Raül Oorthuis⬤ · **Marcel Hürlimann**⬤ · **Jean Vaunat**⬤ · **José Moya**⬤ · **Antonio Lloret**⬤

Monitoring the role of soil hydrologic conditions and rainfall for the triggering of torrential flows in the Rebaixader catchment (Central Pyrenees, Spain)

Role of soil hydrologic conditions and rainfall for torrential flows in the Rebaixader catchment

Abstract Torrential flows (debris flows and debris floods) are mainly triggered by precipitation and soil hydrological processes. Most early warning systems in torrential catchments are rainfall-based. However, this approach can result in frequent false positives, due to its pure black-box nature, in which soil water conditions are neglected. We aim to contribute to the understanding of the conditions required for triggering torrential flows by considering also in situ measurements of soil water content. Herein, monitoring data of 12 years of rainfall and torrential flow occurrence (2009–2020) and 8 years of soil hydrologic conditions (2013–2020) in the Rebaixader catchment (Central Pyrenees, Spain) are analyzed. The dataset includes more than 1000 rainfall events and 37 torrential flows. First, rainfall thresholds using maximum rainfall intensity (I_{max}) and mean intensity (I_{mean}) are defined. For the 2013–2020 dataset, which includes 15 torrential events, the I_{mean} threshold predicted 2 false negatives and 73 false positives (positive predictive value, PPV, of 15.1%) and the best I_{max} threshold predicted also 2 false negatives but only 11 false positives (PPV of 54.2%). However, our observations confirmed quantitatively that the lower is the soil moisture the higher is the rainfall intensity to trigger torrential flows. Then, we combined I_{max} and volumetric water content at 15 and 30 cm depth to define a hydro-meteorological threshold. This latter threshold reduced false negatives to 1 and false positives to 8 and increased the PPV to 63.6%. These results confirm that soil hydrological conditions are key factors for torrential flow triggering and may improve early-warning predictions.

Keywords Monitoring · Torrential flows · Rainfall · Soil moisture · Rainfall threshold · Hydro-meteorological threshold

Introduction

Torrential flows, like debris flows and debris floods, are mixed masses of debris and water that move at high velocities in steep channels in mountainous regions (Godt and Coe 2007; Hungr et al. 2014). The properties of these flows range between those of water floods and dry rock avalanches, depending on how solid and fluid forces interact in the flow dynamics (Iverson 1997). In addition to their relative high velocities, torrential flows are characterized by long runout distances and great impact forces (Jakob and Hungr 2005). The flow characteristics, paired with the fact that human settlements and activities have increased in mountainous regions,

turn torrential flows into one of the most important hazards (e.g., Takahashi 2014). Therefore, understanding the triggering and initiation mechanisms of these flows is crucial to mitigate theirs socio-economical losses (Jakob and Hungr 2005; Chen et al. 2019).

The initiation and triggering mechanisms of debris flows and debris floods involve many factors like sediment availability, slope angle, groundwater conditions (Iverson et al. 1997; Brayshaw and Hassan 2009; Takahashi 2014), and rainfall (Wieczorek and Glade 2005). Moreover, snowmelt can combine with rainfall to trigger torrential flows in late spring and early summer (Church and Miles 1987; Abancó et al. 2016; Mostbauer et al. 2018). Initiation mechanisms, or how debris sediments are transformed into flowing debris, are generally complex and not fully understood. Debris flows or floods can result from the occurrence of slope mass failure that later evolves into torrential flow, or from progressive channel bed and bank erosion due to intense runoff (Pastorello et al. 2020). In the first case, debris flow occurrence is mainly related to rainfall-induced instability of slope superficial debris layers (Iverson et al. 1997; Takahashi 2007; Imaizumi and Sidle 2007; Berger et al. 2011), whereas it is more likely associated to the reach of critical runoff rate (or critical surface discharge) (Berti and Simoni 2005; Coe et al. 2008; Gregoretti and Fontana 2008; Simoni et al. 2020) in the second case.

Whatever is the initiation mechanism, flowing mass mobilization is controlled by slope hydrological response to rainfall. The latter is usually characterized by discriminating between predisposing (e.g., soil moisture, topography, thickness of sediments) and triggering (e.g., rainfall) factors (Bogaard and Greco 2016). Hydrological data can be directly monitored in the field (Walker et al. 2004; Ochsner et al. 2013), or using remote sensing technologies (Zhao et al. 2010) or by model (Ponziani et al. 2012). Alternatively, in absence of measurements, hydrological conditions can be estimated from current and antecedent rainfall records by using simple infiltration (Chleborad et al. 2008) or hydrological regional models (Zhao et al. 2020). This last approach suffers however the limitations that the relationship between antecedent rainfall and antecedent soil moisture is indirect, as the result of the coupled interactions between several processes like infiltration, evapotranspiration, snowmelt, and drainage (Brocca et al. 2008).

As a consequence, monitoring of rainfall and groundwater hydraulic variables (soil moisture, pore water pressure) appears to

be a fundamental task for a better understanding and prediction of slope hydrological response in the initiation zone of torrential flow. This task is however often made difficult by the field conditions prevailing in mountain areas prone to torrential flows: harsh climatic conditions, difficult access, steep slopes covered by unconsolidated debris, rockfalls and other hazardous phenomena, etc. (Berti et al. 2000; Comiti et al. 2014). In addition, full understanding of torrential mechanisms requires monitoring debris mass flow dynamics in the run-out zone (Abancó et al. 2014; Bel et al. 2017; Hürlimann et al. 2019b). When feasible despite of these difficulties, field monitoring provides data of high interest, susceptible to be used for the definition of initiation thresholds in Landslide Early Warning Systems (LEWSs), an essential tool for debris flow detection and prediction.

Current prediction tools rely mostly on rainfall thresholds in terms of rainfall intensity and duration (ID) (or other similar rainfall characteristics) (Caine 1980; Crosta and Frattini 2001; Gregoretti and Fontana 2007; Guzzetti et al. 2008), although it is well-known that the hydrological conditions previous to the triggering rainfall play a crucial role in debris flow initiation. Some recent studies have included subsurface hydrological variables, such as pore water pressure or soil moisture, in the definition of the thresholds (Glade et al. 2000; Godt et al. 2006; Ponziani et al. 2012; Bogaard and Greco 2018; Mirus et al. 2018b; Marino et al. 2020). For example, Mirus et al. (2018a) improved the rainfall threshold defined by Scheevel et al. (2017) by replacing the antecedent rainfall with the average saturation obtained from direct volumetric water content (VWC) measurements over the same timeframe.

This work analyzes the effect of both rainfall and hydrological slope conditions on the triggering of debris flows and debris floods in the Rebaixader catchment (South Central Pyrenees, Spain). For this purpose, we analyze VWC and suction measurements in the initiation area of Rebaixader torrential flows together with rainfalls and mass movement occurrences. In a first part, the relevance of using specific rainfall characteristic values (mean and peak values of different durations) for the computation of triggering thresholds is investigated by statistical analysis of flow events over the period 2009–2020. In a second part, we study the improvement in prediction provided by the use of hydro-meteorological thresholds, which consider both rainfall maximum intensity and current value of VWC at an appropriate location within the slope. It should be highlighted that the definition of hydro-meteorological thresholds is novel in the Rebaixader catchment and intends to improve the triggering thresholds based exclusively on rainfall parameters (i.e., mean rainfall intensity), which have been defined in previous studies (Abancó et al. 2016; Oorthuis et al. 2021). Last but not least, the implementation of both rainfall and hydro-meteorological thresholds into a LEWS for torrential flow prediction is validated using the year 2019 as example.

Materials and methods

The Rebaixader catchment

General settings

The Rebaixader catchment is a small first-order basin tributary of the Noguera Ribagorzana River, located in Southern Central Pyrenees (42°32′57.02" N 0°45′12.57" E), which shows a typical torrential morphology (Figure 1). The source area is a concave scarp draining 0.53 km² with a slope of 35° in average and 70° in maximum, where the debris flows and debris floods initiate. The channel below the initiation zone is not very long (150 m and 8–10 m wide) with an average slope of 21°. At the lower part of the basin, there is a small deposition fan (0.09 km² extent) with a mean slope of 8°. The elevation ranges from 2475 to 1345 m a.s.l., at the fan apex.

The source material is a glacial till deposited during the Last Glacial Maximum, and 15 to 20 m thick, thus providing almost unlimited material for torrential flows (Hürlimann et al. 2014). The bedrock consists of Paleozoic slates and phyllites (Muñoz 1992).

The till is mainly composed by sandy gravels and large boulders. The fraction of fines (silt and clay) is minor but variable. Two layers are distinguished in the deposit. The lower one is a subglacial till overlying the bedrock, which has a higher proportion of fines and a lower porosity than the upper one that is a supraglacial till. Geotechnical properties of the two soils are indicated in Table 1 and Fig. 2. The geotechnical properties were obtained from field and laboratory tests mainly from repacked soil samples after removing the cobbles (<100 mm). Regarding the strength parameters, effective cohesion and effective friction angle were obtained in previous studies from large direct shear tests (Costa 2014).

The climate in the area is mainly affected by the orographic effects of the Pyrenees, the westerly winds from the Atlantic, and the closeness to the Mediterranean Sea. The mean annual precipitation in the site fluctuates between 800 and 1200 mm (Abancó et al. 2016).

Availability of granular sediment and steep slopes predisposes the basin to torrential flows while amount of water input by rainfall provides the main triggering factor in the Rebaixader catchment. Previous analyses of rainfall in the catchment showed that most of the debris flows and debris floods were triggered by short-duration and high-intensity rainfalls, mainly associated with summer convective rainstorms from June to September (Abancó et al. 2016; Oorthuis et al. 2021). Other torrential flow events, as well as rockfalls, have been observed during spring and were affected by snowmelt (Hürlimann et al. 2010, 2012). In addition, some minor torrential activity has been observed during autumn, which is generally triggered by long-duration and medium intensity rainfall events. The frequency of torrential events is close to one debris flow and two debris floods per year. The estimated volumes range from 100 to 16,000 m³, corresponding to the smallest debris flood and the largest debris flow, respectively.

Monitoring description

The Rebaixader site presents a high torrential activity, which, in addition to the limited size of the catchment and the lack of protective measures, makes this catchment an ideal location for hosting a monitoring system. The monitoring in the catchment started in 2009 with the aim of detecting torrential flows such as debris flows and debris floods. For this purpose, two monitoring stations were installed in the channel area and lower part of the catchment: the meteorological station METEO-CHA, and the FLOW-WR station, which focuses on the detection and analysis of the flow dynamics. Since then, the monitoring system has been maintained and further improved. It includes today four active stations (Figure 1c).
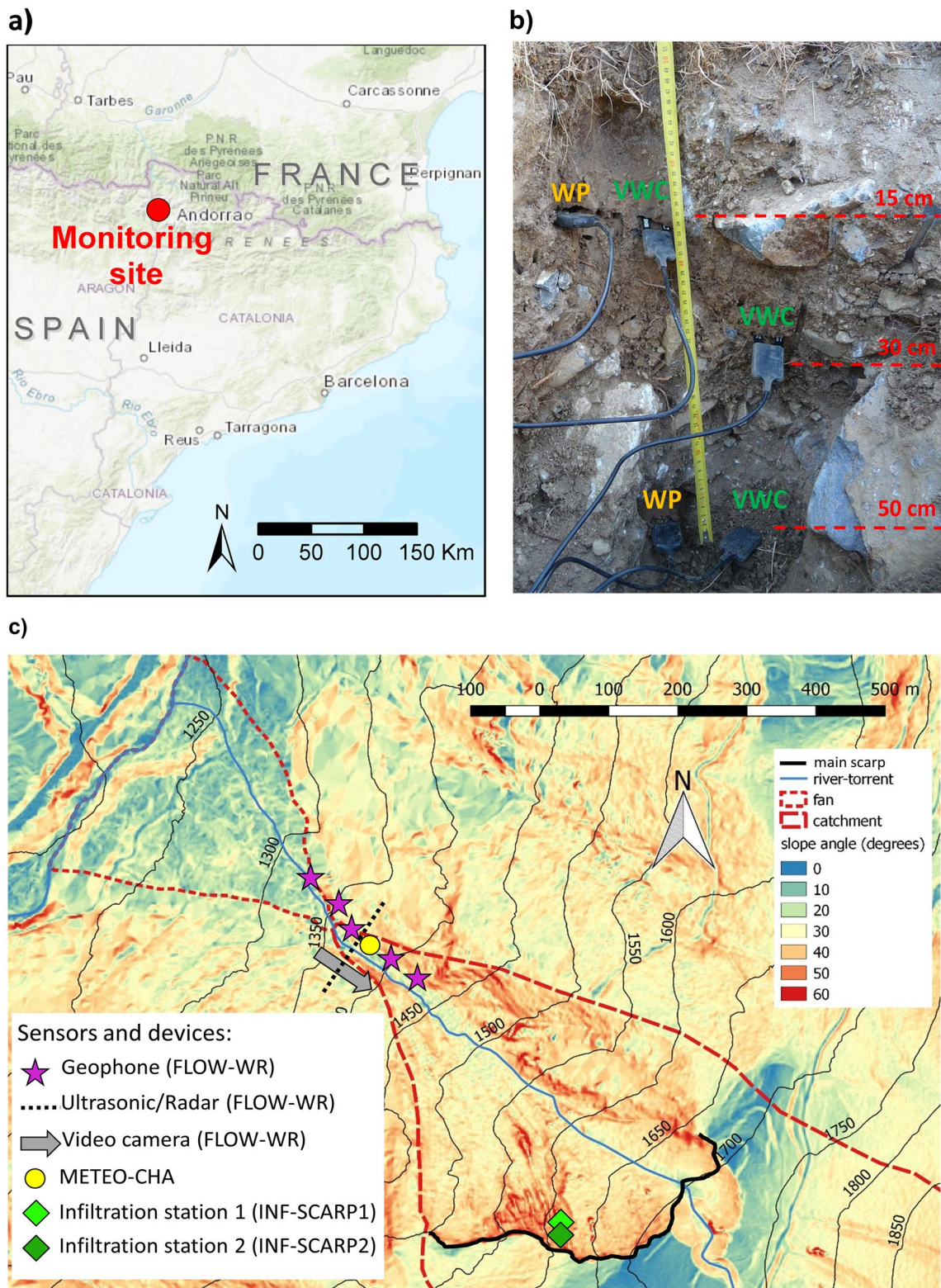
Fig. 1 The Rebaixader catchment and monitoring site. **a** Location of the Rebaixader catchment in the Pyrenees. **b** Detail of the supraglacial till at infiltration station INF-SCARP1 indicating the volumetric water content (VWC) and water potential (WP) sensors (see **c** for location). **c** Slope angle map and location of the infiltration stations (INF-SCARP1 and INF-SCARP2), meteorological station (METEO-CHA), and the specific sensors at the flow dynamics station (FLOW-WR)

Table 1 Soil properties obtained from field and laboratory tests mainly from repacked soil samples and after removing the cobbles (< 100 mm). Strength parameters (ϕ' effective friction angle and c' effective cohesion) were obtained from large direct shear tests (Costa 2014)

Soil property	Material type		Method
	Supraglacial	Subglacial	
Gravel content (%)	54.9–56.3	38.1[*]	GSD, by sieve
Sand content (%)	28.8–41.2	40.5[*]	GSD, by sieve
Fine content (%)	2.1–8.8	21.4[*]	GSD, by sedimentation
Dry density (ρ_d) (Mg/m^3)	1.82[*]	1.66–1.92	Field and laboratory test
Density of solid particles (ρ_s) (Mg/m^3)	2.57–2.62	2.56–2.7	Pycnometer test (lab.)
Soil moisture (ω) (%)	6.1–16.5	6.7–22.2	Oven drying (lab.)
Porosity (n) (−)	0.29[*]	0.25–0.39	Laboratory test
Saturated hydraulic conductivity (K_{sat}) (m/s)	4.06×10^{-6}–2.53×10^{-4}	ND	Infiltration test (lab. and field)
Effective soil friction angle (ϕ') (°)	40–45	ND	Large direct shear test (lab.)
Effective cohesion (c') (kPa)	0–4	ND	Large direct shear test (lab.)

[*]: only one test performed; *ND,* no data; *GSD,* grain-size distribution

The METEO-CHA station is located next to the channel. It consists of a tipping bucket rain gauge with a resolution of 0.2 mm, an air temperature sensor (measurement range − 40° to + 70 °C), and a relative humidity sensor (measurement range 0 to 100%). The FLOW-WR station consists of five geophones distributed along the channel, two flow depth sensors, and a video camera. This is the most important station as it detects the passing torrential flows and allows characterizing its main properties (flow type, velocity, flow depth, and volume). Two infiltration stations were installed within the supraglacial till layer at the highest part of the open scarp on a steep (30–40°) and bare slope, and close to the most active part of the initiation area: INF-SCARP1 station (installed in 2012) and INF-SCARP2 station (installed at the end of 2015). These infiltration stations are recording a total of eight soil volumetric water contents (VWC) at depths between 5 and 50 cm (measurement range 0–0.57 m^3/m^3) and measure soil matric suction (measurement range 5–500 kPa) and soil temperature (measurement range − 40 °C to + 50 °C) at 15 and 50 cm depth. It should be noted that the matric suction sensors installed are an older version than the current available ones, which have a suction range of 9 to 100,000 kPa. Nevertheless, the installed matric suction sensors can exceed the upper bound of 500 kPa, despite lower measurement accuracy. The soil VWC and matric suction sensors, installed in the near surface, aim to improve the knowledge of the effects of soil hydrologic conditions on rainfall infiltration and runoff, which is known to trigger torrential flows in other catchments (Berti and Simoni 2005). In this sense, the precise initiation mechanisms for the triggering of torrential flows are complex and may be divided into two different types: due to shallow landslides connected to the channel network or by

Fig. 2 Grain-size distribution of materials taken at the source area of torrential flows (one sample of subglacial till and three samples of supraglacial till)
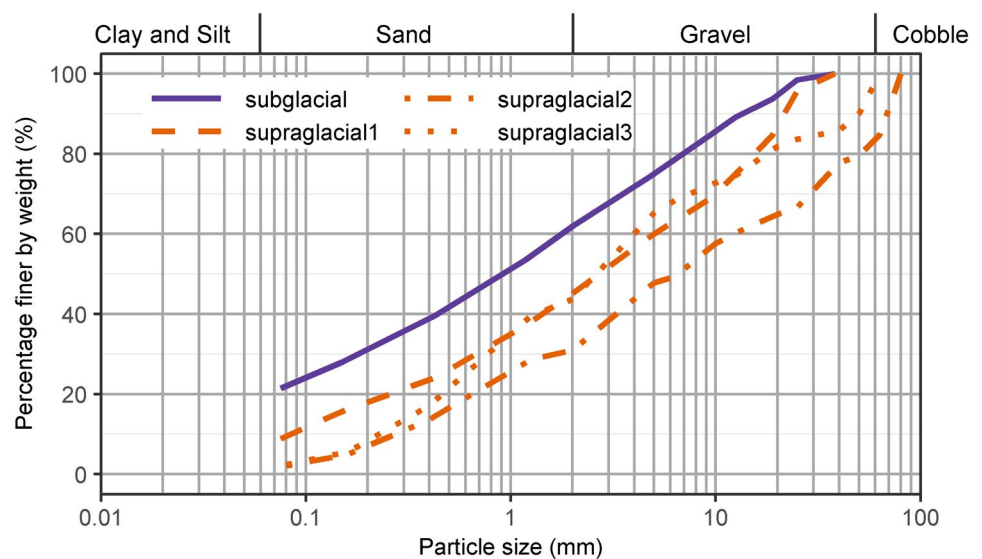
surficial erosion due to intense rainfall runoff. In the first scenario, a shallow landslide caused by rapid infiltration in the surficial soil layer may evolve into a torrential flow if the slope failure is connected to the channel network (Rickenmann 2016). In the second case, torrential flows may develop from intense erosion and progressive sediment entrainment during rainfall runoff (Simoni et al. 2020). That said, the instruments may be not installed sufficiently deep to necessarily account for a shallow perched water table. Figure 1bshows the installation of the soil VWC and matric suction sensors at theinfiltration station INF-SCARP1.

The FLOW-WR station has a low-frequency sampling rate of 2 h due to power and data management limitations. Under non-event conditions, it takes daily images of the channel cross section. Under event conditions, if a given ground vibration threshold is exceeded, FLOW-WR station samples sensors at 1 Hz and turns on the video camera (Abancó et al. 2012; Hürlimann et al. 2014). Conversely, the meteorological and infiltration stations have a constant sampling rate of 5 min. In addition, all the stations installed in the initiation area are connected to a common wireless sensor network (WSN), which uses a long-range wireless technique to transmit the data. All the monitored data are sent to the university server by GSM/GPRS communication, while multiple solar panels and batteries supply power to the monitoring stations. Table 2 summarizes the sensors that are actually installed at the Rebaixader monitoring site; for further details, see Hürlimann et al. (2014, 2019a).

Table 2 List of the sensors installed at Rebaixader monitoring site. More detailed information can be found in Hürlimann et al. (2014, 2019a).

Location	Station	Process (predisposing/triggering/dynamics)	Parameters (range)	Sensor model (nº of sensors)	Installation depth (cm)	Installation year
Open scarp (initiation area)	INF-SCARP 1	Predisposing mechanisms	Water potential (-5 to -500 kPa) and soil temperature (-40 to 60 °C)	Decagon MPS-2 (2)	15	2012
					50	2012
			Volumetric water content (0 to 0.57 m³/m³)	Decagon 10HS (4)	5	2018
					15*	2012
					30*	2012
					50	2012
	INF-SCARP 2 (left vertical profile)	Predisposing mechanisms	Volumetric water content (0 to 0.57 m³/m³)	Decagon 10HS (2)	15	2018
					20	2015
	INF-SCARP 2 (right vertical profile)			Decagon 10HS (2)	10	2018
					30	2015
Channel	METEO-CHA	Predisposing and triggering mechanisms	Rainfall (0.2 mm/tip), air temperature ($-40°$ to $+70$ °C), and relative humidity (0 to 100%)	Decagon ECRN-100 (1)	-	2009
				Campbell Scientific CS215 (1)		2009
	FLOW-WR	Flow dynamics	Ground vibration	Geophones Geospace 20 DX (5)		2009
		Flow dynamics	Flow depth (350 to 6000 mm)	Ultrasonic sensor Pepperl+Fuchs UC6000-30GM-IUR2-V15 (1)		2009
		Flow dynamics	Flow depth (max. 30 m)	Radar sensor VEGAPULS SR 68 (1)		2015
		Flow dynamics	Visual observation	Video camera Mobotix MX-M12D-Sec-DNight-D43N43 (1)		2011

*VWC sensors selected for the definition of hydro-meteorological thresholds

Threshold types and definitions

Thresholds for debris flows and debris floods usually correspond to critical rainfall characteristics, which, when exceeded, are prone to trigger a torrential flow event (Guzzetti et al. 2007). Such thresholds are defined by discriminating rainfalls that trigger an event from those that do not. In this study, we evaluate and compare two types of thresholds. The first and most common type relies only on rainfall characteristics and will be subsequently named as "rainfall threshold." The second type combines both hydrologic soil conditions (predisposing factors) and triggering rainfall characteristics and will be subsequently referred as "hydro-meteorological threshold."

Rainfall versus hydro-meteorological thresholds

Rainfall thresholds can be based on several rainfall parameters. One threshold commonly used is expressed as a relationship between rainfall mean intensity (I_{mean}, mm/h) and duration (D, h). It is drawn as a line in the graph I_{mean}-D where the x and y-axes are plotted in logarithmic scale to capture rainfall data of multiple orders. I_{mean}-D thresholds are generally fitted by a power law equation:

$$I_{mean} = aD^b \tag{1}$$

with a being the intensity of a rainfall event of unit duration, and b the slope of the log-plotted threshold line.

A second type of threshold is expressed by values of maximum rainfall intensity for selected rainfall durations (I_{max_dur}) and takes the form:

$$I_{max\_dur} = c \tag{2}$$

where c is a constant and I_{max_dur} is the value of maximum intensity necessary to trigger torrential flow for duration dur. In this work, six durations have been considered: 5, 10, 15, 20, 30, and 60 min. It results in a graph where maximum rainfall intensities are plotted in the y-axis.

Hydro-meteorological threshold is defined as a relationship between soil volumetric water content (VWC) at the start of the considered rainfall event, reported on the x-axis, and maximum rainfall intensity, plotted in the y-axis. In the present work, VWC monitored at two depths (15 and 30 cm) in a profile located at the scarp of the slope (INF-SCARP1) will be used as they have the longest and most complete VWC time series. To determine the hydro-meteorological threshold, the maximum rainfall intensities given by Eq. (2) will then be plotted against VWC for the six different time durations previously mentioned. This results in six hydro-meteorological thresholds at each sensor depth, expressed by the following linear equation:

$$I_{hydro-meteo} = I_{max\_durVWCdepth} = dVWC + e \tag{3}$$

where $I_{max\_durVWCdepth}$ is the value of maximum intensity necessary to trigger torrential flow for duration dur when considering VWC at the start of the rainfall event at $depth$ 15 or 30 cm, d is the slope, and e is the y-intercept when VWC = 0.

The rainfall intensity used in these different concepts may represent an "instantaneous" measure of the rainfall rate, or an average value of precipitation, depending on the length of the observation period (in: Guzzetti et al. 2007). Particularly, the mean rainfall intensity used in the I_{mean}-D approach does not consider intensity variations during rainfall event, which can lead to underestimate the rainfall intensity that actually triggered the torrential flows. I_{mean}-D thresholds ignore thus other information contained in the rainfall time series, such as peak intensities and antecedent rainfall conditions (Hirschberg et al. 2021). Some studies show that the variability of the rainfall intensity, or the shape of the rainfall hyetographs, can strongly influence the triggering of landslides (D'Odorico 2005; Peres and Cancelliere 2014). Furthermore, maximum rainfall intensities have been shown to have high predictive power at high temporal resolutions (≤10 min) (Abancó et al. 2016; Bel et al. 2017; Hirschberg et al. 2021). This suggests that maximum rainfall intensity at time of landslide triggering would be a better characteristic than the I_{mean} value, which supports the definition of thresholds based on Eqs. (2) or (3).

Available data

Before analyzing the performance of the different thresholds, it is necessary to define criteria for identifying rainfall events and to standardize the way in which threshold variables are computed. Particularly, the computation of rainfall duration is not straightforward, as there could exist several overlapped rainfall events or long episodes of low rainfall before and after a high precipitation event. In this work, according to Abancó et al. (2016), only the rainfall events delimited by a period of no rainfall for at least 1 h before and after the event are considered. Once defined the rainfall events, episodes with total precipitation of less than 0.4 mm were removed. As a result, 1037 events have been analyzed between 2009 and 2020: 1000 without triggers of torrential flows and 37 associated to occurrence of torrential events. In the following, this inventory will be referred as the rainfall dataset and includes all the monitored rainfall events and all the torrential flows detected in the catchment since the monitoring started. Each rainfall event was then characterized with the following parameters: total precipitation, P_{tot} (mm), total duration, D (h), mean rainfall intensity, I_{mean} (mm/h), and by six maximum rainfall intensities, I_{max_dur} (mm/h) for 5, 10, 15, 20, 30, and 60 min. The rainfall dataset (2009–2020) is further used in this work to assess and compare the performances of the I_{mean}-D vs I_{max_dur} rainfall thresholds over a longer period.

In parallel, a hydro-meteorological dataset has been set-up by adding VWC measurements to the rainfall dataset. VWC measurements include all data prior to the start of rainfall events (VWC_i) listed in the rainfall dataset and monitored at 15 and 30 cm depth at INF-SCARP1 station. Due to several technical problems of short duration, VWC recordings present some missing data. Because of that and the fact that the INF-SCARP1 station was installed in the late 2012, the hydro-meteorological dataset has a shorter time period than the rainfall dataset and includes 15 triggering events and 470 non-triggering events with both VWC and rainfall data between 2013 and 2020. The hydro-meteorological dataset (2013–2020) is used to compare the performance of rainfall and hydro-meteorological thresholds.

Finally, in both the rainfall and hydro-meteorological dataset, a correction has been applied on rainfall data during the last torrential flows season (2020). Indeed, the rain gauge clogged twice during this period, leading to time intervals with no rainfall records. Missing data have been then completed by means of a

correlation established between rainfalls observed in 2016 and 2017 at METEO-CHA station and at a rain gauge (https://altaribagorca.smartyplanet.com/ca/estacions/estacio/228/smartis/) located at a distance of 1800 m from our meteorological station.

Evaluation of threshold performance

The rainfall and hydro-meteorological threshold equations (Eq. (1), (2), and (3)) were calibrated and evaluated by means of different scoring metrics from the receiver operating characteristics (ROC) and precision-recall curves (PRC) analysis (Fawcett 2006). First, the confusion matrix was obtained for any given threshold, thus, for any combination of a and b in Eq. (1), c in Eq. (2), and d and e in Eq. (3). The confusion matrix counts the four possible outcomes output by the threshold prediction: TP (true positive or true alarm), FP (false positive or false alarm), TN (true negative or true non-alarm), and FN (false negative or failed alarm). Second, the following evaluation/scoring metrics were computed for each threshold:

$$TPR = True\ positive\ rate\ or\ Sensitivity = Recall = \frac{TP}{TP + FN} \quad (4)$$

$$FPR = False\ positive\ rate\ or\ Specificity = \frac{FP}{FP + TN} \quad (5)$$

$$Precision\ or Positive\ predictive\ value = \frac{TP}{TP + FP} \quad (6)$$

$$F_1 - score = \frac{2}{\frac{1}{Precision} + \frac{1}{TPR}} = 2 \cdot \frac{Precision \cdot TPR}{Precision + TPR} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (7)$$

The most popular evaluation method for landslide thresholds is the ROC curve and the area under the curve (AUC), which relates the trade-off between the TPR and the FPR. This method is, however, not the most robust and easy to interpret when the negative class, or non-triggering events, is the majority class (Fawcett 2006; Saito and Rehmsmeier 2015). In our case, both rainfall and hydro-meteorological datasets are highly imbalanced as they present a very low number of positive events, i.e. events that triggered a torrential flow, compared to the high number of non-triggering events. When comparing thresholds with an elevated number of TN events, the FPR scores are low and the resulting ROC curves and AUC are more difficult to compare and to interpret. Hence, the precision-recall or precision-TPR curves (PRC) are more informative and appropriate for our evaluation, since the large number of TN events does not affect the precision, neither the recall metric.

Torrential flow thresholds aim at maximizing the prediction of TP events while minimizing the FP and FN events, which results in a higher F_1-score. The F_1-score weights precision and TPR equally and is the most often used variable when learning from imbalanced datasets (Weiss 2013). Therefore, we select and optimize every threshold equation (I_{mean}-D, I_{max_dur}, and hydro-meteorological) by maximizing the F_1-score and under the premise that the largest debris flows are predicted as TP events. Another consideration for threshold selection is that the number of torrential events correctly predicted is higher than 85% (TPR > 0.85).

In continuation, we will present an example for the rainfall I_{mean}-D thresholds and the ROC and PRC curves (Fig. 3). A threshold with a perfect skill is represented as a point in the ROC space at TPR = 1 and FPR = 0, while it is defined in the PRC space at Precision = 1 and TPR = 1. First, the I_{mean}-D threshold with the best performance and highest F_1-score is plotted (Fig. 3a, Threshold 1). The best-fitting threshold is represented in the ROC and PRC space as a point (Threshold 1 in Fig. 3b and c, respectively). Second, the ROC and PRC curves are obtained by varying the y-intercept between the minimum (Fig. 3a, Threshold 2) and the maximum (Fig. 3a Threshold 3) y-range and while keeping the slope of the curve constant. This results in a point in the ROC and PRC space for every computed y-intercept and for a fixed threshold slope. The ROC and PRC curves are then plotted by joining the resulting points. Threshold 4 defines the minimum I_{mean}-D threshold for the triggering of torrential flows. The PRC and ROC curves are a valuable tool for the evaluation and interpretation of the threshold performance. However, ROC curves (Fig. 3b) can be misleading when compared to the PRC curves (Fig. 3c): the elevated number of TN events diminishes the FPR and does not truly reflect the high number of FPs or false alarms. This results in a quite optimistic ROC curve, even if the number of FPs is high related to the lower number of TPs, in comparison with the PRC curve. This is not the case of the PRC curve, which is more informative when dealing with imbalanced datasets, as it does not rely on the correctly predicted non-triggering events. Therefore, the PRC approach better represents the variability of the thresholds performance in torrential flow detection.

Results and discussion

Monitoring data analysis

VWC, suction, and daily rainfall time series from January 2016 to October 2020 are shown in Fig. 4. This period covers the most complete and continuous recordings at the infiltration station INF-SCARP1, although some data are missing due to technical problems. Note that the VWC readings at 5 and 50 cm depth are only available since June 2018. In addition, vertical dashed lines indicate the timing of the peak discharge of torrential flows as they pass the radar and US sensors at the FLOW-WR station.

Time series evidence the seasonal fluctuation of soil water variables, driven by meteorological variations of precipitation and also snowmelt and evapotranspiration. Generally, higher VWC and lower suctions are observed during spring and are related to heavy rainfalls and to possible water supply due to snowmelt (Hürlimann et al. 2010). In contrast, high suction values and low VWCs are mainly developed during the summer months. On top of the seasonal fluctuation, daily variations with magnitude of about 25% of the seasonal variations can be observed as the result of the daily cycle of evaporation and, when applicable, hourly rainfalls.

It can be seen from Fig. 4 that torrential flows are triggered between late spring and early fall and not necessarily due to the heaviest rainfalls in terms of daily precipitation. Soil hydrologic state previous and during torrential flows appears also to vary from one event to other: suction and VWC can range from dry (VWC ≈ 0.10 m³/m³ and suction ≥ 1000 kPa) to wet (VWC ≈ 0.34 m³/m³ and suction ≈ 5 kPa) conditions. This demonstrates that
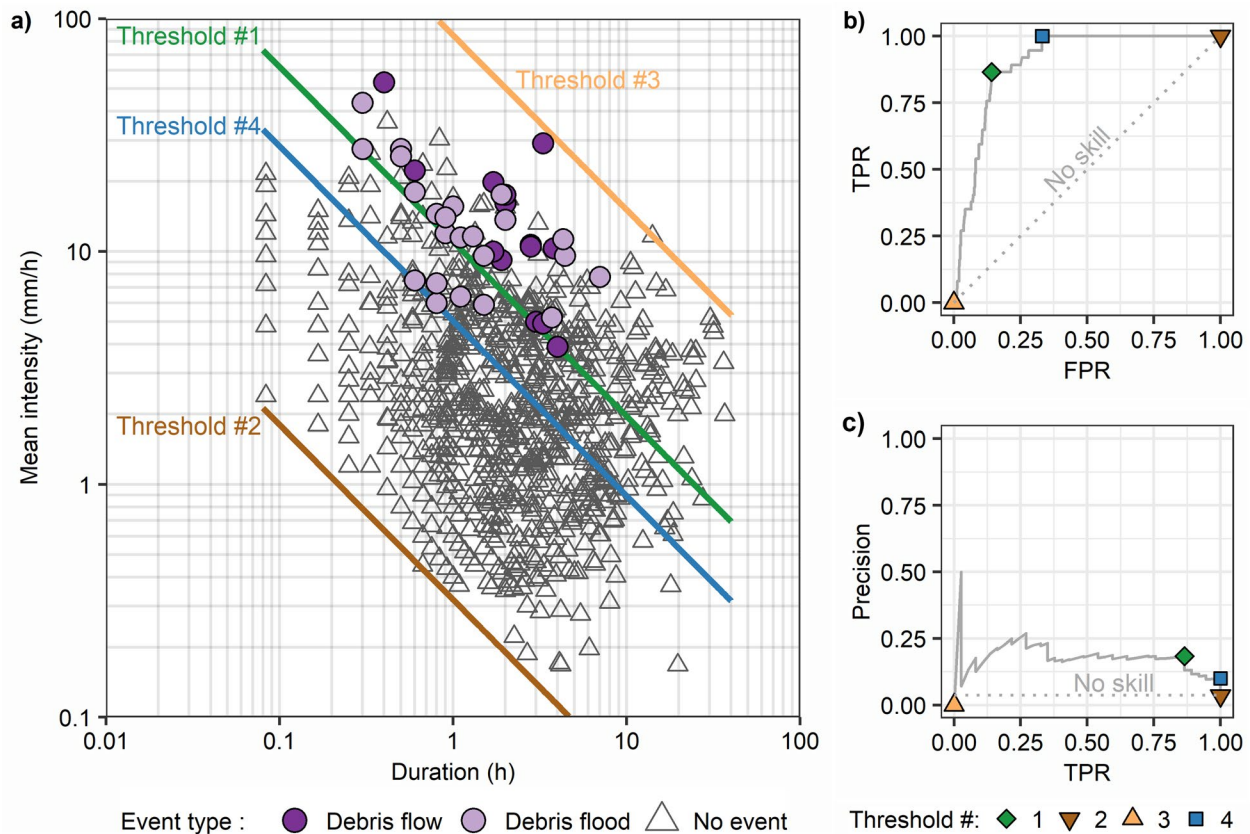
Fig. 3 Methodology for the evaluation and interpretation of the threshold performance using the I_{mean}-D threshold for illustration. **a** Example of rainfall I_{mean}-D thresholds with the best performance (Threshold 1). The ROC (**b**) and precision-recall (**c**) curves are defined by varying the y-intercept while keeping the threshold slope constant. In all plots, Threshold 1 has the best performance; Threshold 2 defines the lower limit in the I_{mean}-D plot, while Threshold 3 defines the higher limit; Threshold 4 defines the minimum I_{mean}-D threshold for torrential flow detection

the full-saturation of the sediment layer in the near surface is not a required condition for the initiation of debris flows and debris floods, as can be stated by the three last torrential flow events of 2017 where both VWC and suction measurements indicated partial saturation. Thus, the hillslopes can remain unsaturated at the monitoring site during the initiation of torrential flows.

Table 3 presents the torrential flow inventory since the installation of the infiltration station INF-SCARP1 (June 2013 to August 2020). Torrential flows of year 2014 and year 2015 were not added, as both the VWC at 15 and 30 cm depth were missing for these years. No torrential events were detected during 2016. As a result, the final inventory includes 7 debris flows and 8 debris floods. For each of them, VWC measurements at 5, 15, 30, and 50 cm depth and rainfall characteristics are indicated.

Data in Table 3 indicate that torrential flows in the Rebaixader catchment are mainly triggered by short duration and high intensity rainfalls between 12:00 and 18:00 h UTC. These rainfalls are strongly related to convective storms, which are generally shorter than 3 h and have maximum rainfall intensities ranging from 48 to 120 mm/h regarding the 5-min recording rate (4 to 10 mm in 5 min). Data also indicate that long duration triggering rainfalls are characterized by lower mean intensity than short duration rainfalls.

This aspect has to be related to VWC measurements. The latter generally reveal that higher rainfall intensities are required to trigger torrential flows when the soil is in a rather dry initial condition (e.g., VWC between 0.07 and 0.18 for flow 7, 8, 9, and 15). Conversely, lower triggering rainfall intensities are observed when the soil is initially wetter (e.g., VWC between 0.22 and 0.30 for flow 1, 3, 4, 5, 12, and 14). In any case, largest debris flows (volume > 9000 m³ for flow 3, 12, and 13) were preceded by medium to high VWCs (in the range of 0.18 to 0.30 at all depths) and were not triggered by the heaviest rainfalls in terms of rainfall intensity and total rainfall amount. All this suggests that the triggering of torrential flows depends on both the hydrologic soil conditions and the rainfall characteristics.

Figure 5 presents a detailed view of soil water content (VWC) time evolution at infiltration stations INF-SCARP1 and INF-SCARP2 during three selected events. The first two events (first two columns in Fig. 5) correspond to short duration and relatively high intensity rainfalls that triggered a debris flow (June 6, 2020) and debris flood (August 28, 2020), respectively. The third event (third column in Fig. 5) corresponds to a long duration and low intensity rainfall, which did not trigger any torrential flow. Soil is quite wet in the 1st and 3rd event and drier in the 2nd event.

Figure 5c and f evidence that triggering rainfall intensity is lower for initially wetter than drier soil conditions. In case of the
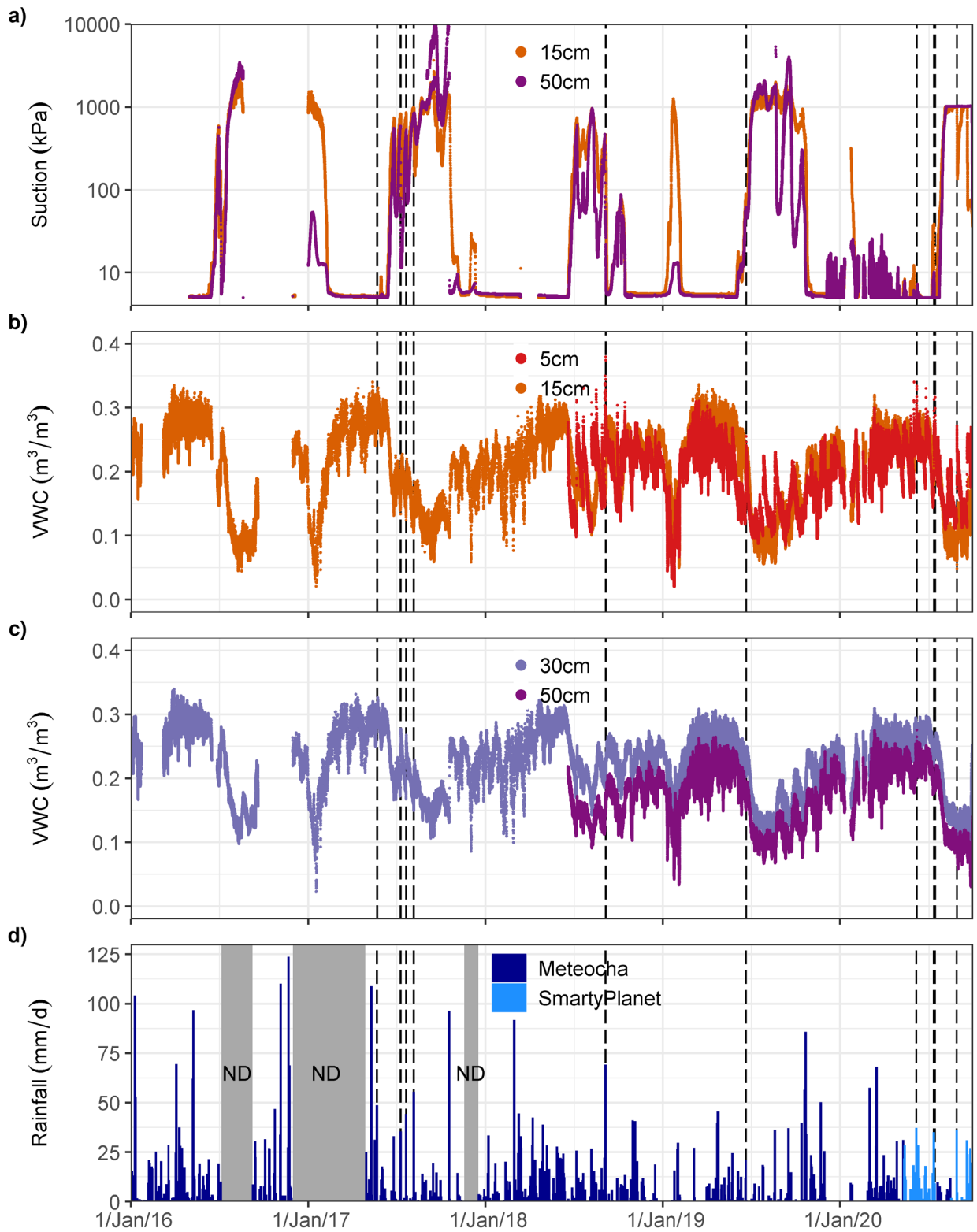
Fig. 4 Volumetric water content (VWC), suction, and daily rainfall time series from January 2016 to October 2020. The dashed vertical lines indicate the timing of the peak discharge of torrential flows at the FLOW-WR station. **a** Matric suction recorded at INF-SCARP1 station at 15 and 50 cm depth; **b** VWC recorded at 5–15 cm of INF-SCARP1 station; **c** VWC at 30–50 cm depth of INF-SCARP1; and **d** daily rainfall. ND stands for no data

Table 3 Debris flow (DFlow) and debris flood (DFlood) inventory combining both the rainfall characteristics and the volumetric water content previous to the triggering rainfall (VWC$_i$) at infiltration station INF-SCARP1. ND stands for no data

Event ID #	Event type	Date and triggering time (UTC)	Volume (m³)	Total rainfall (mm)	Duration (h)	Mean rainfall intensity (mm/h)	I$_{max\_5min}$ (mm/h)	I$_{max\_10min}$ (mm/h)	I$_{max\_15min}$ (mm/h)	I$_{max\_20min}$ (mm/h)	I$_{max\_30min}$ (mm/h)	I$_{max\_60min}$ (mm/h)	VWC$_i$ 5 cm (m³/m³)	VWC$_i$ 15 cm (m³/m³)	VWC$_i$ 30 cm (m³/m³)	VWC$_i$ 50 cm (m³/m³)
1	DFlood	05/06/2013 13:00	2100	54.5	7	7.8	49.2	42.6	36	32.7	28	26.4	ND	0.28	0.25	0.17
2	DFlood	17/06/2013 16:38	100	4.4	0.6	7.5	12	11.4	10.8	10.2	8.6	4.5	ND	0.27	0.24	0.16
3	DFlow	17/07/2013 13:05	10,400	15.8	3.3	4.9	96	58.8	41.6	33	23.2	12.8	ND	0.30	0.27	0.18
4	DFlood	23/07/2013 12:03	600	12.8	0.5	25.6	54	45.6	38	37.8	25.6	12.8	ND	0.29	0.29	0.17
5	DFlood	27/07/2013 17:15	450	14.5	0.3	43.5	70.8	64.2	54.4	43.5	29	14.7	ND	0.27	0.28	0.16
6	DFlood	22/05/2017 15:22	800	48.2	4.3	11.3	72	61.2	48.8	42	37.2	31.2	ND	0.27	0.28	ND
7	DFlow	10/07/2017 02:55	2950	33.2	1.7	19.9	100.8	75.6	66.4	55.2	54	31	ND	0.13	0.20	ND
8	DFlow	21/07/2017 07:08	3100	38.6	3.8	10.3	96	72	52.8	42	32.4	25	ND	0.12	0.17	ND
9	DFlood	06/08/2017 00:14	600	33.8	1.9	17.6	88.8	79.2	69.6	58.8	47.6	31.2	ND	0.10	0.15	ND
10	DFlow	05/09/2018 17:49	2500	22.2	0.4	53.3	117.6	106.8	85.6	66	44.4	22.2	0.24	0.17	0.21	0.15
11	DFlow	21/06/2019 15:23	2000	13	0.6	22.3	55.2	52.8	45.6	36.6	25.6	13	0.22	0.17	0.20	0.16
12	DFlow	06/06/2020 19:13	9600	29.5	2.75	10.7	79.1	63.6	53.8	45.6	32.7	19.2	0.26	0.25	0.28	0.24
13	DFlow	12/07/2020 12:00	10,000	28.8	2.75	10.5	74.7	60.1	50.8	43	34.2	22.4	0.2	0.19	0.23	0.19
14	DFlood	14/07/2020 14:42	700	8.8	1.5	5.9	48.4	38.9	32.9	27.8	22.1	14.5	0.23	0.22	0.25	0.20
15	DFlood	28/08/2020 16:21	1200	27.4	2	13.7	96.7	77.7	65.7	55.7	45	27.2	0.18	0.07	0.13	0.09

non-triggering rainfall event (Fig. 5i), intensity is low and certainly not enough to initiate a torrential flow, although initial VWC are relatively high and accumulated rainfall almost three times higher than for the triggering events. These results confirm that triggering rainfall intensity is strongly related to soil hydrological conditions, while total precipitation seems not to be a critical parameter.

Figure 5a, b, d, and e provide insights into the hydrologic response of the slope under the short duration/high intensity rainfall events. VWCs at shallow depths (5 and 10 cm) appear to react quickly to rainfall while deeper sensors (15 cm to 30 cm for INF-SCARP1, 15 cm to 20 cm for INF-SCARP2) exhibit an attenuated response. This generally expressed the propagation of a sharp front of higher water saturation within the shallowest soil layer. Deepest sensors (50 cm at INF-SCARP1, 30 cm at INF-SCARP2) present curiously higher changes than the ones at intermediate depths, which has been interpreted by modeling as resulting from the establishment of a lateral flow in the underlying layer (Luna 2015). In fact,

non-sequential or non-monotonic wetting, as well as non-uniform responses with depth, has also been attributed to rapid unsaturated zone responses (e.g., Nimmo et al. 2021) that can bypass layers through preferential flow pathways such as fractures, root or animal burrows, and other soil-ped structures. Another interesting aspect is that the increase of water content is higher when the soil is initially in a drier state and rainfall intensity is higher. In such a case, soil will lose stability more quickly. As a final remark, the timing of torrential flow peak discharge at the FLOW-WR station matches in both cases the timing of maximum rainfall intensity and peak in shallow water content changes. Conversely, the variation in VWC is progressive for the low intensity non-triggering rainfall and shows slow increase of water content at all sensors depths (Fig. 5g and h).

These results confirm that initial hydrologic soil condition and short duration rainfall are important controlling factors for the triggering of torrential flows. In the following part of the work, we will study the hydro-meteorological threshold by using VWCs at
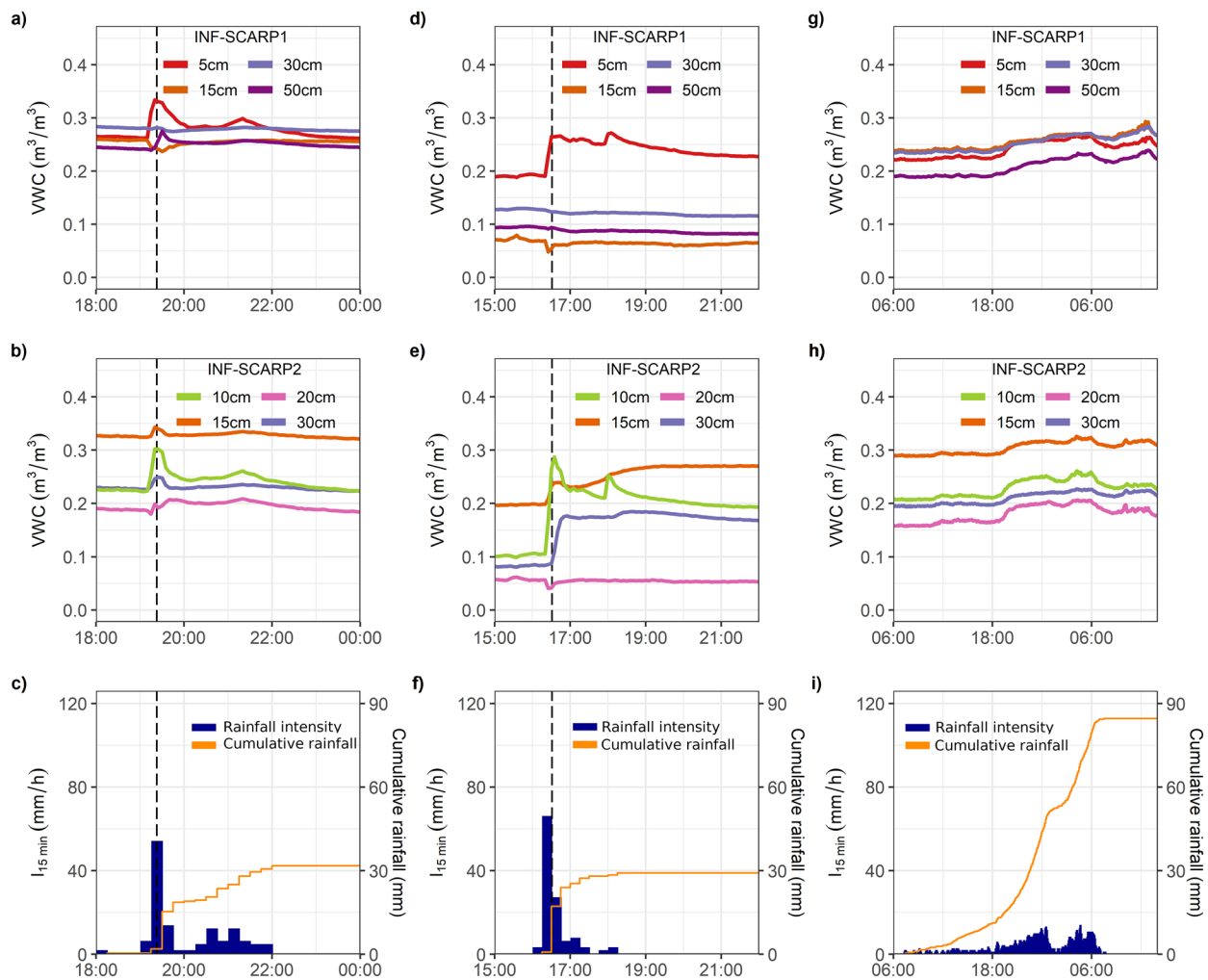

Fig. 5 Time series of soil volumetric water content (VWC) and rainfall showing the slope hydrologic response at the initiation zone of torrential flows during three rainfall events. First column (**a**, **b**, **c**): debris flow of June 6, 2020; second column (**d**, **e**, **f**): debris flood of August 28, 2020; and third column (**g**, **h**, **i**): non-triggering rainfall of April 24, 2019. First and second rows: VWC at the infiltration stations INF-SCARP1 and INF-SCARP2, respectively. Third row: rainfall intensity in 15-min duration (I_{15min}, dark blue color) and cumulative rainfall (orange color). The vertical dashed lines indicate the timing of torrential flow peak discharge at the FLOW-WR monitoring station

Table 4 Comparison of the equations and scoring metrics for the selected rainfall (I_{mean}-D and I_{max_dur}) and hydro-meteorological ($I_{max\_durVWCdepth}$) thresholds. See Eqs. (1–3) in the text for more details on equations. The total number of rainfall events are 1000 non-triggering and 37 triggering events for the 2009–2020 dataset, and 470-non triggering and 15 triggering events for the 2013–2020 dataset

Threshold type		Equation	Dataset	TP	FN	TN	FP	TPR	FPR	Precision	F_1-score
Rainfall: I_{mean}–D (Eq. 1) (Abancó et al. 2016)		$I_{mean} = 6.2 Duration^{-0.36}$	2009–2020	36	1	811	189	0.973	0.189	0.160	0.275
Rainfall: I_{mean}–D (Eq. 1) (Oorthuis et al. 2021)		$I_{mean} = 11 Duration^{-0.74}$	2009–2020	31	6	859	141	0.838	0.141	0.180	0.297
Rainfall: I_{mean}–D (Eq. 1) (this study)		$I_{mean} = 11.03 Duration^{-0.75*}$	2009–2020	32	5	858	142	0.865	0.142	0.184	0.303
Rainfall: I_{max_dur} (Eq. 2) (this study)		$I_{max\_5min} = 31$	2009–2020	32	5	915	85	0.865	0.085	0.274	0.416
		$I_{max\_10min} = 25$		32	5	923	77	0.865	0.077	0.294	0.438
		$I_{max\_15min} = 21*$		32	5	926	74	0.865	0.074	0.302	0.448
		$I_{max\_20min} = 18$		32	5	923	77	0.865	0.077	0.294	0.438
		$I_{max\_30min} = 16$		32	5	942	58	0.865	0.058	0.356	0.504
		$I_{max\_60min} = 9$		33	4	892	108	0.892	0.108	0.234	0.371
Rainfall: I_{mean}–D (Eq. 1) (Abancó et al. 2016)		$I_{mean} = 6.2 Duration^{-0.36}$	2013–2020	15	0	340	130	1	0.277	0.103	0.188
Rainfall: I_{mean}–D (Eq. 1) (Oorthuis et al. 2021)		$I_{mean} = 11 Duration^{-0.74}$	2013–2020	13	2	384	86	0.867	0.183	0.131	0.228
Rainfall: I_{mean}–D (Eq. 1) (this study)		$I_{mean} = 11.39 Duration^{-0.71*}$	2013–2020	13	2	397	73	0.867	0.155	0.151	0.257
Rainfall: I_{max_dur} (Eq. 2) (this study)		$I_{max\_5min} = 48$	2013–2020	14	1	448	22	0.933	0.047	0.389	0.549
		$I_{max\_10min} = 42$		13	2	454	16	0.867	0.034	0.448	0.591
		$I_{max\_15min} = 36$		13	2	457	13	0.867	0.028	0.500	0.634
		$I_{max\_20min} = 32*$		13	2	459	11	0.867	0.023	0.542	0.667
		$I_{max\_30min} = 23$		13	2	456	14	0.867	0.030	0.481	0.619
		$I_{max\_60min} = 12$		14	1	438	32	0.933	0.068	0.304	0.459
15 cm depth	Hydro–meteorological: $I_{max\_durVWCdepth}$ (Eq. 3) (this study)	$I_{max\_5min} VWC_{15cm} = -110VWC + 72$	2013–2020	14	1	452	18	0.933	0.038	0.438	0.596
		$I_{max\_10min} VWC_{15cm} = -103VWC + 70*$		13	2	463	7	0.867	0.015	0.650	0.743
		$I_{max\_15min} VWC_{15cm} = -88VWC + 60*$		13	2	463	7	0.867	0.015	0.650	0.743
		$I_{max\_20min} VWC_{15cm} = -62VWC + 47$		13	2	462	8	0.867	0.017	0.619	0.722
		$I_{max\_30min} VWC_{15cm} = -57VWC + 35$		13	2	458	12	0.867	0.026	0.520	0.650
		$I_{max\_60min} VWC_{15cm} = -35VWC + 22$		13	2	452	18	0.867	0.038	0.419	0.565

Table 4 (continued)

Threshold type	Equation	Dataset	TP	FN	TN	FP	TPR	FPR	Precision	F_1-score
30 cm depth Hydro–meteorological: $I_{max\_durVWCdepth}$ (Eq. 3) (this study)	$I_{max\_5min}VWC_{30cm} = -182VWC + 92$	2013–2020	14	1	455	15	0.933	0.032	0.483	0.636
	$I_{max\_10min}VWC_{30cm} = -169VWC + 81$		14	1	461	9	0.933	0.019	0.609	0.737
	$I_{max\_15min}VWC_{30cm} = -161VWC + 73*$		14	1	462	8	0.933	0.017	0.636	0.757
	$I_{max\_20min}VWC_{30cm} = -145VWC + 66$		13	2	463	7	0.867	0.015	0.650	0.743
	$I_{max\_30min}VWC_{30cm} = -76VWC + 41$		14	1	458	12	0.933	0.026	0.538	0.683
	$I_{max\_60min}VWC_{30cm} = -103VWC + 40$		13	2	453	17	0.867	0.036	0.433	0.578

*Optimal equations for each threshold type and each dataset focusing on the F_1-score

15 cm and 30 cm as representative of slope hydrologic state before rainfall occurrence and rainfall maximum intensity for several short durations as trigger variable.

Threshold analysis

Scoring metrics, ROC, and PRC curves

The scoring metrics of the thresholds with highest performance are listed in Table 4 for both the rainfall thresholds (I_{mean}-D and I_{max_dur} for the 2009–2020 and 2013–2020 datasets) and the hydro-meteorological thresholds ($I_{max\_durVWCdepth}$ regarding the 2013–2020 dataset). The thresholds in this study have been selected on the premise that more than 85% of the torrential events are correctly predicted as TPs (TPR > 0.85) and by maximizing the F_1-score. Furthermore, the performance of the selected I_{mean}-D rainfall thresholds in this work is compared with those defined in previous studies: Abancó et al. (2016) and Oorthuis et al. (2021) for the data registered during 2009–2014 and 2009–2018, respectively. Thus, the present work includes additional records from the last 2 years and spans from 2009 to 2020. The selection of the I_{mean}-D rainfall threshold in Abancó et al. (2016) was done by maximizing the correct prediction of torrential flows (debris flows and debris floods); this is by maximizing TPR. In contrast, Oorthuis et al. (2021) defined the I_{mean}-D rainfall threshold by fitting a power-law trend line using the registered debris flows and modifying the scale parameter until all the debris flows were located above the threshold line. The threshold equation with the best performance is highlighted for each type of threshold. In addition, some of the selected thresholds are presented in the ROC and PRC space (Fig. 6) in order to visualize and better interpret their performance.

First, the results in this study clearly highlight an increase of the I_{mean}-D rainfall threshold performance when compared to the one defined in Abancó et al. (2016). The number of false positives predicted in the present work is strongly reduced at the expense of predicting some debris flood events as false negatives (5 and 2 debris floods under the treshold line for the 2009–2020 and 2013–2020 dataset, respectively). See Fig. 7 for comparison between the I_{mean}-D rainfall threshold in this study and Abancó et al. (2016). In contrast, the I_{mean}-D rainfall threshold equation defined in this work slightly increases the threshold performance in comparison to the defined in Oorthuis et al. (2021), since both threshold equations are very similar.

In the following, we analyzed the rainfall I_{mean}-D thresholds defined in this study and realized that a high number of FPs (false alarms) were predicted, although more than 85% of the torrential flow events were correctly classified as TPs (Table 4). As a result, the precision (ratio between true alarms and the total triggered alarms) is low and around 18% and 15% for the 2009–2020 and 2013–2020 datasets, respectively. In the following, we focused on the comparison of the I_{mean}-D and I_{max_dur} rainfall thresholds. The first thing that stands out is the important reduction of FPs when considering maximum rainfall intensities instead of mean rainfall intensities. Hence, the rainfall I_{max_dur} thresholds mainly revealed higher precision and TPR values relative to the I_{mean}-D thresholds, when the same dataset is considered (Table 4 and

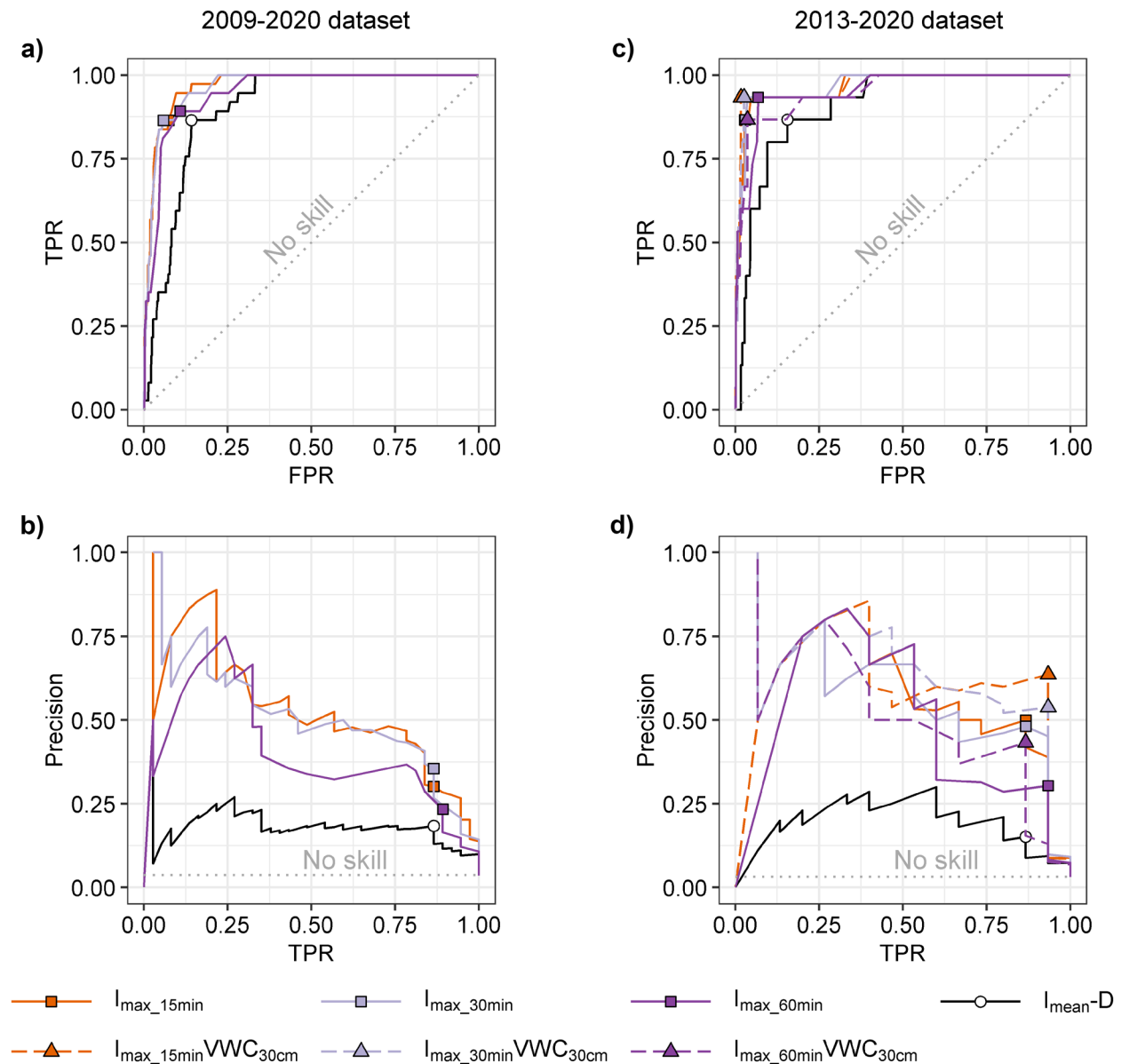**Fig. 6** Receiver operator characteristic (ROC) curves and precision-recall (PRC) curves comparing rainfall thresholds (I_{mean}-D and I_{max_dur}) and hydro-meteorological thresholds ($I_{max\_durVWC30cm}$) regarding the VWC readings at 30 cm depth. **a** ROC curves and **b** PRC curves considering the rainfall thresholds for the 2009–2020 dataset. **c** ROC curves and **d** PRC curves considering the rainfall thresholds and the hydro-meteorological threshold for the 2013–2020 dataset. Note that the hydro-meteorological thresholds, $I_{max\_durVWC30cm}$, and the rainfall thresholds, I_{max_dur}, are represented by the maximum rainfall intensities in 15, 30, and 60 min duration. The data shown corresponds to the thresholds defined in this study

Fig. 6). Therefore, the results point out that maximum rainfall intensities are more appropriate for assessing the triggering rainfall conditions in the Rebaixader catchment, compared to mean rainfall intensities. Furthermore, the best performance is attained with the maximum rainfall intensity for a duration of 15 min (I_{max_15min} threshold) for the 2009–2020 dataset, and of 20 min (I_{max_20min} threshold) for the 2013–2020 dataset. As a result, the precision increased to 30% and 54%, respectively. This duration of 15 and 20 min for torrential flows triggering is similar to the critical duration needed to reach the critical discharge for debris

flow initiation in other torrential catchments (for example 12 and 22 min, Pastorello et al. 2020).

In continuation, we checked whether the rainfall thresholds defined by maximum rainfall intensities, I_{max_dur}, could further improve the performance by including VWC measurements. For this reason, the hydro-meteorological thresholds, $I_{max\_durVWCdepth}$, and rainfall I_{max_dur} thresholds are compared in Fig. 6 for the same dataset (2013–2020). The results clearly highlight that adding subsurface hydrology decreases the number of FPs or false alarms compared to the rainfall I_{max_dur} thresholds, while the

number of TPs remains mainly equal. This produces increased precision and similar TPR compared to the rainfall I_{max_dur} threshold (Fig. 6c, d and Table 4). In any case, the combination of VWC and maximum rainfall intensities improved the thresholds performance and hence its prediction capabilities. Furthermore, one can see that the highest performance is achieved by linking the maximum rainfall intensity in 10–15-min duration with the VWC observations at 15 cm depth (which correspond to the $I_{max\_10minVWC15cm}$ and $I_{max\_15minVWC15cm}$ hydro-meteorological thresholds) and by combining the maximum rainfall intensity in 15-min duration with VWC at 30 cm depth ($I_{max\_15minVWC30cm}$). This duration of 10–15 min for torrential flows triggering, considering the hydro-meteorological thresholds, is also very similar to the 12–22 min proposed by Pastorello et al. (2020) and necessary to reach the critical discharge in other torrential catchments. In this situation, the precision increased up to 65% and 63%, considering VWC at 15 and 30 cm depth, respectively. Nevertheless, the hydro-meteorological thresholds defined by the VWC at 30 cm depth scored slightly higher than those defined at 15 cm depth, even though both predicted better results compared to the rainfall I_{max_dur} thresholds. Anyway, the results confirm that, for such high permeable soils, similar performance is obtained regardless the VWC depth.

In addition, it can be seen that ROC curves of Fig. 6 show less variability and are thus less informative than PRC curves when dealing with imbalanced datasets and high numbers of TN events. In this scenario, the FPRs are low and consequently the ROC curves seem quite optimistic and present similar results. For this reason, the rainfall and hydro-meteorological thresholds defined by maximum rainfall intensities are harder to compare (Fig. 6a and c), in contrast to the higher variability shown in the PRC curves (Fig. 6b and d). In essence, the PRC curves in Fig. 6d clearly illustrate the higher performance of the rainfall I_{max_dur} thresholds relative to the I_{mean}-D rainfall thresholds, and which is further improved by the inclusion of VWC observations in the hydro-meteorological thresholds.

Torrential flow thresholds

Finally, the selected rainfall and hydro-meteorological thresholds are plotted for a better visualization and interpretation of the results. In a first step, the rainfall I_{mean}-D thresholds with the highest performance are presented for the 2009–2020 and 2013–2020 dataset (Fig. 7a and b, respectively). In addition, the rainfall I_{mean}-D threshold selected in this study is compared with a previous threshold defined in Abancó et al. (2016). In the next stage, we compare the best fitting rainfall I_{max_dur} thresholds and hydro-meteorological thresholds considering the VWC at 15 cm depth (Fig. 8a) and the VWC at 30 cm depth (Fig. 8b) for the 2013–2020 dataset.

Regarding the rainfall I_{mean}-D thresholds (Fig. 7) of this study, a high number of FPs events are predicted relative to the number of TPs for both datasets. Nevertheless, the thresholds performance could be increased if the rainfall events with a duration longer than 10 h were not considered, since none of them triggered a torrential flow. Even so, the number of FPs would be higher compared to the rainfall I_{max_dur} and hydro-meteorological thresholds. In any case, the majority of torrential flow events exceed the threshold line and only 5 and 2 debris floods with
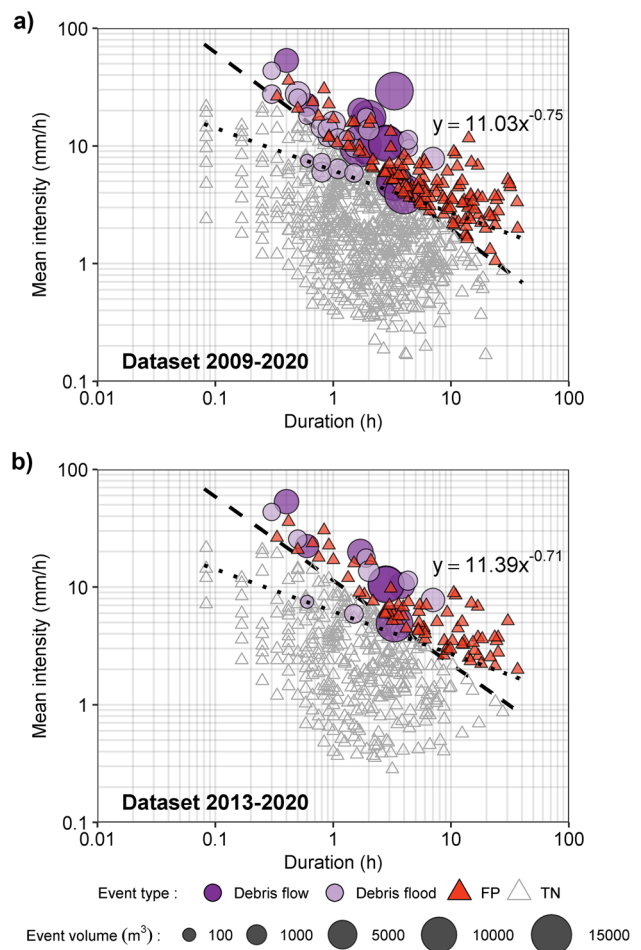

Fig. 7 Mean intensity vs duration (I_{mean}-D) rainfall thresholds regarding the 2009–2020 dataset (**a**) and 2013–2020 dataset (**b**). Dashed line and equation indicate the threshold with best performance defined in this study and selected by maximizing the F_1-score. FP and TN stand for false positive and true negative, respectively, and correspond to the thresholds defined in the present study. Dotted line indicates the previous threshold defined in Abancó et al. (2016) ($y = 6.2x.^{-0.36}$)

small volumes (< 1000 m^3) did not fulfill the threshold condition for the 2009–2020 and 2013–2020 dataset, respectively. Therefore, the results in Fig. 7 demonstrate that, regardless of the longer (2009–2020) or shorter (2013–2020) dataset, similar thresholds equations are obtained for the rainfall I_{mean}-D dataset. Contrarily, the rainfall I_{mean}-D threshold in Abancó et al. (2016) greatly increases the number of false positives since the correct prediction of torrential flows prevails over the false alarms. It is worth to note that very close threshold results were also described in a previous study spanning the period 2009–2018 (Oorthuis et al. 2021) and, therefore, this threshold was not plotted in Fig. 7. Furthermore, the results in Fig. 7 confirm the hypothesis that torrential flows in the Rebaixader catchment are triggered by short lasting rainfalls, generally shorter than 3 h in duration.

In relation to the rainfall I_{max_dur} and hydro-meteorological thresholds (Fig. 8), the results clearly illustrate that, in general,
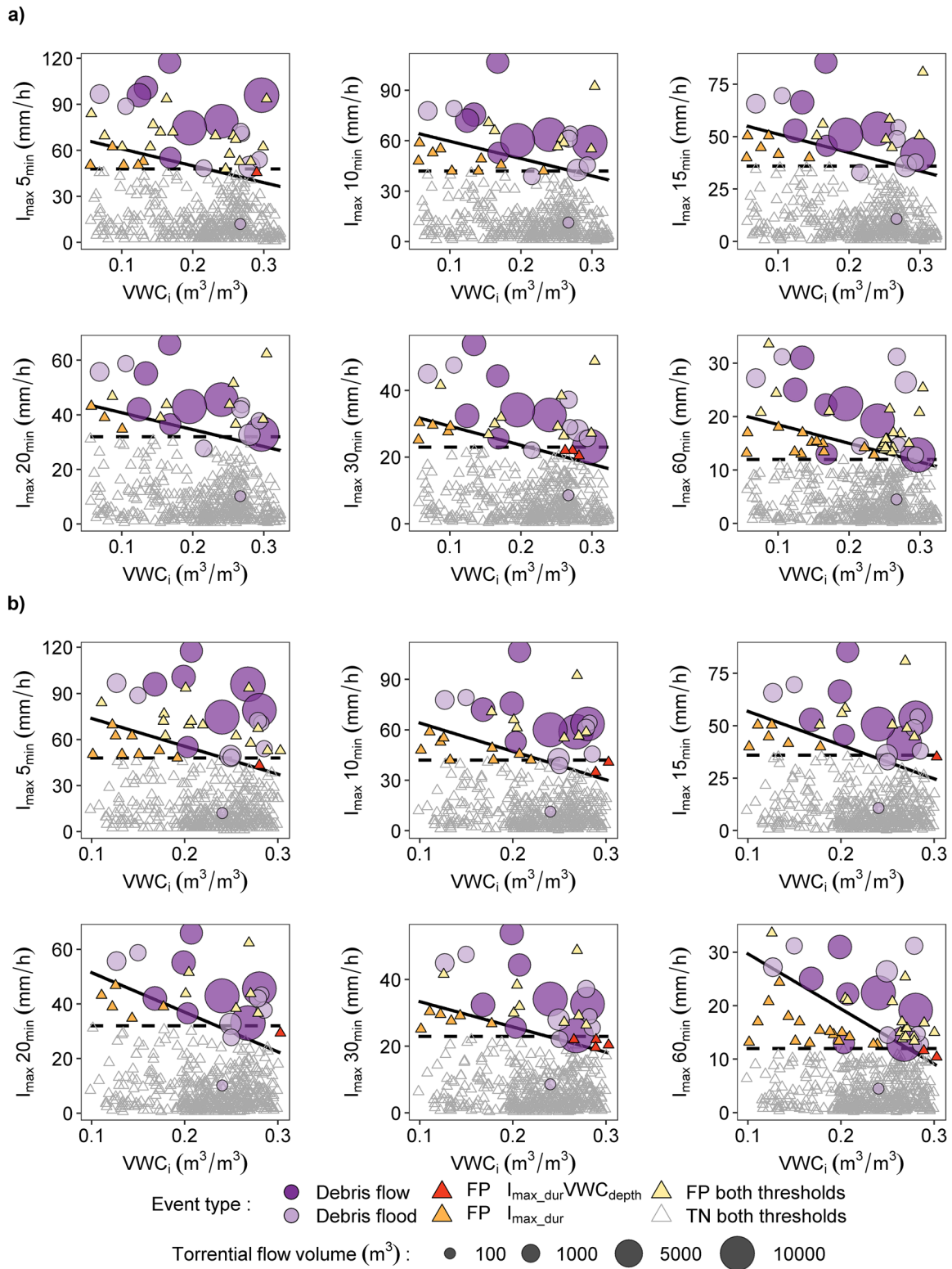
Fig. 8 Comparison of the rainfall I_{max_dur} thresholds (dashed lines) and the hydro-meteorological $I_{max\_durVWCdepth}$ thresholds (solid lines) regarding the VWC prior to the rainfall event (VWC_i) at **a** 15 cm depth and **b** 30 cm depth. FP and TN stand for false positive and true negative, respectively. Thresholds correspond to the ones defined in this study

higher maximum rainfall intensities are necessary for the triggering of torrential flows when compared to the non-triggering rainfalls. Moreover, one may distinguish a slight trend when looking closer at the hydro-meteorological thresholds: generally, higher rainfall intensities are required for the triggering of torrential flows when the soil is initially drier, while lower rainfall intensities are necessary for initially wetter soil conditions. In conclusion, the results in Fig. 8 confirm the hypothesis that including the soil hydrologic conditions (VWC in this study) prior to the triggering rainfall reduces the number of FPs compared to thresholds relying exclusively on rainfall parameters.

Implementation in a landslide early warning system

Landslide early warning systems (LEWSs) are a significant option among the diverse mitigation solutions available to reduce the risk related to landslides (Segoni et al. 2018). In contrast to active measures or structural engineering works, the installation of a LEWS is often a cost-effective and sustainable risk mitigation strategy (Glade and Nadim 2014). LEWSs are used to monitor one or more variable related to landslide triggering in order to generate and disseminate timely and meaningful warning information to enable authorities and communities to act appropriately and in sufficient time to reduce the landslide risk (UN/ISDR 2006). LEWSs can be classified into alarm, warning, and forecasting systems, depending on the lead time between alarm and landslide triggering (Sättele et al. 2012). Warning and forecasting systems monitor the precursors or predisposing factors in order to predict landslide triggering (e.g., rainfall and/or soil hydrologic conditions) and have higher lead times (Sättele et al. 2015). An important part of LEWS is the definition of thresholds for landslide initiation.

Here, we check the implementation of our previous results into a LEWS by means of rainfall and hydro-meteorological thresholds. This is an exercise for research purposes, since there are no significant risks related to torrential activity in the catchment. The
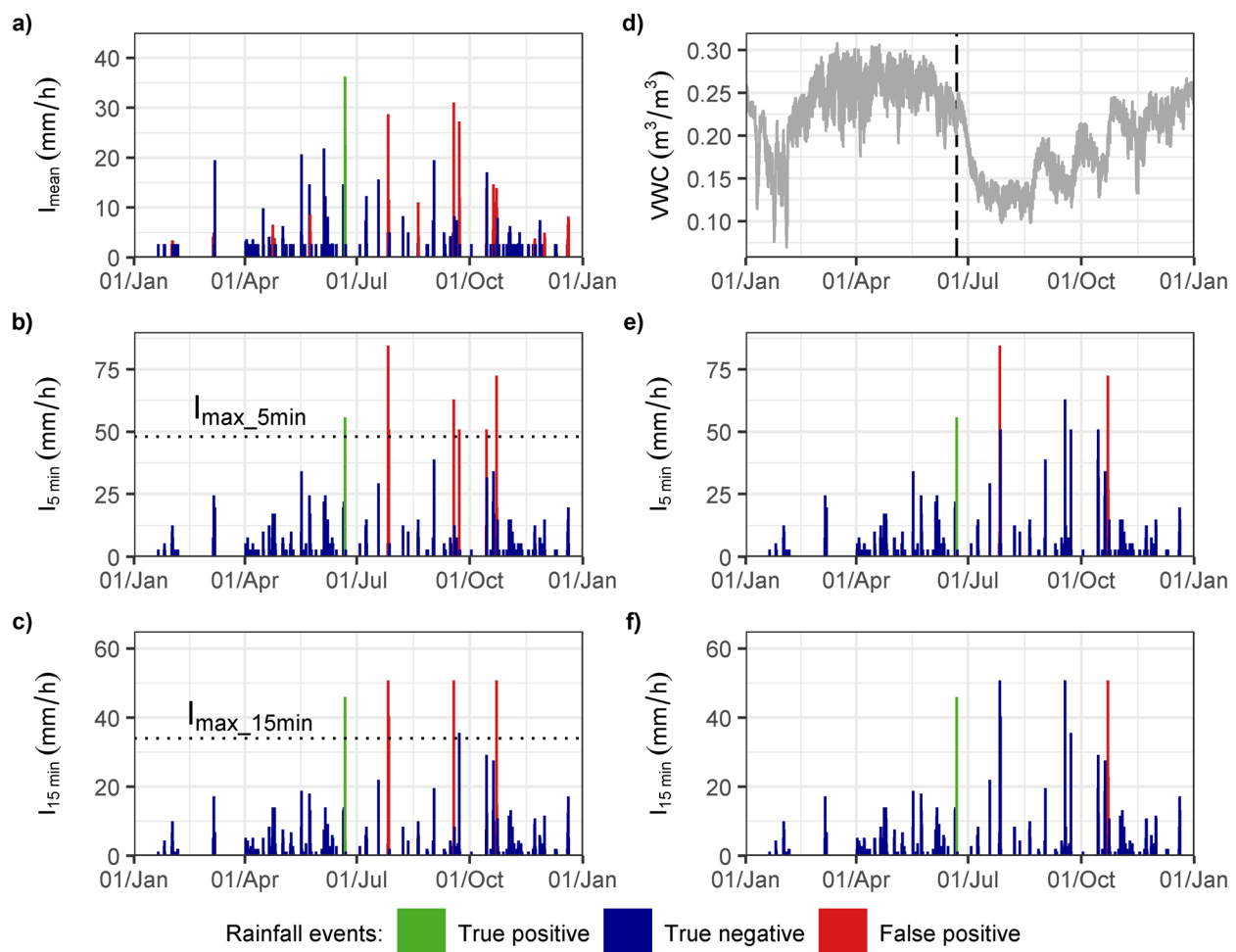

Fig. 9 Time series of rainfall and volumetric water content (VWC) showing landslide occurrence and threshold exceedance in 2019. Predictions of rainfall thresholds in left column and hydro-meteorological thresholds in right column. Rainfall thresholds predictions for the I_{mean}-D threshold (**a**) and maximum rainfall intensity thresholds for 5 min (**b**) and 15 min (**c**) of duration. **d** VWC at INF-SCARP1 infiltration station and at 30 cm depth. Hydro-meteorological threshold predictions combining VWC at 30 cm at INF-SCARP1 and maximum rainfall intensities at 5 min (**e**) and 15 min (**f**) of duration. The horizontal dotted line in **b** and **c** shows the maximum intensities defined in the maximum rainfall intensity thresholds. The vertical dashed line in **d** indicates the timing of torrential flow detection at the FLOW-WR monitoring station. Thresholds correspond to the ones defined in this study

performance of the two thresholds is compared using the year 2019 as example (Fig. 9). Both thresholds were defined considering the 2013–2020 dataset (see Table 4).

Regarding rainfall thresholds, the analysis of the I_{mean}-D threshold predictions (Fig. 9a) reveals a high number of false positives, 16 in total, despite correctly predicting the occurrence of the only torrential flow detected on that year. This trade-off between correct and incorrect predictions limits the success of implementing the rainfall I_{mean}-D threshold in a LEWS, since with a high number of false alarms, the so-called cry-wolf effect may induce populations to ignore the issued alarms (Barnes et al. 2007; Peres et al. 2018). Conversely, the rainfall thresholds defined with maximum intensities of 5 and 15 min duration (Fig. 9b and c) reduce the false positives to 6 and 4, respectively, while torrential flow occurrence is still correctly predicted. Hence, when relying on rainfall measurements or rainfall forecasts, the maximum rainfall intensity threshold should be considered for a LEWS in front of traditional rainfall I_{mean}-D thresholds. Nevertheless, the precision of the predicted alarms is still low which may be a drawback for implementing these thresholds into a LEWS.

Regarding the hydro-meteorological thresholds, the results demonstrate that including VWC measurements improves the predictions compared to rainfall-only thresholds. False positives reduce to 2 and 1 when considering maximum rainfall intensities in 5 and 15 min of duration in combination with the VWC at 30 cm depth (Fig. 9e and f). This reduction in the FPs occurs mainly during the summer and with initially low VWCs, which is especially favorable for the correct prediction of torrential flows in the Rebaixader catchment, since debris flows and floods are mainly triggered between June and September (Oorthuis et al. 2021).

These results encourage further investigation of the role that soil hydrologic conditions play in triggering torrential flows or other slope mass movements and justify the definition of hydro-meteorological thresholds for landslide warning. We suggest considering the hydrological soil conditions in combination with rainfall in other catchments or landslide-prone areas to improve or implement new or existing LEWS.

Conclusions

Our study evaluates the role of hydrologic soil conditions and rainfall characteristics in the triggering of torrential flows at the Rebaixader catchment through monitoring close to the initiation zone of these flows. The results highlight the advances of considering both soil hydrologic conditions and rainfall characteristics for torrential flows initiation. The contributions of combining the hydrologic soil state with rainfall characteristics are compared in terms of predictive capability with rainfall thresholds based solely on the rainfall as triggering factor. As rainfall-based thresholds neglect the antecedent hydrologic soil conditions, their predictive performance is often low and results in many false positives, which may induce to ignore the issued alarms.

Regarding the hydrologic soil measurements (VWC and suction in this study), the results show that soil moisture can range from dry (VWC $\approx$ 0.10 m³/m³ and suction $\geq$ 1000 kPa) to wet (VWC $\approx$ 0.34 m³/m³ and suction $\approx$ 5 kPa) conditions prior to the initiation of triggering rainfalls. Therefore, the complete saturation of the sediment debris layer is not necessary for the initiation of these flows. The analysis of rainfall characteristics demonstrates that short duration (less than 3 h) and high intensity (4–10 mm in 5 min

duration, which corresponds to 48–120 mm/h) summer rainfalls between 12:00 and 18:00 h UTC triggered most of torrential flows. Conversely, long duration (e.g., more than 10 h) rainfalls did not trigger torrential flows. In addition, the three largest debris flows (volume > 9000 m³) were preceded by medium to high VWCs (0.18 to 0.30 at all depths) and were not triggered by the heaviest precipitations in terms of rainfall intensity and total rainfall amount, which suggests that both hydrologic soil conditions and rainfall characteristics affect the triggering of torrential flows.

The implications of rainfall and VWC measurements in the definition of only-rainfall and hydro-meteorological thresholds for torrential flow prediction have been evaluated and the performance of each threshold has been analyzed. With respect to rainfall thresholds based exclusively on precipitation parameters, the maximum rainfall intensity thresholds obtained higher prediction accuracy compared to the traditional rainfall mean intensity—duration (I_{mean}-D) threshold. The rainfall I_{mean}-D threshold is not really the most suitable for the prediction of torrential flows in the Rebaixader catchment due to the elevated number of false positives. This confirms that peak rainfall intensities better represent the triggering rainfall characteristics compared to mean rainfall intensity which clearly diminishes for long duration rainfalls. In addition, the duration for the development of triggering peak rainfall intensities is very similar to the critical duration necessary to reach the critical discharge for torrential flows initiation in other torrential catchments. Finally, the combination of VWC and maximum rainfall intensities in the hydro-meteorological thresholds provides an improved precision compared to only-rainfall thresholds, as important antecedent conditions and rainfall intensity variations or bursts are taken into account. The improved precision of hydro-meteorological thresholds, which are novel in the Rebaixader catchment, justifies the testing of this approach in other torrential catchments or landslide-prone areas where continuous hydrologic soil conditions are being monitored. In addition, for both peak rainfall intensity thresholds and hydro-meteorological thresholds, similar durations for torrential flow triggering were obtained when compared to the critical duration necessary to reach the critical discharge in other torrential catchments.

Our results justify the definition of hydro-meteorological thresholds for application in LEWS including VWC. The hydro-meteorological thresholds confirm that the hydrologic conditions of the soil affect the maximum rainfall intensity necessary for torrential flow triggering. Generally, lower rainfall peak intensities are required for torrential flow triggering when the soil is initially wetter at the start of a rainfall event, and vice versa. This confirms that soil moisture is an important predisposing factor for torrential flows initiation and, therefore, a key parameter in LEWSs for torrential flow warning.

Declarations

Conflict of interest   The authors declare no competing interests.

References

Abancó C, Hürlimann M, Fritschi B et al (2012) Transformation of ground vibration signal for debris-flow monitoring and detection in alarm systems. Sensors 12:4870–4891. https://doi.org/10.3390/s120404870

Abancó C, Hürlimann M, Moya J (2014) Analysis of the ground vibration generated by debris flows and other torrential processes at the Rebaixader monitoring site (Central Pyrenees, Spain). Nat Hazards Earth Syst Sci 14:929–943. https://doi.org/10.5194/nhess-14-929-2014

Abancó C, Hürlimann M, Moya J, Berenguer M (2016) Critical rainfall conditions for the initiation of torrential flows. Results from the Rebaixader catchment (Central Pyrenees). J Hydrol 541:218–229. https://doi.org/10.1016/j.jhydrol.2016.01.019

Barnes LR, Gruntfest EC, Hayden MH et al (2007) False alarms and close calls: a conceptual model of warning accuracy. Weather Forecast 22:1140–1147. https://doi.org/10.1175/WAF1031.1

Bel C, Liébault F, Navratil O et al (2017) Rainfall control of debris-flow triggering in the Réal Torrent, Southern French Prealps. Geomorphology 291:17–32. https://doi.org/10.1016/j.geomorph.2016.04.004

Berger C, McArdell BW, Schlunegger F (2011) Sediment transfer patterns at the Illgraben catchment, Switzerland: implications for the time scales of debris flow activities. Geomorphology 125:421–432. https://doi.org/10.1016/j.geomorph.2010.10.019

Berti M, Genevois R, LaHusen R et al (2000) Debris flow monitoring in the acquabona watershed on the dolomites (Italian alps). Phys Chem Earth, Part B Hydrol Ocean Atmos 25:707–715. https://doi.org/10.1016/S1464-1909(00)00090-3

Berti M, Simoni A (2005) Experimental evidences and numerical modelling of debris flow initiated by channel runoff. Landslides 2:171–182. https://doi.org/10.1007/s10346-005-0062-4

Bogaard T, Greco R (2018) Invited perspectives: hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds. Nat Hazards Earth Syst Sci 18:31–39. https://doi.org/10.5194/nhess-18-31-2018

Bogaard T, Greco R (2016) Landslide hydrology: from hydrology to pore pressure. Wiley Interdiscip Rev Water 3:439–459. https://doi.org/10.1002/wat2.1126

Brayshaw D, Hassan MA (2009) Debris flow initiation and sediment recharge in gullies. Geomorphology 109:122–131. https://doi.org/10.1016/j.geomorph.2009.02.021

Brocca L, Melone F, Moramarco T (2008) On the estimation of antecedent wetness conditions in rainfall–runoff modelling. Hydrol Process 22:629–642. https://doi.org/10.1002/hyp.6629

Caine N (1980) The rainfall intensity: duration control of shallow landslides and debris flows. Geogr Ann Ser a, Phys Geogr 62:23. https://doi.org/10.2307/520449

Chen J, Chen X, Zhao W, You Y (2019) Debris-flow mitigation measures and an application case in a small-scale watershed in China. Debris-Flow Hazards Mitig Mech Monit Model Assess - Proc 7th Int Conf Debris-Flow Hazards Mitig pp. 928–935. https://doi.org/10.25676/11124/173125

Chleborad AF, Baum RL, Godt JW, Powers PS (2008) A prototype system for forecasting landslides in the Seattle, Washington, area. In: Landslides and engineering geology of the Seattle, Washington, Area. Geological Society of America, Seattle, Washington, USA, pp. 103–120. https://doi.org/10.1130/2008.4020(06)

Church M, Miles MJ (1987) Meteorological antecedents to debris flow in southwestern British Columbia; some case studies. pp. 63–80. https://doi.org/10.1130/REG7-p63

Coe JA, Kinner DA, Godt JW (2008) Initiation conditions for debris flows generated by runoff at Chalk Cliffs, central Colorado. Geomorphology 96:270–297. https://doi.org/10.1016/j.geomorph.2007.03.017

Comiti F, Marchi L, Macconi P et al (2014) A new monitoring station for debris flows in the European Alps: first observations in the Gadria basin. Nat Hazards 73:1175–1198. https://doi.org/10.1007/s11069-014-1088-5

Costa A (2014) Resistència d'un sòl glacial: aplicació a l'anàlisi d'estabilitat d'una zona font de corrents d'arrossegalls a Senet, Alta Ribagorça. Minor thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, pp. 81. http://hdl.handle.net/2099.1/21697

Crosta G, Frattini P (2001) Rainfall thresholds for triggering soil slips and debris flow. In: A. Mugnai, F. Guzzetti GR (eds) Proceedings 2nd EGS Plinius Conference on Mediterranean Storms. Siena, Italy, pp. 463–488

D'Odorico P (2005) Potential for landsliding: dependence on hyetograph characteristics. J Geophys Res 110. https://doi.org/10.1029/2004JF000127

Fawcett T (2006) An introduction to ROC analysis. Pattern Recognit Lett 27:861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Glade T, Crozier M, Smith P (2000) Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical "Antecedent Daily Rainfall Model." Pure Appl Geophys 157:1059–1079. https://doi.org/10.1007/s000240050017

Glade T, Nadim F (2014) Early warning systems for natural hazards and risks. Nat Hazards 70:1669–1671. https://doi.org/10.1007/s11069-013-1000-8

Godt JW, Baum RL, Chleborad AF (2006) Rainfall characteristics for shallow landsliding in Seattle, Washington, USA. Earth Surf Process Landforms 31:97–110. https://doi.org/10.1002/esp.1237

Godt JW, Coe JA (2007) Alpine debris flows triggered by a 28 July 1999 thunderstorm in the central Front Range, Colorado. Geomorphology 84:80–97. https://doi.org/10.1016/j.geomorph.2006.07.009

Gregoretti C, Fontana GD (2007) Rainfall threshold for the initiation of debris flows by channel bed failure of the Dolomites. Debris-flow hazards mitigation: mechanics, prediction and assessment. Millpress, Rotterdam, Chengdu, China, pp 11–21

Gregoretti C, Fontana GD (2008) The triggering of debris flow due to channel-bed failure in some alpine headwater basins of the dolomites: analyses of critical runoff. Hydrol Process 22:2248–2263. https://doi.org/10.1002/hyp.6821

Guzzetti F, Peruccacci S, Rossi M, Stark CP (2008) The rainfall intensity–duration control of shallow landslides and debris flows: an update. Landslides 5:3–17. https://doi.org/10.1007/s10346-007-0112-1

Guzzetti F, Peruccacci S, Rossi M, Stark CP (2007) Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorol Atmos Phys 98:239–267. https://doi.org/10.1007/s00703-007-0262-7

Hirschberg J, Badoux A, McArdell BW et al (2021) Evaluating methods for debris-flow prediction based on rainfall in an Alpine catchment.

Nat Hazards Earth Syst Sci 21:2773–2789. https://doi.org/10.5194/nhess-21-2773-2021

Hungr O, Leroueil S, Picarelli L (2014) The Varnes classification of landslide types, an update. Landslides 11:167–194. https://doi.org/10.1007/s10346-013-0436-y

Hürlimann M, Abancó C, Moya J (2010) Debris-flow initiation affected by snowmelt. Case study of the Senet monitoring site, Eastern Pyrenees. In: Proceedings of the Mountain Risks International Conference: Bringing Science to Society. Firenze, Italy, 24–26 November 2010a, pp. 81–86

Hürlimann M, Abancó C, Moya J (2012) Rockfalls detached from a lateral moraine during spring season. 2010 and 2011 events observed at the Rebaixader debris-flow monitoring site (Central Pyrenees, Spain). Landslides pp. 385–393. https://doi.org/10.1007/s10346-011-0314-4

Hürlimann M, Abancó C, Moya J, Vilajosana I (2014) Results and experiences gathered at the Rebaixader debris-flow monitoring site, Central Pyrenees, Spain. Landslides 11:939–953. https://doi.org/10.1007/s10346-013-0452-y

Hürlimann M, Coviello V, Bel C et al (2019a) Debris-flow monitoring and warning: review and examples. Earth-Science Rev 199 10.1016 / j. earscirev .2019a.102981

Hürlimann M, Oorthuis R, Abancó C et al (2019b) Monitoring of rainfall and soil moisture at the Rebaixader catchment (Central Pyrenees). In: Kean JW, Coe JA, Santi PM, Guillen BK (eds) International conference on debris-flow hazards mitigation. "Debris-Flow Hazards Mitigation: Mechanics, Monitoring, Modeling, and Assessment: proceedings of the Seventh International Conference on Debris-Flow Hazards Mitigation. Association of Environmental & Engineering Geologists (AEG), Golden, Colorado, USA, pp. 131–137. https://doi.org/10.25676/11124/173227

Imaizumi F, Sidle RC (2007) Linkage of sediment supply and transport processes in Miyagawa Dam catchment. Japan J Geophys Res 112:F3. https://doi.org/10.1029/2006JF000495

Iverson RM (1997) The physics of debris flows. Rev Geophys 35:245–296. https://doi.org/10.1603/033.046.0206

Iverson RM, Reid ME, LaHusen RG (1997) Debris-flow mobilization from landslides. Annu Rev Earth Planet Sci 25:85–138. https://doi.org/10.1146/annurev.earth.25.1.85

Jakob M, Hungr O (2005) Debris-flow hazards and related phenomena. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 739. https://doi.org/10.1007/b138657

Luna BQ (2015) Infiltración de lluvia en un suelo no saturado. Auscultación y simulación numérica en la Cuenca el Rebaixader (Pirineo Central). MSc. Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, pp. 126.

Marino P, Peres DJ, Cancelliere A et al (2020) Soil moisture information can improve shallow landslide forecasting using the hydrometeorological threshold approach. Landslides 17:2041–2054. https://doi.org/10.1007/s10346-020-01420-8

Mirus BB, Becker RE, Baum RL, Smith JB (2018a) Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning. Landslides 15:1909–1919. https://doi.org/10.1007/s10346-018-0995-z

Mirus BB, Morphew MD, Smith JB (2018b) Developing hydro-meteorological thresholds for shallow landslide initiation and early warning. Water (switzerland) 10:1–19. https://doi.org/10.3390/W10091274

Mostbauer K, Kaitna R, Prenner D, Hrachowitz M (2018) The temporally varying roles of rainfall, snowmelt and soil moisture for debris flow initiation in a snow-dominated system. Hydrol Earth Syst Sci 22:3493–3513. https://doi.org/10.5194/hess-22-3493-2018

Muñoz JA (1992) Evolution of a continental collision belt: ECORS-Pyrenees crustal balanced cross-section. In: McClay KR (ed) Thrust tectonics. Springer Netherlands, Dordrecht, pp. 235–246. https://doi.org/10.1007/978-94-011-3066-0_21

Nimmo JR, Perkins KS, Plampin MR et al (2021) Rapid-response unsaturated zone hydrology: small-scale data, small-scale theory, big problems. Front Earth Sci 9. https://doi.org/10.3389/feart.2021.613564

Ochsner TE, Cosh MH, Cuenca RH et al (2013) State of the art in large-scale soil moisture monitoring. soil sci soc am j 77:1888–1919. https://doi.org/10.2136/sssaj2013.03.0093

Oorthuis R, Hürlimann M, Abancó C et al (2021) Monitoring of rainfall and soil moisture at the Rebaixader catchment (Central Pyrenees). Environ Eng Geosci 27:221–229. https://doi.org/10.2113/EEG-D-20-00012

Pastorello R, D'Agostino V, Hürlimann M (2020) Debris flow triggering characterization through a comparative analysis among different mountain catchments. CATENA 186:104348. https://doi.org/10.1016/j.catena.2019.104348

Peres DJ, Cancelliere A (2014) Derivation and evaluation of landslide-triggering thresholds by a Monte Carlo approach. Hydrol Earth Syst Sci 18:4913–4931. https://doi.org/10.5194/hess-18-4913-2014

Peres DJ, Cancelliere A, Greco R, Bogaard T (2018) Influence of uncertain identification of triggering rainfall on the assessment of landslide early warning thresholds. Nat Hazards Earth Syst Sci 18:633–646. https://doi.org/10.5194/nhess-18-633-2018

Ponziani F, Pandolfo C, Stelluti M et al (2012) Assessment of rainfall thresholds and soil moisture modeling for operational hydrogeological risk prevention in the Umbria region (central Italy). Landslides 9:229–237. https://doi.org/10.1007/s10346-011-0287-3

Rickenmann D (2016) Methods for the quantitative assessment of channel processes in torrents (steep streams), 1st edn. CRC Press, London

Saito T, Rehmsmeier M (2015) The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE 10:e0118432. https://doi.org/10.1371/journal.pone.0118432

Sättele M, Bründl M, Straub D (2012) Classification of warning systems for natural hazards. In: Moormann C, Huber M, Proske D (eds) Proceedings of the 10th International Probabilistic Workshop. Institut für Geotechnik der Universitt Stuttgart, Stuttgart, Germany, pp 257–270

Sättele M, Bründl M, Straub D (2015) Reliability and effectiveness of early warning systems for natural hazards: concept and application to debris flow warning. Reliab Eng Syst Saf 142:192–202. https://doi.org/10.1016/j.ress.2015.05.003

Scheevel CR, Baum RL, Mirus BB, Smith JB (2017) Precipitation thresholds for landslide occurrence near Seattle, Mukilteo, and Everett, Washington: U.S. Geological Survey Open-File Report 2017–1039, 51 pp. https://doi.org/10.3133/ofr20171039

Segoni S, Piciullo L, Gariano SL (2018) Preface: landslide early warning systems: monitoring systems, rainfall thresholds, warning models, performance evaluation and risk perception. Nat Hazards Earth Syst Sci 18:3179–3186. https://doi.org/10.5194/nhess-18-3179-2018

Simoni A, Bernard M, Berti M et al (2020) Runoff-generated debris flows: observation of initiation conditions and erosion–deposition dynamics along the channel at Cancia (eastern Italian Alps). Earth Surf Process Landforms 45:3556–3571. https://doi.org/10.1002/esp.4981

Takahashi T (2014) Debris flow. Mechanics, Prediction and Countermeasures, 2nd edition. Taylor & Francis, London, pp. 572. https://doi.org/10.1201/b16647

Takahashi T (2007) Debris flow. Mechanics, Prediction and Countermeasures, 1st edition. Taylor & Francis, London, pp 448. https://doi.org/10.1201/9780203946282

United Nations International Strategy for Disaster Reduction (UN/ISDR) (2006) Developing early warning systems: a checklist. Third International Conference on Early Warning (EWC III). Bonn, Germany, pp 1–13

Walker JP, Willgoose GR, Kalma JD (2004) In situ measurement of soil moisture: a comparison of techniques. J Hydrol 293:85–99. https://doi.org/10.1016/j.jhydrol.2004.01.008

Weiss GM (2013) Foundations of imbalanced learning. In: Haibo He YM (ed) Imbalanced learning: foundations, algorithms, and applications. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp 13–41. https://doi.org/10.1002/9781118646106.ch2

Wieczorek GF, Glade T (2005) Climatic factors influencing occurrence of debris flows. In: Debris-flow hazards and related phenomena. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 325–362. https://doi.org/10.1007/3-540-27129-5_14

Zhao B, Dai Q, Han D et al (2020) Application of hydrological model simulations in landslide predictions. Landslides 17:877–891. https://doi.org/10.1007/s10346-019-01296-3

Zhao S, Yang Y, Qiu G et al (2010) Remote detection of bare soil moisture using a surface-temperature-based soil evaporation transfer coefficient. Int J Appl Earth Obs Geoinf 12:351–358. https://doi.org/10.1016/j.jag.2010.04.007

Raül Oorthuis (✉) · Marcel Hürlimann · José Moya
BarcelonaTECH, Department of Civil and Environmental Engineering,
Universitat Politècnica de Catalunya, Barcelona, Spain
Email: raul.oorthuis@upc.edu

Jean Vaunat · Antonio Lloret
BarcelonaTECH, International Centre for Numerical Methods
in Engineering, Department of Civiland Environmental Engineering,
Universitat Politècnica de Catalunya, Barcelona, Spain