$$

for all $i, j = 1, \ldots, s$. Q_i represents the proportion of the sample in the rectangle with the origin $(0, \ldots, 0)$ and $(X_1^i, \ldots, X_d^i)$ as lower-left and upper-right corners, respectively. The empirical Kendall function can be then estimated as:

$$K_n\left(t\right) = \frac{1}{s}\sum_{i=1}^{s} 1_{\left[Q_i \leq t\right]} \tag{4.14}$$

with $t \in \mathbb{I}$.

A function strictly related to the Kendall one is the λ-function that is defined as:

$$\lambda\left(t\right) = t - K_C\left(t\right) \tag{4.15}$$

with $t \in \mathbb{I}$ (e.g. Michiels et al., 2011) while its empirical version is:

$$\lambda_n\left(t\right) = t - K_n\left(t\right) \tag{4.16}$$

In terms of exceedance probabilities, it's possible to introduce the concept of joint *survival* CDF defined as:

$$\overline{F}\left(x_1, \ldots, x_d\right) = P\left\{X_1 > x_1, \ldots, X_d > x_d\right\} \tag{4.17}$$

that can be written in terms of copula as:

$$\overline{F}\left(x_1, \ldots, x_d\right) = \hat{C}\left(\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)\right) \tag{4.18}$$

in which $\hat{C}$ is the survival copula of the random variables $X_1, \ldots, X_d$ and $\overline{F}_k(x_k) = 1 - F_k(x_k)$ with $k = 1, \ldots, d$ is the survival function of the marginal distribution of X_k (e.g. Salvadori et al., 2016). According to Joe (2014), by applying the inclusion-exclusion probability law and introducing the survival function $\overline{C}$ related to C defined as:

$$\overline{C}\left(1 - u_1, \ldots, 1 - u_d\right) = \hat{C}\left(u_1, \ldots, u_d\right) \tag{4.19}$$

for all $(u_1, \ldots, u_d) \in \mathbb{I}^d$, $\hat{C}$ can be express in terms of C as:

$$\overline{C}\left(u_1, \ldots, u_d\right) = 1 - \sum_{i=1}^{d} u_i + \sum_{S \in \mathcal{P},\, \#S \geq 2} (-1)^{\#S} C_S\left(u_i \,:\, i \in S\right) \tag{4.20}$$

in which $\mathcal{P}$ is the set of all subsets of $\{1, 2, \ldots, d\}$, $\#S$ represents the cardinality of S and $C_S\left(u_i \,:\, i \in S\right)$ is the marginal copula of C of dimension $\#S$ (Salvadori et al., 2013). For example, considering the bivariate case, namely assuming $d = 2$, the relation between $\hat{C}$, $\overline{C}$ and C is:

$$\hat{C}\left(u_1, u_2\right) = \overline{C}\left(1 - u_1, 1 - u_2\right) = u_1 + u_2 - 1 + C\left(1 - u_1, 1 - u_2\right) \tag{4.21}$$

with $(u_1, u_2) \in \mathbb{I}^2$.

Finally, different copula classes exist in the literature, with the elliptical and archimedean classes being the most studied and used in applications (Durante et al., 2015). As a result, these copulas classes are considered in this study. Table 4.1 shows the bivariate expression of the main copula families belonging to the elliptical and archimedean classes: Gaussian and t copula are elliptical copulas, Clayton, Gumbel, Frank and Joe are one-parameter archimedean copulas, while the BB copulas are two parameters archimedean copulas.

We refer the reader to Nelsen (2006), Joe (2014) and Durante and Sempi (2016), and reference therein, for more details on copulas theory and properties.

Name	$C(u_1, u_2)$	Parameter(s) range
Gaussian	$\int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\theta^2}} \exp\left(\frac{2\theta x_1 x_2 - x_1^2 - x_2^2}{2(1-\theta^2)}\right) dx_1 dx_2$	$\theta \in [-1, 1]$
t	$\int_{-\infty}^{t_{\theta_2}^{-1}(u_1)} \int_{-\infty}^{t_{\theta_2}^{-1}(u_2)} \frac{\Gamma(\frac{\theta_2+2}{2})}{\Gamma(\frac{\theta_2}{2})\sqrt{(\pi\theta_2)^2(1-\theta_1^2)}} \left(1 + \frac{x_1^2 - 2\theta_1 x_1 x_2 + x_2^2}{\theta_1(1-\theta_2^2)}\right)^{-\frac{\theta_1+2}{2}} dx_1 dx_2$	$\theta_1 \in [-1, +1]$ $\theta_2 \in (0, +\infty)$
Clayton	$\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-\frac{1}{\theta}}$	$\theta \in [-1, +\infty) \backslash \{0\}$
Gumbel	$\exp\left[-\left((-\ln u_1)^\theta + (-\ln u_2)^\theta\right)^{\frac{1}{\theta}}\right]$	$\theta \in [1, +\infty)$
Frank	$\frac{1}{\theta} \ln\left(1 - \frac{(1-e^{-\theta u_1})(1-e^{-\theta u_2})}{1-e^{-\theta}}\right)$	$\theta \in \mathbb{R} \backslash \{0\}$
Joe	$1 - \left[(1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta\right]^{\frac{1}{\theta}}$	$\theta \in [1, +\infty)$
BB1	$\left[1 + \left((u_1^{-\theta_1} - 1)^{\theta_2} + (u_2^{-\theta_1} - 1)^{\theta_2}\right)^{\frac{1}{\theta_2}}\right]^{-\frac{1}{\theta_1}}$	$\theta_1 \in (0, +\infty)$ $\theta_2 \in [1, +\infty)$
BB6	$1 - \left[1 - \exp\left(-\left((-\ln(1-(1-u_1)^{\theta_1}))^{\theta_2} + (-\ln(1-(1-u_2)^{\theta_1}))^{\theta_2}\right)^{\frac{1}{\theta_2}}\right)\right]^{\frac{1}{\theta_1}}$	$\theta_1 \in [1, +\infty)$ $\theta_2 \in [1, +\infty)$
BB7	$1 - \left[1 - \left((1-(1-u_1)^{\theta_1})^{-\theta_2} + (1-(1-u_2)^{\theta_1})^{-\theta_2} - 1\right)^{-\frac{1}{\theta_2}}\right]^{\frac{1}{\theta_1}}$	$\theta_1 \in [1, +\infty)$ $\theta_2 \in (0, +\infty)$
BB8	$\theta_2^{-1}\left[1 - \left(1 - (1-(1-\theta_2)^{\theta_1})^{-1}(1-(1-\theta_2 u_1)^{\theta_1})(1-(1-\theta_2 u_2)^{\theta_1})\right)^{\frac{1}{\theta_1}}\right]$	$\theta_1 \in [1, +\infty)$ $\theta_2 \in (0, 1]$

Table 4.1: Bivariate elliptical and archimedean copula families along with the related parameter(s) range. In Gaussian copula, Φ^{-1} is the inverse of a standard univariate Gaussian distribution. In the t-copula, $t_{\theta_2}^{-1}$ is the inverse of the univariate t distribution with θ_2 degrees of freedom and $\Gamma(\cdot)$ is the gamma function.

4.2.2 Dependence measures

According to Balakrishnan and Lai (2009), a dependence measure reveals, in some way, how closely two (or more) random variables are related. In literature, several measures of dependence exist (see e.g. Joe, 1997, Nelsen, 2006) and each is able to quantify only an aspect of the dependence structure of the investigated variables. In this context, among others, the Kendall's rank correlation coefficient (or Kendall's tau) and the tail dependence are considered. In the following, since the method used to construct the multivariate distribution (i.e. Pair-Copula Construction - PCC) is based on bivariate copulas (as described in Section 4.2.3), we briefly introduce these dependence measures considering the bivariate case.

Kendall's tau is a non parametric measure that quantify the concordance between variables based on ranks (e.g. Genest and Favre, 2007). Considering a random sample $(X_1, Y_1), \ldots, (X_s, Y_s)$ of the variables (X, Y), the pairs (X_i, Y_i), (X_j, Y_j) are defined concordant if $(X_i - X_j)(Y_i - Y_j) > 0$ otherwise if $(X_i - X_j)(Y_i - Y_j) < 0$ the two pairs are discordant. Hence, the empirical version of the Kendall's tau is defined as:

$$\tau_n = \frac{P - Q}{\binom{s}{2}} = \frac{4}{s\,(s-1)} P - 1 \tag{4.22}$$

in which P and Q are the number of pairs concordant and discordant, respectively, and $\tau_n \in [-1, 1]$ (Genest and Favre, 2007).

Kendall's tau can be also be written in terms of copula. Indeed, considering the variables (X, Y) with copula C and continuous marginal distributions, Kendall's tau can be express as:

$$\tau_C = 4 \int_{\mathbb{I}^2} C\,(u, v)\, dC\,(u, v) - 1 \tag{4.23}$$

with $(u, v) \in \mathbb{I}^2$ and $\tau_C \in [-1, 1]$. Kendall's tau is thus a property of the copula (e.g. Pappadà, 2014) and theoretical formulas of τ_C as a function of the copula parameter(s) exist for elliptical and archimedean copulas (see e.g. Frees and Valdez, 1998, Nadarajah et al., 2017).

As evident from both Eq. (4.22) and (4.23), Kendall's tau, and its empirical version, provides an overall concordance measure of observations and copula. However, if the extreme values of variables are the focus of the analysis, it's advisable to introduce other dependence measures that describe the degree of dependence of the observations and copula in the upper-right-quadrant tail and lower-left-quadrant tail (Frahm et al., 2005). The extremes degree of dependence can be estimated with the so-called tail

dependence coefficients (TDC), suggested in Sibuya et al. (1960) among others. Considering two random variables X and Y with continuous marginal CDF F_X and F_Y, respectively, the upper TDC λ_U and the lower TDC λ_L of (X, Y) are defined as:

$$\lambda_U = \lim_{t \to 1^-} P\left\{Y > F_Y^{-1}(t) \mid X > F_X^{-1}(t)\right\} \tag{4.24}$$

$$\lambda_L = \lim_{t \to 0^+} P\left\{Y \leq F_Y^{-1}(t) \mid X \leq F_X^{-1}(t)\right\} \tag{4.25}$$

with $\lambda_U,\ \lambda_L \in [0, 1]$. Hence, λ_U (λ_L) denotes the asymptotic limit of the probability that Y exceeds (does not exceed) a high (low) quantile given that X does (does not) (e.g. Durante et al., 2015). Introducing then the copula C of (X, Y), λ_U and λ_L can be written in terms of C as:

$$\lambda_U = \lim_{t \to 1^-} \frac{\hat{C}(1-t, 1-t)}{1-t} = \lim_{t \to 1^-} \frac{1 - 2t + C(t, t)}{1-t} \tag{4.26}$$

$$\lambda_L = \lim_{t \to 0^+} \frac{C(t, t)}{t} \tag{4.27}$$

and, considering the related empirical copula C_n, their empirical versions are:

$$\lambda_{Un} = \lim_{t \to 1^-} \frac{1 - 2t + C_n(t, t)}{1-t} \tag{4.28}$$

$$\lambda_{Ln} = \lim_{t \to 0^+} \frac{C_n(t, t)}{t} \tag{4.29}$$

with $\lambda_{Un},\ \lambda_{Ln} \in [0, 1]$.

Hence, by definition, TDC, and their empirical version, provide an asymptotic approximation of the copula and observations behaviour in the tail. However, investigating the tail behavior at some points close to the unit square's (i.e. $\mathbb{I}^2$) corners can provide a more exhaustive overview on tail dependence (Durante et al., 2015). To this end, the so-called tail concentration function (TCF) (e.g. Venter et al., 2002), useful to provide a representation of the tail behavior of both copula and observations (Pappadà, 2014), is considered. The TCF of the copula C is defined as:

$$TCF_C(t) = \frac{C(t, t)}{t} \cdot 1_{(0, 0.5]} + \frac{1 - 2t + C(t, t)}{1-t} \cdot 1_{(0.5, 1)} \tag{4.30}$$

and the related empirical version is:

$$TCF_n(t) = \frac{C_n(t, t)}{t} \cdot 1_{(0, 0.5]} + \frac{1 - 2t + C_n(t, t)}{1-t} \cdot 1_{(0.5, 1)} \tag{4.31}$$

4.2.3 Pair-copula construction and vine copulas

The Pair-Copula Construction (PCC) is a flexible method to built multivariate copulas using exclusively bivariate copulas (e.g. Bedford and Cooke, 2001, 2002, Joe, 1996, Kurowicka and Cooke, 2006) by recursively decomposing a multivariate PDF into conditional densities products (Aas et al., 2009, Czado, 2010). Considering d random variables $X_1, \ldots, X_d$ with continuous marginal CDF $F_1, \ldots, F_d$ and joint PDF $f(x_1, \ldots, x_d)$, the latter can be factorised as:

$$f(x_1, \ldots, x_d) = f_d(x_d) \cdot f(x_{d-1}|x_d) \cdot f(x_{d-2}|x_{d-1}, x_d) \cdot \ldots \cdot f(x_1|x_2, \ldots, x_d) \tag{4.32}$$

in which $f(\cdot|\cdot)$ is the conditional density.

Considering then Eq. (4.5), it's possible to write each factor of the right-hand side of Eq. (4.32) as a product between a proper bivariate copula (i.e. *pair-copula*) density and a conditional density applying this general formula:

$$f(x|\boldsymbol{\nu}) = c_{x\nu_j|\boldsymbol{\nu}_{-j}}\left(F(x|\boldsymbol{\nu}_{-j}), F(\nu_j|\boldsymbol{\nu}_{-j})\right) \cdot f(x|\boldsymbol{\nu}_{-j}) \tag{4.33}$$

in which $c_{x\nu_j|\boldsymbol{\nu}_{-j}}$ is the appropriate bivariate copula density, $F(\cdot|\cdot)$ is the conditional CDF, $\boldsymbol{\nu}$ is a vector of d dimension, ν_j is a component of $\boldsymbol{\nu}$ chosen arbitrarily and $\boldsymbol{\nu}_{-j}$ indicates all the component of $\boldsymbol{\nu}$ but the chosen one (Aas et al., 2009). For instance, the factor $f(x_{d-2}|x_{d-1}, x_d)$, using Eq. (4.33), can be written as:

$$f(x_{d-2}|x_{d-1}, x_d) = c_{x_{d-2}x_{d-1}|x_d}\left(F(x_{d-2}|x_d), F(x_{d-1}|x_d)\right) \cdot f(x_{d-2}|x_d) \tag{4.34}$$

in which, according to the symbols used in Eq. (4.33), $x = x_{d-2}$, $\boldsymbol{\nu} = (x_{d-1}, x_d)$, $\nu_j = x_{d-1}$ and $\boldsymbol{\nu}_{-j} = (x_d)$. Then, similarly, $f(x_{d-2}|x_d)$ can be expressed as:

$$f(x_{d-2}|x_d) = c_{x_{d-2}x_d}\left(F_{d-2}(x_{d-2}), F_{d-1}(x_{d-1})\right) \cdot f_{d-2}(x_{d-2}) \tag{4.35}$$

Hence, substituting Eq. (4.35) in Eq. (4.34), the latter becomes:

$$\begin{aligned} f(x_{d-2}|x_{d-1}, x_d) = & c_{x_{d-2}x_{d-1}|x_d}\left(F(x_{d-2}|x_d), F(x_{d-1}|x_d)\right) \\ & \cdot c_{x_{d-2}x_d}\left(F_{d-2}(x_{d-2}), F_{d-1}(x_{d-1})\right) \cdot f_{d-2}(x_{d-2}) \end{aligned}$$

Therefore, writing each term of Eq. (4.32) using Eq. (4.33), the joint PDF is expresses as a product of pair-copulas, evaluated at marginal conditional CDF, and marginal PDF.

According to Joe (1996), for each $\nu_j \in \boldsymbol{\nu}$, the marginal conditional CDF can be written in the following generic form:

$$F(x|\boldsymbol{\nu}) = \frac{\partial C_{x\nu_j|\boldsymbol{\nu}_{-j}}\left(F(x|\boldsymbol{\nu}_{-j}), F(\nu_j|\boldsymbol{\nu}_{-j})\right)}{\partial F(\nu_j|\boldsymbol{\nu}_{-j})} \tag{4.36}$$

in which $C_{x\nu_j|\boldsymbol{\nu}_{-j}}$ is a bivariate copula. For instance, Eq. (4.36) implies:

$$F\left(x_{d-1}|x_d\right) = \frac{\partial C_{x_{d-1}x_d}\left(F\left(x_{d-1}\right), F\left(x_d\right)\right)}{\partial F\left(x_d\right)} \tag{4.37}$$

The factorisation of Eq. (4.32) is not unique. This means that the same joint PDF can be obtained applying different pair-copula constructions. Bedford and Cooke (2001, 2002) introduced a model to classify them in a graphical way considering copula densities and nested trees sequence. This method is named *vines* and the multivariate copulas built in this way are called *regular vines* or R-vines copulas (e.g. Czado, 2010, Joe and Kurowicka, 2011). A general d-dimensional R-vine structure is characterised by $d-1$ trees, i.e., T_j with $j = 1, \ldots, d-1$ and each tree T_j has $d+1-j$ nodes and $d-j$ edges. Each node represents a variable, in T_1, or a marginal conditional CDF, in the other trees, while the edges are pair-copulas.

Two R-vines subclasses are the so-called *canonical vines* or C-vines and *drawable vines* or D-vines. A C-vine structure is characterised by a root node for each tree to which the other nodes are connected. Each edge represent a pair-copula. A four-dimensional example of C-vine trees is shown in Figure 4.1. The general d-dimensional C-vine copula density is the following (e.g. Brechmann and Schepsmeier, 2013):

$$f\left(x_1, \ldots, x_d\right) = \prod_{k=1}^{d} f_k\left(x_k\right) \cdot \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{i\ i+j|1:(i-1)} \tag{4.38}$$

in which the following general abbreviation for a bivariate conditional copula density of x_i and x_j given $x_{i_1}, \ldots, x_{i_k}$ is introduced (Czado, 2010):

$$c_{i\ j|i_1,\ldots,i_k} = c_{i\ j|i_1,\ldots,i_k}\left(F\left(x_i|x_{i_1}, \ldots, x_{i_k}\right), F\left(x_j|x_{i_1}, \ldots, x_{i_k}\right)\right) \tag{4.39}$$

with $i, j, i_1, \ldots, i_k$ arbitrary distinct indices with $i < j$ and $i_1 < \ldots < i_k$.

Instead, a D-vine has trees with a path structure and the general d-dimensional D-vine copula density is the following (e.g. Brechmann and Schepsmeier, 2013):

$$f\left(x_1, \ldots, x_d\right) = \prod_{k=1}^{d} f_k\left(x_k\right) \cdot \prod_{i=1}^{d-1} \prod_{j=1}^{d-i} c_{j\ j+1|(j+1):(j+i-1)} \tag{4.40}$$

An example of a four-dimensional D-vine trees is shown in Figure 4.2.

We refer the reader to Joe (1996), Bedford and Cooke (2001), Bedford and Cooke (2002) and Kurowicka and Cooke (2006) for more details on PCC and vines copulas.

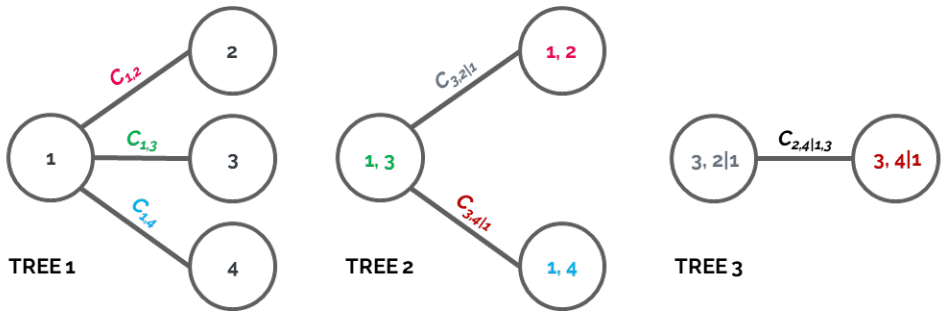

Figure 4.1: Example of four-dimensional C-vine trees. The numbers indicate the subscript of the variables (in tree 1) and the subscript of the variables involved in the marginal conditional CDF (in trees $2, 3$).

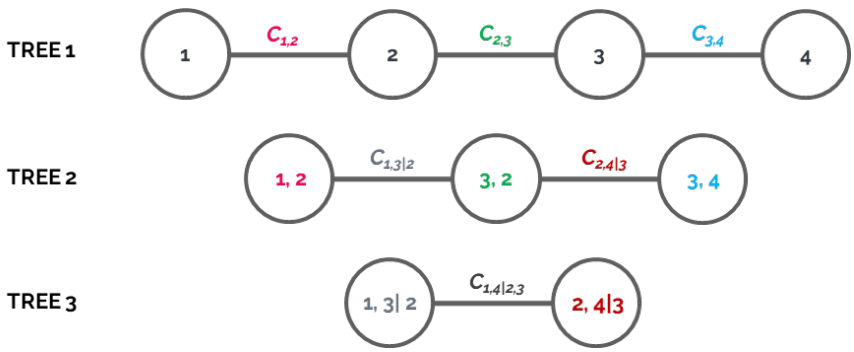

Figure 4.2: Example of four-dimensional D-vine trees. The numbers indicate the subscript of the variables (in tree 1) and the subscript of the variables involved in the marginal conditional CDF (in trees $2, 3$).

4.2.4 Simplified Metastatistical Extreme Value

The Simplified Metastatistical Extreme Value (SMEV) framework, introduced in Marra et al. (2019), is a simplification of the original Metastatistical Extreme Value (MEV), presented in Marani and Ignaccolo (2015), and is based on the assumption that extremes emerge from the distribution of the *ordinary* events, i.e. from the distribution of all the independent samples of the stochastic process under investigation (Formetta et al., 2022, Marani and Ignaccolo, 2015). According to the SMEV formulation, the cumulative distribution function (CDF) of rainfall extreme values (i.e. the yearly non-exceedance probabilities) can be written as:

$$\Omega\left(x\right) = G\left(x, \boldsymbol{\theta}\right)^n \tag{4.41}$$

in which $G(x, \boldsymbol{\theta})$ is the cumulative distribution of the ordinary events, $\boldsymbol{\theta}$ is the vector of the distribution parameters and n is the annual average number of ordinary events (Miniussi and Marra, 2021). The formulation of Eq. (4.41) is obtained by neglecting the inter-annual variability in the ordinary events distribution and number, originally taking into account in the MEV approach (Marra et al., 2019). Thanks to this simplification, the parameters of the distribution can be estimated considering all ordinary events of the time series, allowing you to:

- use short rainfall records (Marra et al., 2018, 2022);
- obtain a robust estimate of the parameters (Marra et al., 2020, Miniussi and Marani, 2020, Schellander et al., 2019);
- focus on representing the right tail of the distribution (i.e. extremes) through left-censoring the ordinary events (Marra et al., 2019, 2020).

Moreover, as highlighted in Serinaldi et al. (2020), the SMEV approach is equivalent to a maxima standard distribution considering n-size blocks with independence condition.

Previous SMEV application shown that a two-parameter Weibull distribution is suitable to describe the right tail of the precipitation ordinary events distribution (Formetta et al., 2022, Marra et al., 2019, Miniussi and Marra, 2021). The CDF of a two-parameter Weibull distribution can be written as:

$$G\left(x; \lambda, \kappa\right) = 1 - e^{-\left(\frac{x}{\lambda}\right)^{\kappa}} \tag{4.42}$$

in which λ and κ are the scale and shape parameter, respectively, and both of them vary in the interval $]0, +\infty[$ (Marra et al., 2019). Hence, using Eq.

(4.42) in Eq. (4.41), the rainfall extremes CDF becomes:

$$\Omega\left(x\right) = \left[1 - e^{-\left(\frac{x}{\lambda}\right)^{\kappa}}\right]^{n} \tag{4.43}$$

which is nothing but an exponential Weibull CDF (e.g. Pal et al., 2006).

We refer the reader to Marra et al. (2019) for more details on the SMEV formulation.

4.2.5 Multivariate return period

In a multivariate framework, several definitions of return period exist (e.g. Salvadori et al., 2016, Serinaldi, 2015, Vandenberghe et al., 2011). In general, the return period T can be express as (Serinaldi, 2015):

$$T = \frac{\mu}{p} \tag{4.44}$$

in which μ is the average inter-arrival time between two process realizations and p is the probability that a process realization is greater than an arbitrary value. Considering the SMEV approach (i.e. Eq. (4.43)), Eq. (4.44) becomes:

$$T = \frac{1}{1 - t^{n}} \tag{4.45}$$

in which t is the non-exceedance probability related to the ordinary events distribution.

Considering the general formula of Eq. (4.44), as highlighted in Tilloy et al. (2020), the expression (and meaning) of T changes based on the type of exceedance probability p considered, namely on how the investigated random variables combine and which type of compound events they yield. In this context, among others, three different types of exceedance probability are considered that, introducing d random variables $X_1, \ldots, X_d$ with marginal distributions $F_1(x_1), \ldots, F_2(x_2)$ and copula C, can be generally defined as follow (e.g. Salvadori et al., 2016, Serinaldi, 2015, Tilloy et al., 2020):

- the *OR* exceedance probability:

$$\begin{aligned} p_{OR} &= P\{X_1 > x_1 \vee \ldots \vee X_d > x_d\} \\ &= 1 - C\left(F_1(x_1), \ldots, F_d(x_d)\right) \\ &= 1 - C\left(u_1, \ldots, u_d\right) \end{aligned} \tag{4.46}$$

  that can also be express as:

$$p_{OR} = P\{(u_1, \ldots, u_d) \in \mathcal{S}_{\mathbf{u}}^{\vee}\} \tag{4.47}$$

in which

$$\mathcal{S}_{\mathbf{u}}^{\vee} = \bigcup_{i=1}^{d} (\mathbb{I} \times \ldots \times (u_i, 1) \times \ldots \times \mathbb{I}) \tag{4.48}$$

Hence, p_{OR} describes the probability to observe events in the hypersolid $\mathcal{S}_{\mathbf{u}}^{\vee}$ (Serinaldi, 2015) (bivariate example in Figure 4.3(a)).

- the *AND* exceedance probability:

$$\begin{aligned} p_{AND} &= P\{X_1 > x_1 \wedge \ldots \wedge X_d > x_d\} \\ &= \hat{C}\left(\overline{F}_1(x_1), \ldots, \overline{F}_d(x_d)\right) \\ &= \hat{C}\left(1 - u_1, \ldots, 1 - u_d\right) \\ &= \overline{C}\left(u_1, \ldots, u_d\right) \end{aligned} \tag{4.49}$$

that can also be written as:

$$p_{AND} = P\{(u_1, \ldots, u_d) \in \mathcal{S}_{\mathbf{u}}^{\wedge}\} \tag{4.50}$$

with

$$\mathcal{S}_{\mathbf{u}}^{\wedge} = \bigcap_{i=1}^{d} (\mathbb{I} \times \ldots \times (u_i, 1) \times \ldots \times \mathbb{I}) \tag{4.51}$$

In this case, p_{AND} is the probability that events in the hyperrectangle $\mathcal{S}_{\mathbf{u}}^{\wedge}$ occur (bivariate example in Figure 4.3(b)).

- the *Kendall* function-based exceedance probability:

$$\begin{aligned} p_K &= P\{P\{X_1 \leq x_1 \wedge \ldots \wedge X_d \leq x_d\} > t\} \\ &= P\{C\left(u_1, \ldots, u_d\right) > t\} \\ &= 1 - K_C\left(t\right) \end{aligned} \tag{4.52}$$

with $t \in \mathbb{I}$. It's possible to write p_K also as:

$$p_K = P\{(u_1, \ldots, u_d) \in \mathcal{S}_t^K\} \tag{4.53}$$

where

$$\mathcal{S}_t^K = \{(u_1, \ldots, u_d) \in \mathbb{I}^d \ : \ C\left(u_1, \ldots, u_d\right) > t\} \tag{4.54}$$

Hence, p_K indicates the probability to observe events belonging to the hyperspace $\mathcal{S}_t^K$ characterised by a copula level greater than t (bivariate example in Figure 4.3(c)).

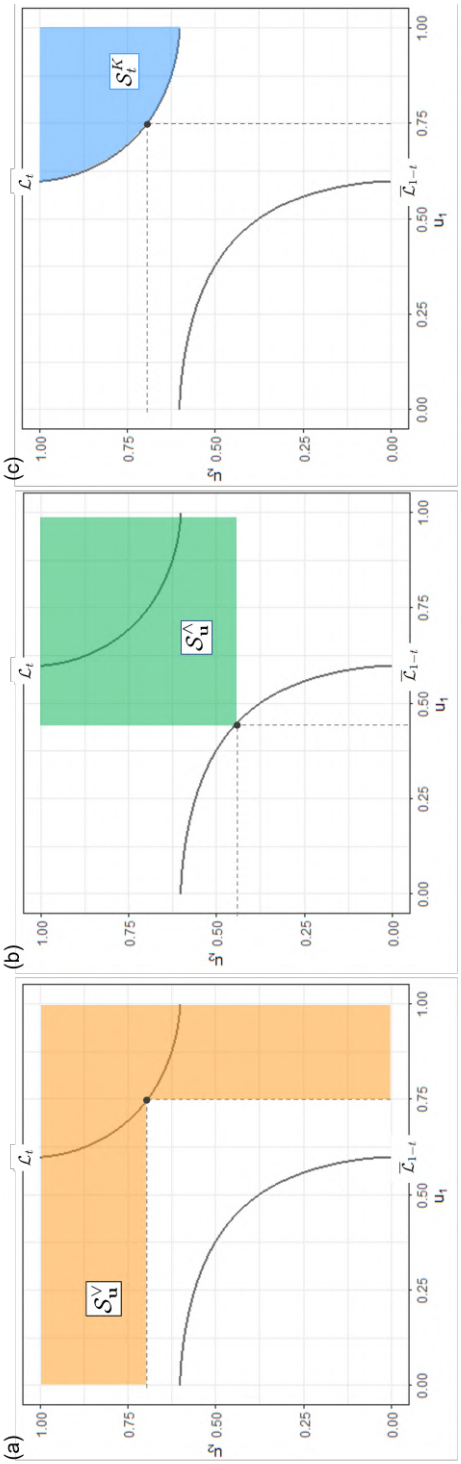

Figure 4.3: Bivariate example of exceedance probability types and related critical regions (colored area). (a) OR case, (b) AND case and (c) Kendall case. $\mathcal{L}_t$ and $\overline{\mathcal{L}}_{1-t}$ are the critical and survival critical layer of level t.

These three exceedance probabilities lead to as many related definitions of return period: T_{OR}, T_{AND} and T_K.

As evident from Eqs. (4.46) and (4.52), all copula points that are characterised by the same copula value have an equal return period, T_{OR} or T_K based on the considered probability type. Similarly, considering Eq. (4.49), the same T_{AND} is associated to all copula points with an equal survival function value. For this reason it is useful to introduce the concepts of *critical layer* and *survival critical layer* defined respectively as (e.g. Salvadori et al., 2011, 2013):

$$\mathcal{L}_t = \{(u_1, \ldots, u_d) \in \mathbf{I}^d \,:\, C\,(u_1, \ldots, u_d) = t\} \tag{4.55}$$

$$\overline{\mathcal{L}}_{1-t} = \{(u_1, \ldots, u_d) \in \mathbf{I}^d \,:\, \overline{C}\,(u_1, \ldots, u_d) = 1 - t\} \tag{4.56}$$

To avoid ambiguity, these definitions are written assuming conventionally that the level t represents a non-exceedance probability. Hence, $\mathcal{L}_t$ ($\overline{\mathcal{L}}_{1-t}$) is the $(d-1)$-dimensional isohypersurface along which C ($\overline{C}$) equals the constant value t $(1-t) \in \mathbb{I}$.

4.3 Study area and data

A severe storm, the so-called Vaia storm, affected northeastern Italy in October 2018, causing widespread and serious damage (e.g. Giovannini et al., 2021). The Sole Valley, located in the northwestern part of the Province of Trento (Italy), was one of the hardest affected areas, with the main river (i.e. the Noce river) experiencing a flood wave and a very intense debris flow occurring in one of its right tributaries (i.e. the Rotiano creek) (Figures 4.4 and 4.5), among others. This extreme event highlighted the possibility of compound phenomena occurring in this area and, for this reason, the confluence between the Noce river and the Rotiano creek, along with the related catchment, is considered as the study area for this work.

The Rotiano creek is an Alpine stream with a catchment of about 2.5 km^2, characterised by an altitude that varies between 2050 m a.s.l. and 800 m a.s.l. (at the alluvial fan) and is susceptible to debris flows (Piccolroaz et al., 2021). Most of the Rotian basin is characterised by high hydraulic conductivity and thick soils. The Noce catchment, with outlet point located just upstream of the confluence with the Rotiano creek, has an area of about 370 km^2 and its altitude ranges between 3750 m a.s.l. and 785 m a.s.l. with mean value equals to 2124 m a.s.l.. As a whole, the study area is characterised by an alpine climate and mean annual precipitation of approximately 1100 mm.

Rain gauge	Abbreviation	Altitude (m a.s.l.)	Record from	Mean annual precipitation (mm)
Pradalago	P	2084	1991	1400
Mezzana	M	905	1921	950
Passo Tonale	PT	1875	1987	1200
Careser	C	2600	1930	1000

Table 4.2: Rain gauges chosen for the analysis, along with their abbreviation, altitude, start year of recording and mean annual precipitation, estimated considering the available data from 1991. The location of the selected rain gauges is shown in Figure 4.5.

As for the precipitation, several rain gauges are present in the study area (Figure 4.5) and the rainfall data are available with a temporal resolution equal to 5 minutes. Among the available, the Pradalago rain gauge (P) is chosen as the reference rain gauge for the Rotiano catchment since it is the closest one with an altitude comparable with the one at the basin's highest part, where a debris flow could trigger. As for the Noce catchment, the considered rain gauges are Mezzana (M), Passo Tonale (PT) and Careser (C). These rain gauges are selected based on their location, altitude and data availability: as evident from Figure 4.5 and Table 4.2, these rain gauges are well dispersed in the Noce catchment, they are placed at different altitudes and their recording period includes the P one with an overall limited number of missing data. Three rain gauges are assumed to be a reasonable number resulting from the compromise between the desire to depict the spatial variability of rainfall in the Noce catchment and the degree of analysis complexity.

Since the aim is to perform a multivariate analysis, we exclude the calendar years with a missing data rate higher than 20% in at least one of the four selected rain gauges. Overall, 18 years of precipitation record, 13 of which continuous, are considered suitable for the analysis.

Moreover, the streamflow data in the Noce river provided by the Malé gauging station, located just downstream of the confluence (Figure 4.5), are considered for the event definition.

4.4 Method

The multivariate model with a multi-hazard analysis purpose is constructed and estimated with a copula-based approach.

The construction and estimate of the model are based on observations that are relevant to the phenomena under investigation in each selected

Figure 4.4: Effects of the Vaia storm on the Rotiano alluvial fan.

rain gauge, defined with a proper strategy. The model is then built in order to reproduce both the values and the dependence structure between the observations and it includes the following components:

- the marginal distributions, that are the univariate distributions describing the observations values in each rain gauge, independently;
- the dependence structure model, which consists of a multivariate distribution capable of reproducing the dependence structure between observations.

Once these components are defined and validated, the resulting model is used to estimate scenarios with different values and definitions of the return period, consisting of rainfall values in each rain gauge.

The analyzes described below have been done using R (R Core Team, 2013) and the packages: *copula* (Hofert et al., 2020), *VineCopula* (Schepsmeier et al., 2021), *CDVine* (Brechmann and Schepsmeier, 2013) and *rvinecopulib* (Nagler and Vatter, 2021).

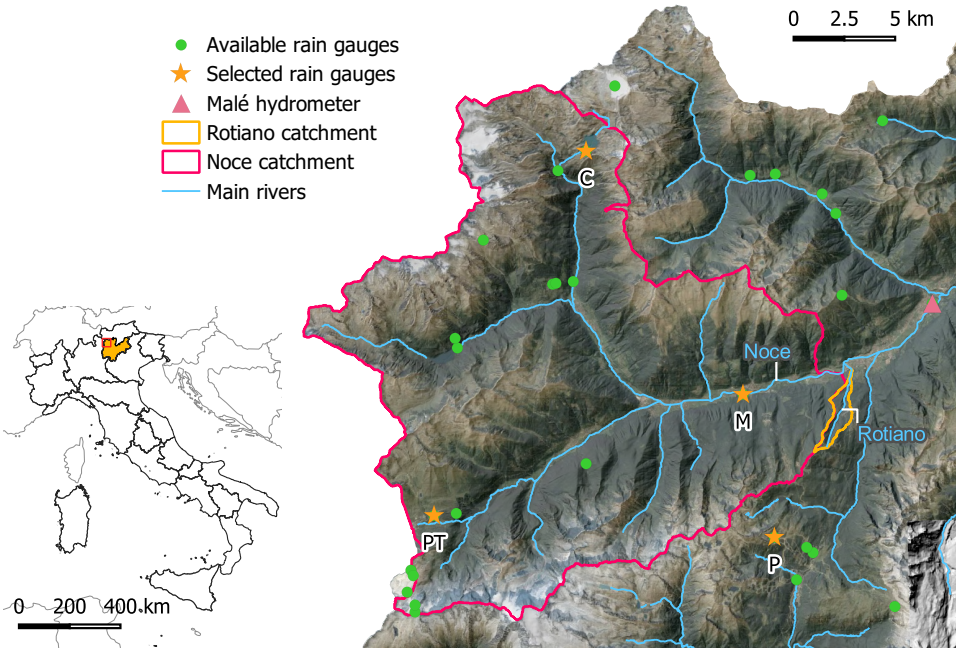

Figure 4.5: Location of the study area within the Province of Trento (Italy), with indication of the Noce and Rotiano catchments, the available rain gauges, the selected ones and the Malé hydrometer.

4.4.1 Event and related observations definition

The 5-minutes rainfall data series provided by the four selected rain gauges are aggregated in hourly time intervals thus obtaining $I_P(\tau)$, $I_M(\tau)$, $I_{PT}(\tau)$ and $I_C(\tau)$ in which the subscripts indicate the rain gauges, according to the abbreviations reported in Table 4.2, and $\tau = 1, \ldots, m$, with m total amount of hours in the rainfall records. $I_P(\tau)$, $I_M(\tau)$, $I_{PT}(\tau)$ and $I_C(\tau)$ are hence subdivided in independent rainfall events or "storms" considering dry separation intervals of at least 24 hours (Formetta et al., 2022, Gao et al., 2021, Miniussi and Marra, 2021). Moreover, according to Marra et al. (2020), only the independent rainfall events that last more than 30 minutes are considered to avoid to take into account isolated rain gauges spikes as storms.

Then, a strategy to define synchronous events, namely a rainfall events that occurred simultaneously in all selected rain gauges, and to obtain the related observations, is introduced. First, for each investigated hazardous phenomenon, a rainfall duration that is significant for its occurrence is identified. A rainfall duration of 24h is assumed to be representative for both debris flows and floods in the Rotiano and Noce catchments, respectively. Indeed, given the conductivity and thickness of most Rotiano basin soils, long precipitations are needed to reach a high degree of soil saturation and, therefore, the conditions for potential debris flows triggering (e.g. Hungr et al., 2002, Rosatti et al., 2019). This hydrologic response of the catchment is also highlighted by the debris flow that occurred during the Vaia storm. As for the Noce catchment, its extension and morphology make it reasonable to consider precipitation of 24h to observe a flood wave in the river at the confluence. For example and in support of these assumptions, the one hour rainfall intensities recorded by the four selected rain gauges and the discharges of the Noce river at the Malé gauging station during the Vaia storm are shown in Figure 4.6. Hence, the variables considered in this analysis are 24h cumulative rainfall and are named H_P, H_M, H_{PT} and H_C in which each subscript indicates the related rain gauge.

Second, the occurrence of a debris flow during the peak of a river flood is regarded as the most dangerous compound event. This timing is considered to identify synchronous events. It is assumed that if a storm in P results in a debris flow, the phenomenon triggering time coincides with the instant in which the hourly intensity in P is maximum (Iadanza et al., 2016, Rosatti et al., 2019). Moreover, the propagation time of the potential debris flow in the Rotian catchment is considered and it is assumed to be equal to 3 hours.

Al più ci mette tre ore per raggiungere la confluenza...intorno al picco

si genera

Hence, according to these assumptions, given a storm k in P, the hourly maximum rainfall intensity within the storm, $\hat{I}_P(k)$, and the related instant, $\hat{\tau}_P(k)$, are identified. Then, according to Callau Poduje and Haberlandt (2018), a time window is introduced to identify synchronous events: in this analysis, we define the time window as the interval $[\hat{\tau}_P(k) - 20,\ \hat{\tau}_P(k) + 3]$. The time window is then intersected with the independent events of M, PT and C and if the intersection is non-zero in all three rain gauges, the overall event is defined as synchronous.

Finally, the cumulative rainfall in 24h is computed for all synchronous events in each rain gauge. Given a synchronous event w, the cumulative rainfall for P is calculated as:

$$H_P^w = \sum_{j=-23}^{0} I_P\left(\hat{\tau}_P(w) + j\right) \tag{4.57}$$

and the time window $[\hat{\tau}^P(w) - 20,\ \hat{\tau}^P(w) + 3]$ is used to calculate the cumulative precipitation in the other three rain gauges:

$$H_M^w = \sum_{j=-20}^{3} I_M\left(\hat{\tau}_P(w) + j\right) \tag{4.58}$$

$$H_{PT}^w = \sum_{j=-20}^{3} I_{PT}\left(\hat{\tau}_P(w) + j\right) \tag{4.59}$$

$$H_C^w = \sum_{j=-20}^{3} I_C\left(\hat{\tau}_P(w) + j\right) \tag{4.60}$$

In this way, all elements of the synchronous events set $(H_P^w, H_M^w, H_{PT}^w, H_C^w)$, in which $w = 1, \ldots, N$ with N number of synchronous events, are defined.

It is worth noting that the definition of the synchronous events set made it necessary to introduce some (strong) assumptions in terms of rainfall duration and investigated phenomena timing. However, these assumptions can be seen as a first attempt. As will be highlighted later, the goodness of the provided assumptions can be evaluated and, if necessary, optimized based on the results of the analysis.

4.4.2 Marginal distributions

In this study, the analysis of extremes is conducted applying the SMEV approach, which, as mentioned in Section 4.2.4, builds the yearly non-exceedance probabilities distribution on the basis of all ordinary events of

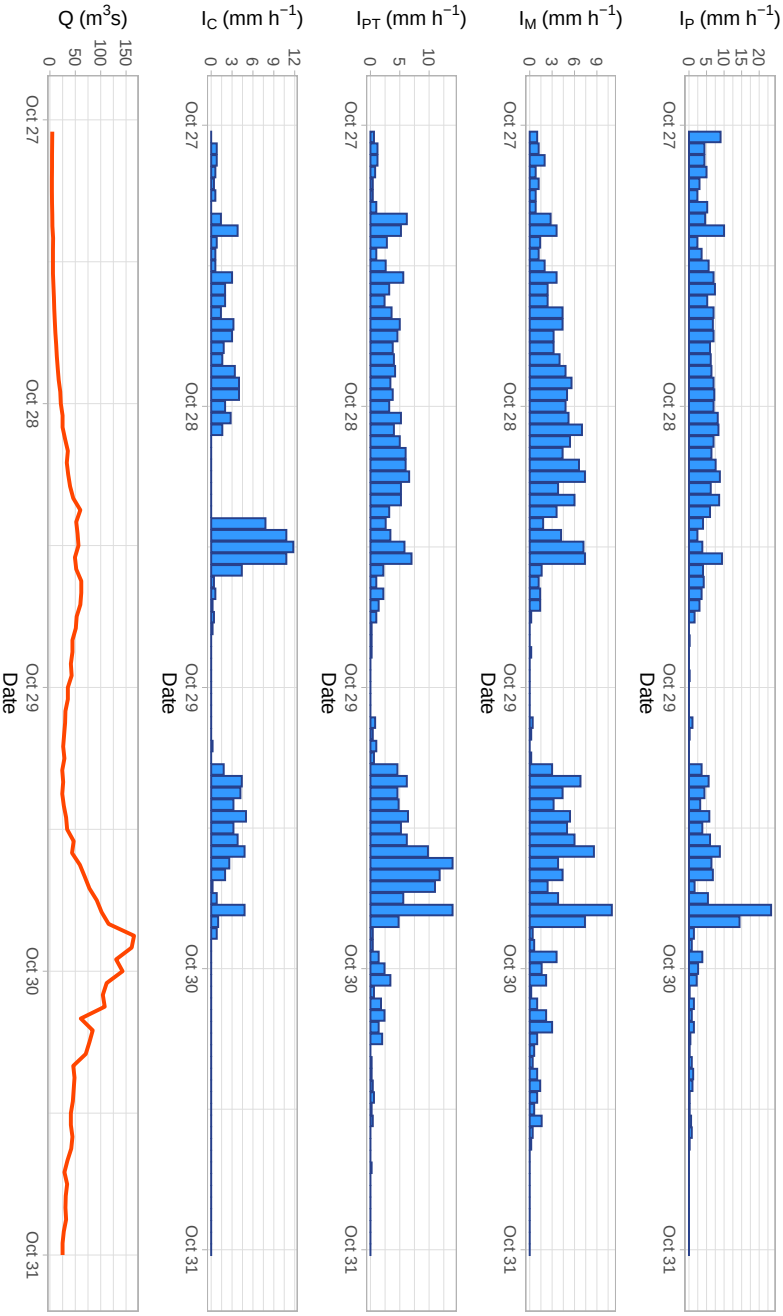

Figure 4.6: Hourly rainfall intensities at the selected rain gauges and hyetogram of the Noce river at the Malé gauging station recorded during the Vaia storm.

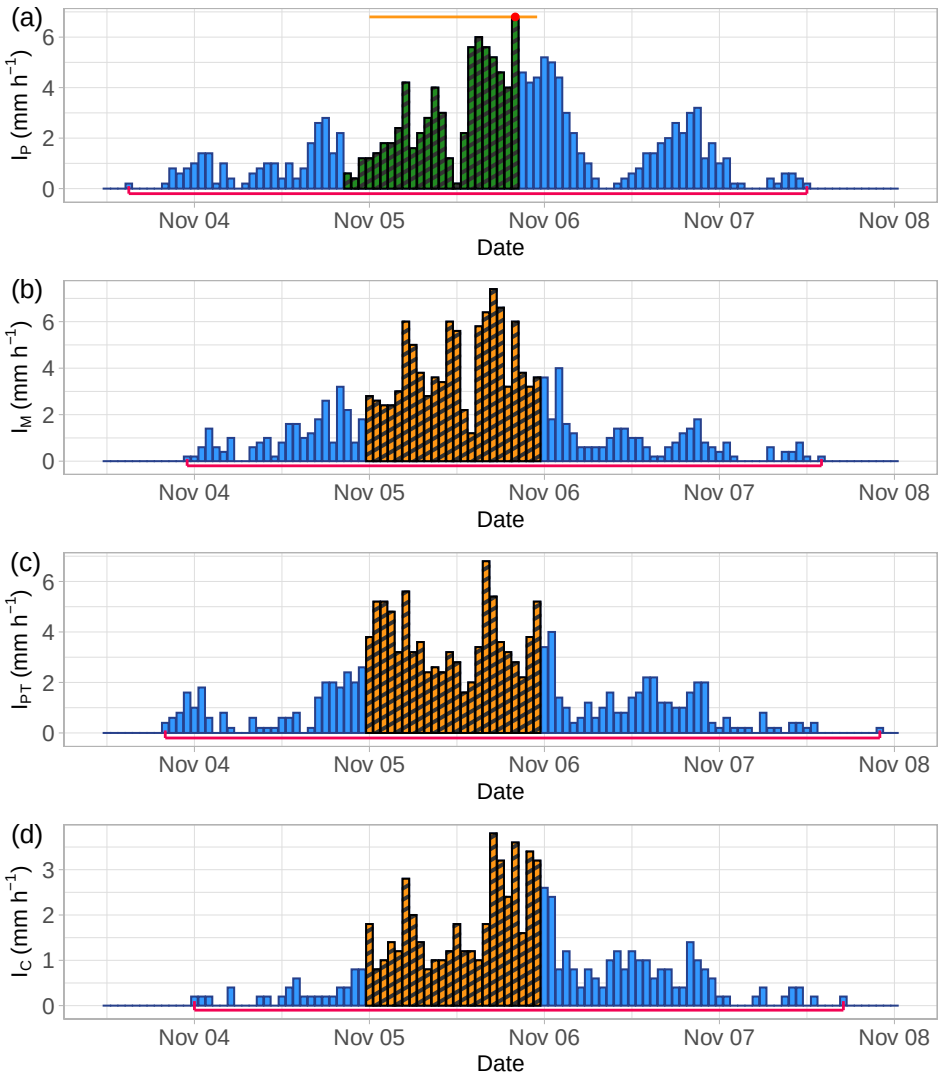

Figure 4.7: Computation example for one of the synchronous events set's elements. (a) Hourly rainfall intensities of a storm in P along with the identification of the maximum intensity instant (red dot) and the time window (orange segment). The green striped columns are the rainfall intensities considered to compute the 24h cumulative rainfall in P. (b), (c) and (b) hourly rainfall intensities of storms in M, PT and C respectively. The orange striped columns are the rainfall intensities within the time window and hence are the ones used to calculate the 24h cumulative rainfall in M, PT and C. For each rain gauge, the pink segment highlights the interval of the storm.

the investigated process. This approach is in contrast to classical EVT, that studies the yearly maxima distribution based on some selected observations of the process, namely directly annual maxima or realizations that exceed an established threshold (e.g. Davison and Smith, 1990). As a result, by avoiding this observations selection, the SMEV framework allows us to use all of the information offered by the realizations of the investigated process, for both univariate marginal and dependence structure analysis.

The SMEV approach is applied independently to each considered rain gauge to obtain the related marginal distributions considering, as ordinary events, the observations of the related variable. Given a rain gauge, the parameters of the Weibull distribution are estimated left-censoring the related ordinary events, considering as threshold the 75th percentile (Formetta et al., 2022, Marra et al., 2019), and adopting the maximum likelihood estimation (MLE) technique (e.g. Balakrishnan and Kateri, 2008, Delignette-Muller and Dutang, 2015). The result is four fitted Weibull CDF: $G_P(x_P)$, $G_M(x_M)$, $G_{PT}(x_{PT})$ and $G_C(x_C)$, and related extremes CDF: $\Omega_P(x_P)$, $\Omega_M(x_M)$, $\Omega_{PT}(x_{PT})$ and $\Omega_C(x_C)$.

Then, for each variable, to check the adequacy of the Weibull distribution in representing the right tail of the distribution of the ordinary events, the 90% sampling uncertainty is quantified generating 1000 random samples, of size equal to the ordinary events one, from the resulting fitted distribution (Miniussi and Marra, 2021). The assumption of having a Weibull tail above the threshold, used to censure the observations, is to be rejected if more than 10% of the left-censored observations lie outside the 90% sampling confidence interval.

Finally, the uncertainty in the quantiles (obtainable by inverting Eq. (4.43)), due to the uncertainty in the estimate of the distribution's parameters, is quantified using the bootstrap method with replacement. According to Marra et al. (2019) and Formetta et al. (2022), the procedure consists in (i) randomly select, with replacement, a number of years equal to the ones available in the record, (ii) consider as ordinary events the ones observed during the sampled years and (iii), based on these, estimate the distribution parameters. This procedure is performed 1000 times and the 90% confidence interval is used to quantify the quantiles uncertainty.

4.4.3 Dependence structure estimate, modeling and validation

Once the marginal distributions are defined, the dependence structure between the observations of the synchronous events set is assessed and modelled through a copula-based approach.

The dependence structure of the observations is characterized by esti-

mating τ_n and TCF_n of all possible variables pairs (Eqs. 4.22 and 4.31, respectively). To compute these dependence measures, all observations are transformed in pseudo-observations performing a rank order transformation as (Graler et al., 2013):

$$S(x) = \frac{rank(x)}{N+1} \tag{4.61}$$

in which N is the total number of observations in each rain gauge. This transformation allow us to transform the variables H_P, H_M, H_{PT} and H_C in the variables U, V, W and Z, respectively, each approximately uniformly distributed on $\mathbb{I}$. Hence, the variables pairs (U, V), (U, W), (U, Z), (V, W), (V, Z) and (W, Z) are considered.

Hence, the empirical dependence structure has to be reproduced introducing a proper multivariate copula. In general, constructing high dimensional copulas is widely acknowledged to be a difficult problem (Aas et al., 2009). Indeed, unlike bivariate copulas, the number of higher-dimensional copulas is very limited in literature. For this reason, the PCC and the related vine graphical representation are used in this analysis to model the dependence structure between the four variables under investigation.

According to Dißmann et al. (2013), to fit a four-dimensional R-vine copula to the pseudo-observations of the variables U, V, W and Z, the following steps are needed for each tree:

1. select the tree structure;

2. define a bivariate copula family for each edge present in the selected tree structure. According to Yu et al. (2020), in this analysis, the families of copulas considered from which to choose are the elliptical and archimedean copulas (Table 4.1). It is worth noting that, for all trees, before selecting each (conditional) bivariate copula family, the test of independence proposed in Genest and Favre (2007) is performed to assess if the independence copula has to be chosen;

3. estimate the parameters of each chosen copula.

These three steps are performed following the procedure proposed in Dißmann et al. (2013) and implemented in the R package "VineCopula" (Schepsmeier et al., 2015). Starting from the first tree T_1, empirical Kendall's tau of all possible variables pairs are considered. Then, the T_1 structure that maximizes the sum of the absolute empirical Kendall's tau is selected. To this end, a maximum spanning tree algorithm (e.g. Cormen et al., 2009) is implemented, namely the following general optimization problem is solved:

$$\max \sum_{e=\{ij\} \text{ in spanning tree}} |\tau_{n,ij}| \tag{4.62}$$

Then, given the T_1 structure, the bivariate copula families corresponding to each edge is selected using the Akaike Information Criterion (AIC) (Akaike, 1973), namely for each pair, the one with the smallest AIC is selected and the related parameter(s) is estimated with the MLE technique. Then, the three marginal conditional CDF to be used as nodes in T_2 are estimated using Eq. (4.36) and empirical Kendall's tau is computed for each conditional variables pair that can be present in T_2. As for T_1, the T_2 structure is, then, selected maximizing the sum of the absolute empirical Kendall's tau just calculated, using the optimisation of Eq. (4.62). Hence, the AIC values are used to select the conditional bivariate copula families of the obtained T_2 structure and the related parameters are estimated with the MLE technique. Finally, Eq. (4.36) is used to compute the two marginal conditional CDF to be used as nodes in T_3, the conditional bivariate copula is selected with the AIC method and the copula parameter(s) is estimated with the MLE technique.

In this way, each tree of the R-vine structure, as well as the six (conditional) bivariate copula families and their associated parameters, are defined.

Once the pair-copula decomposition is estimated, a strategy to validate the obtained dependence structure model is introduced. The aim is to evaluate how adequate the model obtained is in reproducing the dependence structure between the observations. According to Brunner et al. (2019), quantitative tests are not available to validate a multivariate distribution. As a result, in this analysis, the validation is based on the following copula-related quantities:

- Kendall function and related λ-function;
- TCF;
- Kendall's tau.

and, in particular, on the comparison between their empirical values and the theoretical ones provided by the copula. The adequacy of each (conditional) bivariate copula, present in the resulting R-vine structure, and the overall four-dimensional R-vine copula is assessed.

To validate an arbitrary copula, the empirical values of each quantity are estimated based on observations of the variables of which the copula models the dependence structure, using Eqs. (4.14), (4.16), (4.22) and (4.31). The

theoretical formulas are then employed to compute the related theoretical values of of each quantity, namely considering Eqs. (4.10), (4.15), (4.23) and (4.30). If a theoretical formula of a quantity is not available in closed-form for the considered copula, its theoretical values are estimated by simulation based on sample of size 10^4 generated from the investigated copula. This method is applied for Kendall functions and λ-functions of elliptical copulas and four-dimensional R-vine copula and Kendall's tau between variables pairs modelled by the four-dimensional R-vine copula.

As regards the TCF of the R-vine copula, this quantity is not considered in this analysis given its complexity and the related large number of possible conditional probabilities definitions (i.e. 14 for a four-dimensional copula) (for more details see e.g. Luca and Rivieccio, 2012). However, the bivariate version of the TCF, related to the variables pairs not directly modelled by the R-vine copula and hence not addressed in the validation of the (conditional) bivariate copulas, is considered to validate the overall dependence structure model.

Once the theoretical and empirical values of the considered copula-related quantities are computed for all copulas, the adequacy of the obtained dependence structure model is assessed graphically and by introducing proper performance measures. For all measures but Kendall's tau, the adequacy is quantified through the root mean square error normalized with respect to the empirical mean (NRMSE), that, in a generic form, can be written as:

$$NRMSE = \frac{\sqrt{\frac{1}{s}\sum_{i=1}^{s}(h^i - \tilde{h^i})^2}}{\frac{1}{s}\sum_{i=1}^{s} h^i} \tag{4.63}$$

where h^i is the ith empirical values of the vector $\mathbf{h}$ of dimension s and $\tilde{h^i}$ is the related theoretical value predicted by the considered model. Instead, the theoretical and empirical Kendall's tau are compared by calculating their absolute difference normalized with respect to the empirical value (NAD) since they are a single value for each variables-pair. The generic expression of NAD is the following:

$$NAD = \frac{|h - \tilde{h}|}{h} \tag{4.64}$$

with h and $\tilde{h}$ empirical and theoretical value, respectively. It is worth noting that Eq. (4.64) can be obtained from Eq. (4.63) by imposing $s = 1$. Both the considered performance measures are calculated in a normalized form to allow us to compare their values.

Finally, to assess the adequacy of each copula in simulating the dependence structure of the considered variables, the related 90% sampling uncertainty interval of each considered quantity is estimated performing 1000 repetitions of their empirical version computation using an observations-size random sample generated from the considered copula.

4.4.4 Scenarios

Once the marginal distributions and the dependence structure model are defined and validated, it is possible to use the resulting multivariate model to estimate scenarios with assigned return period, consisting of a non-exceedance probability, namely a 24h cumulative rainfall value, for each investigated variable.

In this analysis, to compute a scenario (or design event) for a given value and definition of the return period T, the following steps are performed:

1. identify the non-exceedance probability level t related to T. Considering the OR case, t can be computed combining Eqs. (4.45), (4.46) and (4.55):

$$t_{OR} = \left[1 - \frac{1}{T_{OR}}\right]^{1/n} \tag{4.65}$$

Similarly, for the AND case, considering Eqs. (4.45), (4.49) and (4.56), t can be written as:

$$t_{AND} = \left[1 - \frac{1}{T_{AND}}\right]^{1/n} \tag{4.66}$$

As for the Kendall case, combining Eqs. (4.45), (4.52) and (4.55), t is defined as:

$$t_K = K_C^{-1}\left(\left(1 - \frac{1}{T_K}\right)^{1/n}\right) \tag{4.67}$$

in which $K_C^{-1}(\cdot)$ is the inverse of K_C that allow us to compute the *Kendall's quantile* generically defined as:

$$\begin{aligned} q &= \inf\{l \in \mathbb{I} \ : \ K_C(l) = b\} \\ &= K_C^{-1}(b) \end{aligned} \tag{4.68}$$

with $b \in \mathbb{I}$ (Salvadori et al., 2011). Since a closed-from for K_C is not available, t_K is estimated using a bootstrap technique, namely algorithm 2 proposed in Salvadori et al. (2011), considering 1000 repetitions with sample size 10^5;

2. identify the isohypersurface related to t ($\mathcal{L}_t$ for the OR and Kendall case while $\overline{\mathcal{L}}_{1-t}$ for the AND one). Considering the analyzed variables U, V, W and Z with R-vine copula C and density c, for the OR and Kendall case, the four-dimensional points $(u, v, w, z) \in \mathbb{I}^4$ that satisfy respectively the following equations have to be identified:

$$\int_0^u \int_0^v \int_0^w \int_0^z c(u, v, w, z)\, dudvdwdz = t_{OR} \tag{4.69}$$

$$\int_0^u \int_0^v \int_0^w \int_0^z c(u, v, w, z)\, dudvdwdz = t_K \tag{4.70}$$

Instead, for the AND case, considering Eqs. (4.20) and (4.56), the equation to be solved is the following (Durante et al., 2008, Lux and Papapantoleon, 2017):

$$\begin{aligned} &1 - u - v - w - z + C(u, v, 1, 1) + C(u, 1, w, 1) + C(u, 1, 1, z) + \\ &C(1, v, w, 1) + C(1, v, 1, z) + C(1, 1, w, z) - C(u, v, w, 1) - \\ &C(u, v, 1, z) - C(u, 1, w, z) - C(1, v, w, z) + C(u, v, w, z) \\ &= 1 - t_{AND} \end{aligned} \tag{4.71}$$

in which each copula term is computed with a four-dimensional integral with integration from 0 to 1 for the variable equals to 1. For instance, the term $C(1, v, 1, z)$ is:

$$C(1, v, 1, z) = \int_0^1 \int_0^v \int_0^1 \int_0^z c(u, v, w, z)\, dudvdwdz \tag{4.72}$$

The isohypersurfaces of level t_{OR} and t_K are computed considering 15000 points for which the values of three variables are uniformly generated between $(0, 1)$ and the fourth variable value is estimated in order to satisfy Eq. (4.69) and Eq. (4.70), respectively. The four-dimensional integral is estimated with a Quasi-Monte Carlo approach considering the Halton sequence to generate 10^6 sample points for the integration (e.g. Cambou et al., 2017, Nagler and Vatter, 2021, Owen, 2003). The same approach is used to find the points of the isohypersurface of level t_{AND}. However, for the AND case, the points considered are 5000 due to the high computational times to solve Eq. (4.71);

3. select the scenario related to T, namely one copula point of the estimated isohypersurface, introducing a proper approach to identify

it. The approach considered in this analysis is the *most-likely design realization* (MLDR) one, introduced by Salvadori et al. (2011). According to this approach, among all possible copula points that belong to the isohypersurface related to t, the one with the highest joint PDF is selected. Considering the analyzed variables $U = G_P(H_P)$, $V = G_M(H_M)$, $W = G_{PT}(H_{PT})$ and $Z = G_C(H_C)$ with copula C and joint PDF f, the MLDR for the OR, AND and Kendall case are defined respectively as:

$$(u, v, w, z)_{OR} = \underset{C(u,v,w,z)=t_{OR}}{\arg\max} f\left(G_P^{-1}(u), G_M^{-1}(v), G_{PT}^{-1}(w), G_C^{-1}(z)\right) \tag{4.73}$$

$$(u, v, w, z)_{K} = \underset{C(u,v,w,z)=t_{K}}{\arg\max} f\left(G_P^{-1}(u), G_M^{-1}(v), G_{PT}^{-1}(w), G_C^{-1}(z)\right) \tag{4.74}$$

$$(u, v, w, z)_{AND} = \underset{\overline{C}(u,v,w,z)=1-t_{AND}}{\arg\max} f\left(G_P^{-1}(u), G_M^{-1}(v), G_{PT}^{-1}(w), G_C^{-1}(z)\right) \tag{4.75}$$

Moreover, considering a level t and a definition of T, the effect on the MLDR selection of the points used to define the related isohypersurface is assessed by applying a bootstrap method with replacement. Given an isohypersurface, M points are randomly drawn from the M available ones and, among these, the MLDR is identified. The replication of this procedure 1000 times allows us to quantify the MLDR uncertainty through the 90% confidence interval;

4. finally, the copula points are transformed to real scale. Given a copula point (u, v, w, z), the corresponding quantiles of the variables H_P, H_M, H_{PT} and H_C, namely the point $(x_P, x_M, x_{PT}, x_C) \in \mathbb{R}^4$, are computed as:

$$(x_P, x_M, x_{PT}, x_C) = \left(G_P^{-1}(u), G_M^{-1}(v), G_{PT}^{-1}(w), G_C^{-1}(z)\right) \tag{4.76}$$

In this analysis, for comparative purposes, $T_{OR} = T_K = T_{AND} = 30$ years are considered. Moreover, $T_K = 100$ and 200 years are assessed to obtain the typical scenarios considered in hazard mapping for the Kendall case.

It is worth noting that, given its definition, p_{OR} is not actually adequate to describe the event object of this analysis, i.e. the co-occurrence of the investigated hazardous phenomena at the confluence. However, a scenario related to T_{OR} is considered for comparative purposes.

Finally, the uncertainty in the MLDR values estimate due to uncertainty in the dependence structure estimate and modeling is not taken into account in this research. Indeed, the high computational times required to define an isohypersurface make it impossible to assess the effect of this type of uncertainty on the MLDR values, using for example the method provided in Dung et al. (2015), in an acceptable amount of time.

4.5 Results and discussion

4.5.1 Observations and empirical dependence measures

The synchronous events analysis results in $N = 557$ observations in each rain gauge. The monthly and yearly distribution of observations, as well as the variability of the number of observations per year, are depicted in Figure 4.8. Coherently with the rainfall regime of the investigated area, the months with the greatest number of synchronous events are July, August, May and October (Figure 4.8(a)). As regards the yearly result, 2009 had the fewest number of synchronous events (i.e. 22), whilst 2013 had the most (i.e. 41) (Figure 4.8(b)). However, it is worth noting that the 2009 rainfall record is characterized by missing data at the C rain gauge and this contributes to have a fewer number of events. Considering hence the distribution of the annual number of observations, the summary statistics are shown in Figure 4.8(c) and the mean value is $n = 31$.

As regards the values of the observations, some statistics of each variable of the synchronous events set are reported in Table 4.3. H_P is the variable characterized by the largest mean and high skewness and kurtosis, that suggest a right-heavy tail. This is due to both event definition, which tends to maximize H_P values, and rainfall at P, which is on average higher than in other rain gauges (Table 4.2). H_C has the lowest mean and the skewness and kurtosis values describe a less right-heavy tail than other variables. A mean value intermediate between the H_P and H_C ones characterize the observations of the variables H_M and H_{PT}, with the latter very close to the H_P one. Both H_M and H_{PT} observations present a right-heavy tail but it is less marked than the H_P observations one. These considerations on the distributions of the observations values are visually evident from the box plots shown in Figure 4.9.

The observations values may be explained considering both the rainfall regime at each rain gauge (Table 4.2) and the distance from P (Figure 4.5). By comparing the timing of storms in P and the related storms in the other rain gauges, it is possible to assess if the distance from P may effect the contemporaneity of storms' most intense part and, hence, the consequent

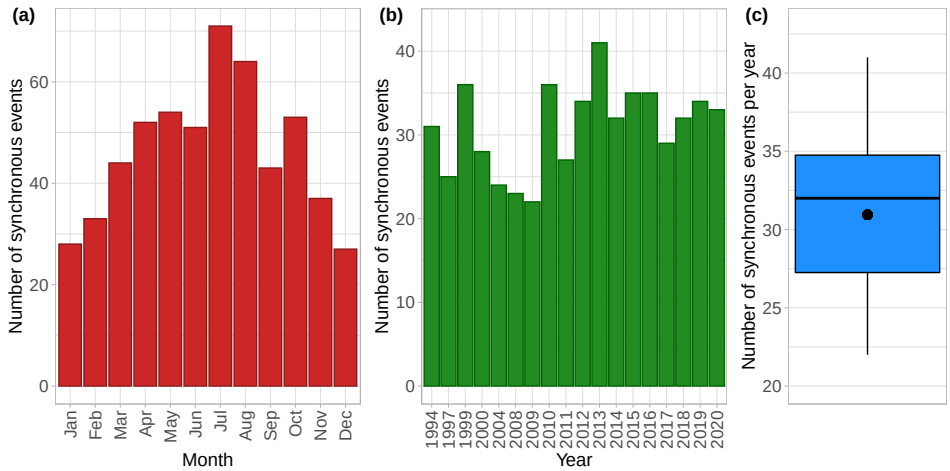

Figure 4.8: Number of synchronous events: (a) monthly and (b) annual distribution and (c) boxplot of the number of observations per year. The black dot indicates the mean value n.

	H_P	H_M	H_{PT}	H_C
25th percentile (mm)	5.2	3.4	5.2	3.2
Mean (mm)	16.8	13.4	16	10.9
75th percentile (mm)	21.4	17.6	21.4	15.4
Maximum (mm)	173.2	98.8	112.6	66.4
Skewness	3.2	1.9	2	1.5
Kurtosis	15.6	4.6	5.1	2.3

Table 4.3: Some statistical quantities of the observations of the variables H_P, H_M, H_{PT} and H_C, respectively.

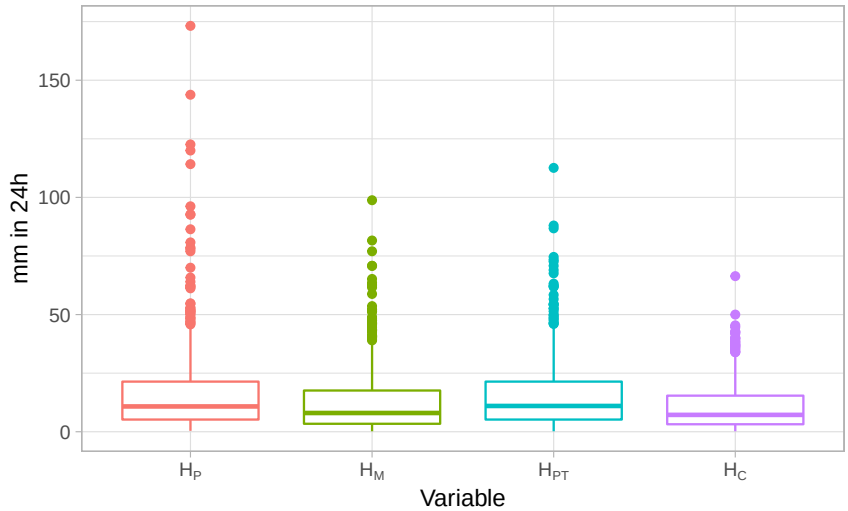

Figure 4.9: Box plots of the observations values of the variables H_P, H_M, H_{PT} and H_C.

magnitude of the cumulative rainfall. To assess this aspect, the hourly maximum rainfall intensity instant, namely the rainfall peak instant, of a given rainfall event is considered as the representative instant of the storm's most intense part. Then, for all rain gauges and storms, it is evaluated if the rainfall peak instant occurs within the time window used to calculate the related 24h cumulative rainfall. Finally, for each rain gauge, the percentage of synchronous events with the rainfall peak instant within the related time window is calculated and shown in Table 4.4. As evident, the greater the rain gauge is distant from P, the less frequent is the contemporaneity of the most intense part of its storms and the P ones. Hence, combining this information with the rainfall at the rain gauges, it is possible to make the following considerations:

- H_M refers to the rain gauge with the lowest mean annual precipitation but its great closeness to P may justify the right-heavy tail that characterizes its observations values: in comparison to the other rain gauges, the most intense part of the storms in M most frequently occur close in time at the related rainfall peaks in P, namely within the connected time windows;

- PT has the highest mean annual precipitation and a distance from P that is larger than that between P and M but less than that between P and C. These characteristics may justify the high values of H_{PT}

Rain gauge	Rainfall peaks instant within the related time window
M	69%
PT	64%
C	63%

Table 4.4: For each rain gauge, the percentage of synchronous events whose hourly maximum rainfall intensity (i.e. rainfall peak) instant is included in the related time window used to calculate the 24h cumulative rainfall.

observations that, however, are characterized by right tail comparable to the H_M one. More frequently than M, in fact, the most intense part of the storms in PT does not fall within the relative time window. Hence, considering the pluviometric regime in PT, this results in extreme values that are lower than expected.

- H_C refers to the rain gauge with mean annual precipitation very close to the M one but the greatest distance from P. These aspects may explain why the observations of H_C have the lowest values and are characterized by a less right-heavy tail than other variables: the lowest frequency of contemporary occurrence of the most intense part of the storms in C and the related rainfall peaks in P results in relatively lower extremes values.

Regarding the observations correlation structure, its graphical overview is provided by the scatter plots reported in Figure 4.10. It's possible to visually notice a positive correlation between the observations of all variables pairs and it seems to be strongest and weakest between the observations of the variables H_M and H_{PT} (Figure 4.10(d)) and H_P and H_C (Figure 4.10(c)), respectively, both in overall terms and in the upper tail.

These visually-based considerations are confirmed by the Kendall's tau and TCF_n computed considering the pseudo-observations and shown in Figure 4.11 and 4.12, respectively. The empirical Kendall's tau values range between 0.46 and 0.57 with the minimum correlation between the pseudo-observations of U and Z and the maximum one between the pseudo-observations of V and W. Hence, all variables pairs present a strong correlation.

Regarding the empirical upper tail dependence, the strongest one subsists between the pseudo-observations of the variables V and W whose TCF_n, for instance, is about 0.61 with $t = 0.95$. The pseudo-observations of the variables U and Z present the weakest upper tail dependence with the

TCF_n that tends to zero fast with a value equal to 0.26 for $t = 0.95$. The remaining pairs are characterized by a TCF_n that convergences to zero slower than the (U, Z) one: considering $t = 0.95$, the TCF_n of the pairs (U, W) and (W, Z) is equal to about 0.44 while the TCF_n of the pairs (U, V) and (V, Z) is about 0.4. It is worth noting that the spikes present in the TCF_n of the variables pairs (V, W), (V, Z) and (W, Z) for values of t very close to 1^- are due to both the pseudo-observations behavior in the upper-right corner of the unit square and their rather limited number used to compute these values of TCF_n.

The obtained empirical dependence structure may be explained considering the location of the investigated rain gauges and their mutual distances (see Figure 4.5). Considering the empirical dependencies between U and the other three variables, Kendall's tau (first row of Figure 4.11) slightly decrease by increasing the mutual distance between P and the second rain gauge considered (i.e. M, PT or C). This possible effect of the distance between P and the other rain gauges is less evident considering the empirical upper tail dependence. It can only be noted that the weakest upper tail dependence subsists between the observations in P and C. Moreover, the location of M and PT within the same valley may explain the strongest correlation and upper tail dependence between the observations of the variables H_M and H_{PT}. However, more rain gauges would be required to fully assess how their mutual distances and positions effect the correlation structure.

4.5.2 Fitted marginal distributions

The univariate Weibull distributions fitting the left-censored observations of each variable are shown in Figure 4.13 along with the related 90% sampling uncertainty interval. The graphs' coordinates make the Weibull distributions straight lines. To better visualize the adequacy of the Weibull distribution in representing the upper portion of the observations, the fitted distributions and sampling uncertainty intervals are also shown in Q-Q plots in Figure 4.14.

As evident from Figure 4.13, for all variables, the observations greater than the 75th percentile are characterized by an approximately linear behavior satisfactorily represented by the related fitting Weibull distributions. The linear behavior is most pronounced in the upper portion of the observations of H_P, H_{PT} and H_M (Figure 4.13(a), (b) and (c)) while the largest values of H_C tend to deviate slightly from the related fitted distribution (Figure 4.13(d)). Ordinary events below the 75th percentile tend to deviate from the obtained distributions for all variables.

Moreover, as evident from Figure 4.14(a), the Weibull distribution de-

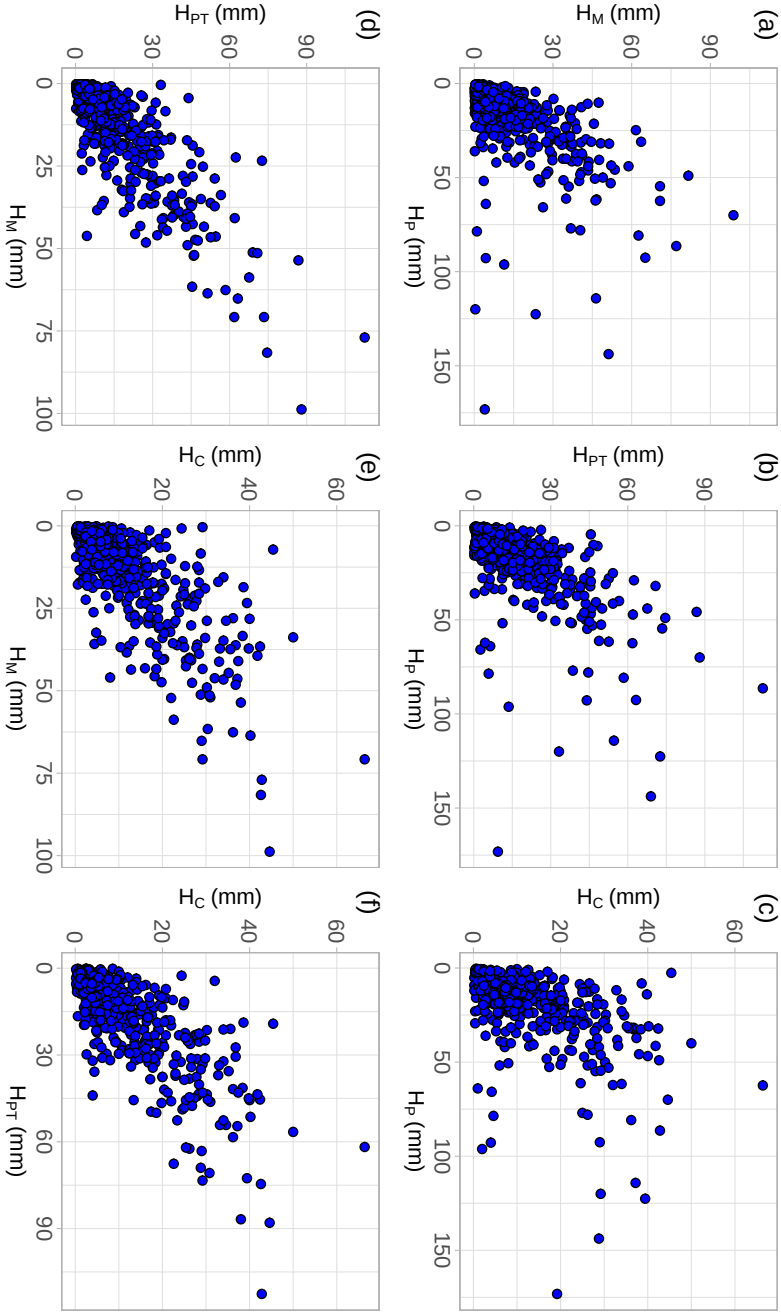

Figure 4.10: Observations scatter plots of all possible variables pairs.

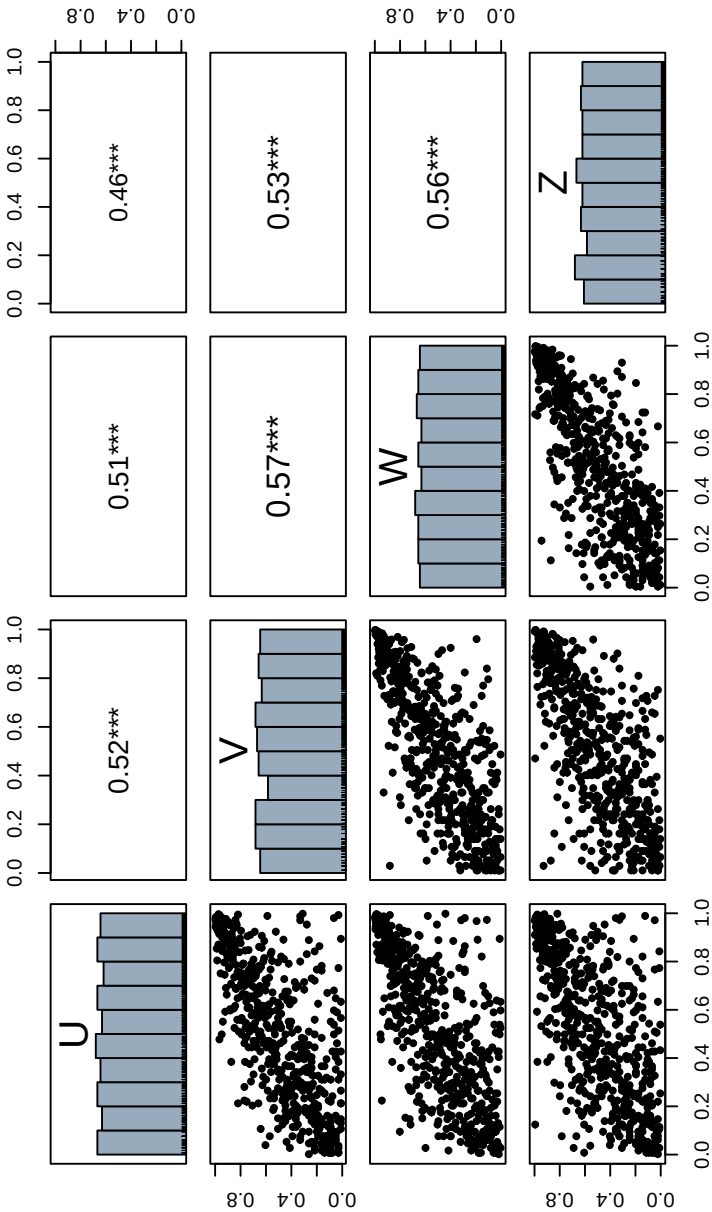

Figure 4.11: Empirical Kendall's tau values. Lower panel: pseudo-observations scatter plots of all possible variables pairs. Diagonal: histograms of pseudo-observations. Upper panel: empirical Kendall's tau of all possible variables pairs, statistically significant at 1% level.

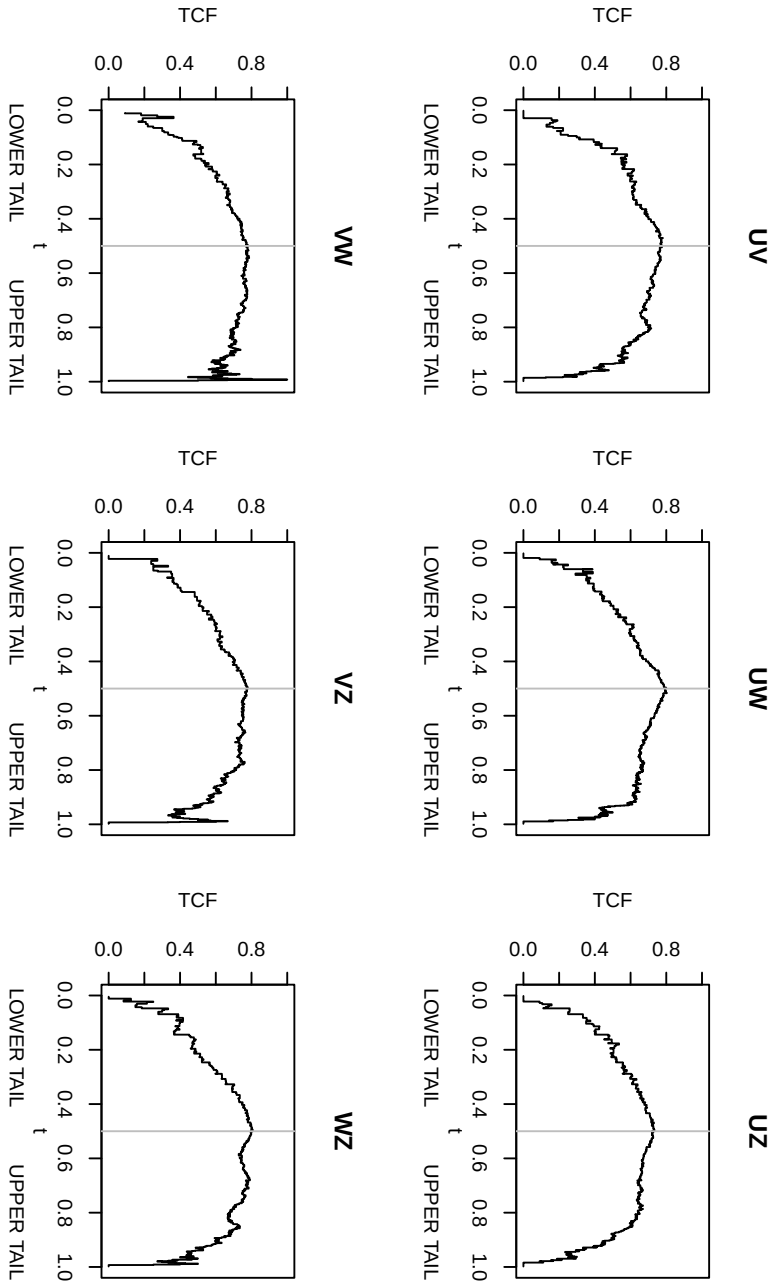

Figure 4.12: Empirical TCF of pseudo-observations for all possible variables pairs. The title of each graph indicates the variables pair considered.

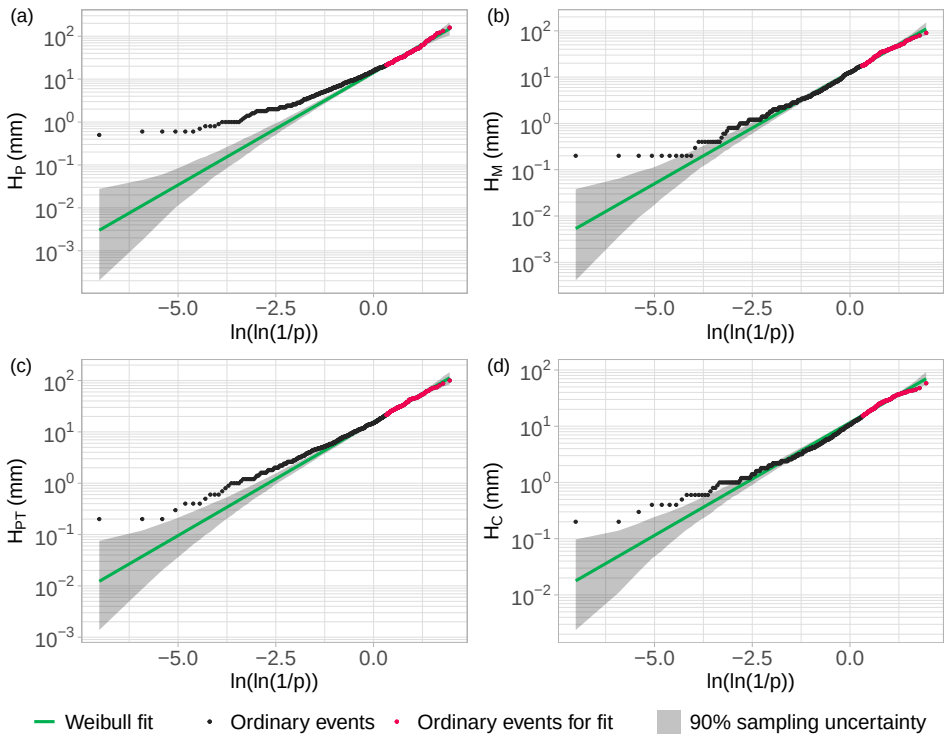

Figure 4.13: Resulting Weibull distributions reproducing the upper portion of the observations of each variable: (a) H_P, (b) H_M, (c) H_{PT} and (d) H_C. For each plot, the green solid line is the Weibull distribution fitting the left-censored ordinary events, represented with magenta points, the black points are all the other ordinary events and the gray shaded area is the 90% sampling uncertainty interval. p is the exceedance probability provided by the Weibull distribution (i.e. $1 - G(x)$).

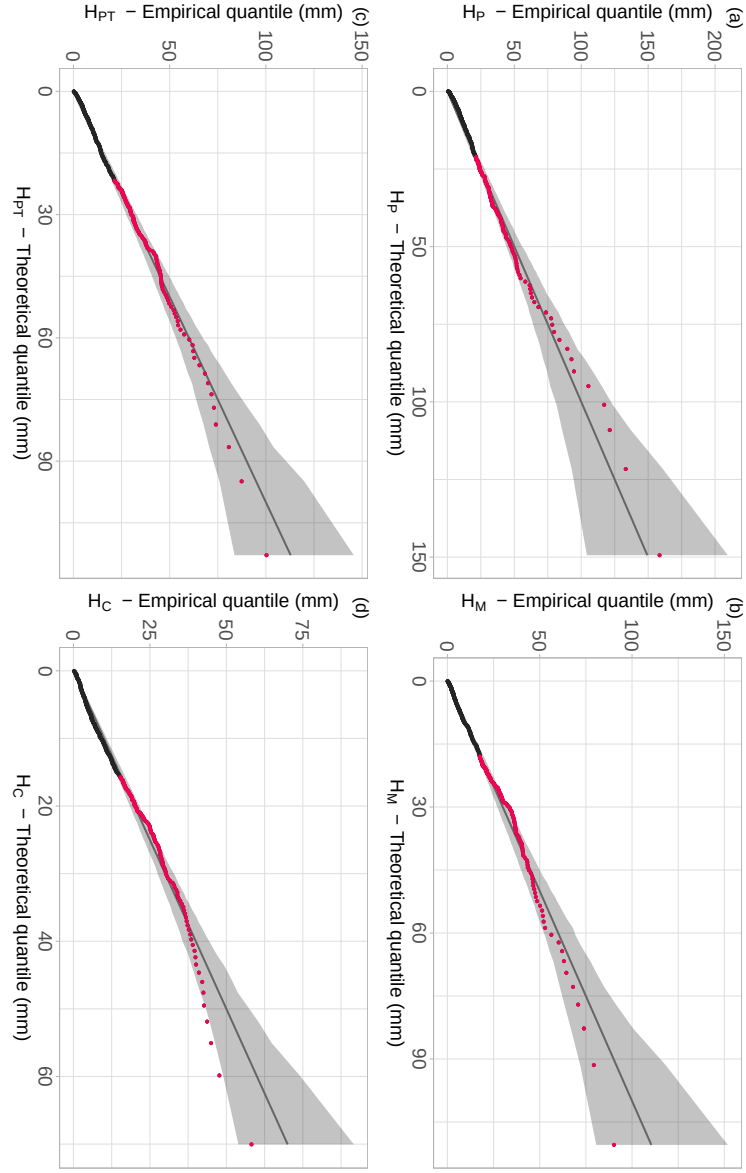

Figure 4.14: Q-Q plot for each variable: (a) H_P, (b) H_M, (c) H_{PT} and (d) H_C. In each graph, the theoretical quantiles computed with the fitted Weibull distribution are compared with the related empirical ones. The magenta points represent the observations used for the fit and are expected to follow the bisector, the black ones are the other ordinary events and the gray shaded area is the 90% sampling uncertainty interval.