UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA EXPERIMENTAL YACHAY

Escuela de Ciencias de la Tierra, Energía y Ambiente

TÍTULO: Raifall Thresholds for Landslide Initiation in the Northern Andes of Ecuador

Trabajo de integración curricular presentado como requisito para la obtención del título de Geólogo

Autor:

Luna Celín Leandra Abigail

Sangucho Luje Willian Alberto

Tutor:

PhD. Pineda Ordóñez Luis Eduardo

Urcuquí, Abril 2023

AUTORÍA

Yo, Leandra Abigail Luna Celín, con cédula de identidad 1004201859, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así cómo, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad de el/la autora (a) del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay.

Urcuquí, Abril 2023

Leandra Abigail Luna Celín
CI.1004201859

AUTORÍA

Yo, Willian Alberto Sangucho Luje, con cédula de identidad 1722739065, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así cómo, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad de el/la autora (a) del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay Tech.

Urcuquí, Abril 2023

_____
Willian Alberto Sangucho Luje
CI: 1722739065

AUTORIZACIÓN DE PUBLICACIÓN

Yo, Leandra Abigail Luna Celín, con cédula de identidad 1004201859, cedo a la Universidad de Investigación de Tecnología Experimental Yachay, los derechos de publicación de la presente obra, sin que deba haber un reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita de la Universidad.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Urcuquí, Abril 2023

Leandra Abigail Luna Celín
CI.1004201859

AUTORIZACIÓN DE PUBLICACIÓN

Yo, Willian Alberto Sangucho Luje, con cédula de identidad 1722739065, cedo a la Universidad de Investigación de Tecnología Experimental Yachay, los derechos de publicación de la presente obra, sin que deba haber un reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita de la Universidad.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Urcuquí, Abril 2023

Willian Alberto Sangucho Luje
CI: 1722739065

To my entire family and all my loved ones, especially my parents: Nelson Luna and Martha Celín. For the unconditional support and belief in me during this trajectory. To my brother Santiago Luna, who constantly teaches me never to give up and give the better in all the things, in which I can be involved. To my best friends Karen and Diana, who inspire me to find and fight for the goals that will make me happy. To my beautiful dog Tomas, which makes me happy on the harder days and always gives me unconditional love. Finally, to all of my special friends during my University life for all the good anecdotes and special memories.

Leandra Abigail Luna Celín

For my family, especially for my parents who have always supported me unconditionally, they encourage me to achieve my goals and always try to achieve more than what I have proposed. To my sisters who always try to guide me to make good decisions. And finally for my colleagues who were in my good and bad times.

Willian Alberto Sangucho Luje

Acknowledgments

We want to express our gratitude to Yachay Tech University, especially the School of Earth Sciences, Energy and Environment for allowing us to develop this project.

To all our professors, but in special to Luis Eduardo Pineda Ordónez for sharing with us the knowledge, and supporting us during the development of this graduate project.

Finally, to our classmates for all the good memories during this adventure, for the long nights of study, and also for the support during fieldwork. All the anecdotes and new experiences of the geofamilly will are a souvenir to us for the next new stage.

Leandra Abigail Luna Celín and Willian Alberto Sangucho Luje

Abstract

In this study, 354 shallow landslides were used which occurred in the Andes of northern Ecuador in the provinces of Pichincha, Imbabura, and Carchi (PIC) during a 5-year period (2013-2018). For each of these landslides, the duration (D) and the accumulated precipitation (E) were determined. The precipitation estimates were obtained from GPM satellite grid data and calibrated by means of a cross-correlation analysis to identify time lag and bias. By calculating the relative bias, a factor was obtained to correct the rough data. To calculate the rainfall thresholds an algorithm (CTRL-T) developed to automate their estimation was used, the algorithm classifies and characterizes precipitation events through a set of parameters and relates them to each landslide. By analyzing rainfall time series obtained from the GPM, it was possible to identify two weather seasons (wet and dry). This analysis was carried out for each point in the gridded data since the set of parameters changes according to the weather station. For the calculation of thresholds, the CTRL-T algorithm was used which uses a bootstrapping technique and a power law model to determine the conditions that trigger landslides. Using this algorithm with the adjusted set of parameters, rainfall thresholds were calculated for the entire study area (PIC) and for 27 subcategories determined by political division, climatic seasons, precipitation regions, land cover, soil type, lithology, and hydrographic units. We identify five E-D thresholds groups. The first groups (I and II) require higher rainfall conditions to trigger a shallow landslide. Then, group III needs intermediate rainfall conditions for landslide occurrence. Group IV is considered to demand intermediate rainfall conditions, but with a shorter rainfall duration (75 h). Finally, group V needs smaller rainfall conditions to trigger a landslide, and it covers 71% shallow landslide events. The landslide occurrence in the area of E-D thresholds for this group is the easiest.

Key words: Northern of Ecuador, landslide, E-D rainfall threshold, CTRL-T

Resumen

En este estudio se utilizaron 354 deslizamientos superficiales ocurridos en los Andes del norte de Ecuador en las provincias de Pichincha, Imbabura y Carchi (PIC) durante un período de 5 años (2013-2018). Para cada uno de estos deslizamientos se determinó la duración (D) y la precipitación acumulada (E). Las estimaciones de precipitación se obtuvieron de los datos de la red satelital GPM y se calibraron mediante un análisis de correlación cruzada para identificar el desfase temporal y el sesgo. Al calcular el sesgo relativo, se obtuvo un factor para corregir la sobreestimación de los datos. Para calcular los umbrales de lluvia se utilizó un algoritmo (CTRL-T) desarrollado para automatizar el calculo, el algoritmo clasifica y caracteriza los eventos de precipitación a través de un conjunto de parámetros y los relaciona con cada deslizamiento. Mediante el análisis de las series de tiempo obtenidas del GPM se determinó que existen dos estaciones climáticas (wet and dry), este análisis se realizó para cada punto en los datos grillados ya que el conjunto de parametros cambian según la estación climática. Para el cálculo de umbrales se utilizó el algoritmo CTRL-T que utiliza la técnica de bootstrapping y el modelo de power law para determinar para modelar las condiciones que desencadena los deslizamientos. Usando este algoritmo con los parámetros P ajustados, se calculó los umbrales de lluvia para toda el área de estudio (PIC) y para 27 subcategorías determinadas por la división política, temporadas climáticas, regiones de precipitación, cobertura de suelo, tipo de suelo, litología y cuencas hidrográficas. Identificamos cinco grupos de umbrales de lluvia. Los primeros grupos (I y II) requieren condiciones de mayor precipitación para desencadenar un deslizamiento superficial. Luego, el grupo III necesita condiciones de lluvia intermedias para la ocurrencia de deslizamientos. Se considera que el grupo IV exige condiciones de lluvia intermedias, pero con una duración de lluvia más corta (75 h). Finalmente, el grupo V necesita condiciones de lluvia más pequeñas para desencadenar un deslizamiento de tierra y cubre el 71% de los eventos de deslizamientos superficiales. La ocurrencia de deslizamientos en el área de los umbrales E-D para este grupo es más fácil.

Palabras clave: Norte del Ecuador, Deslizamiento, Umbrales de luvia E-D, CTRL-T

Acronym

AD	alluvial deposit group
C	climatic conditions
CL	conglomerate group
CORRA	Combined Radar-Radiometer Precipitation Algorithms
CTRL-T	Calculation of Thresholds for Rainfall induced Landslides Tool
D	rainfal event Duration
DesInventar	Disaster Inventory System
DPR	Dual-frequency Precipitation Radar
E	cumulated event rainfall
ews	end warm season
FD	flysch deposit group
FONAG	Fondo para la protección del Agua
GPM	Global Precipitation Measurement
IAD	Inter-Andean Depression
IGM	Instituto Geografico Militar
ID	Intensity-Duration
IMERG	Integrated Multi-satellite Retrievals for GPM
INAMHI	Instituto Nacional de Meteorología e Hidrología
ITCZ	Inter Tropical Convergence Zone
MAATE	Ministerio del Ambiente, Agua y Transicion Ecológica
MAG	Ministerio de Agricultura y Ganaderia
MAGAP	Ministerio de Agricultura, Ganadería, Acuacultura y Pesca
MPRC	maximum probability rainfall conditions
MRC	multiple rainfall conditions
MR	metamorphic rocks group
RMSE	root mean square
sws	start warm season
TRMM	Tropical Rainfall Measuring Mission
VO	Volcanics group

Contents

List of Figures

List of Tables

1 | Introduction

In Ecuador, one common geo-hazard is landslides due to the geomorphological and climatic conditions of the country. The consequences of this phenomenon are often casualties and infrastructure losses. According to Disaster Inventory System (DesInventar) in Ecuador, since 1970, around 7717 landslides have been recorded. The cost of human lives was 1642 dead, 12880 affected and 15126 evacuated. While in infrastructure 2456 homes were destroyed, 7851 homes were affected, and 1419 km of roads needed repair, representing a loss of around 900 million dollars. Therefore, it is important to understand the geophysical triggering mechanisms to support disaster management plans and avoid possible damages.

According to the Servicio Nacional de Gestión de Riesgos y Emergencias, in Ecuador, most landslides occur during the rainy seasons (March-April and October-November) when soil mechanical properties change due to the high variability of water content. The main factors that affect the amount of water in the soil are the rainfall intensity and wet antecedent soil moisture conditions since both modify the pore water pressure, which can decrease the soil strength and increase the stress (Ray & Jacobs, 2007).

Despite the complexity of mechanisms underlying landslide formation in Ecuador, few studies have investigated the necessary geophysical conditions that trigger landslide initiation. In the Andes of southern Ecuador, landslide susceptibility to several confounders, including road construction and topographic, climatic, and geological factors, was found to be more than one order of magnitude higher near high ways. This susceptibility persists along the strong climate gradient of the eastern cordillera (Brenning,

1

Schwinn, Ruiz-Páez, & Muenchow, 2015). Also, in the same region, a multi-method study that integrates geophysical and geotechnical methods, mineralogical studies, and analysis of rainfall data on an intra-montane basin (Soto et al., 2017) reported that the wet high-altitude tropical climate induces weak slope instability. Both studies, however, focused on limited areas of southern Ecuador. To our knowledge, nobody had investigated the antecedent rainfall characteristics of landslide initiation over the complex Andean landscape of northern Ecuador. Such a study would shed light on one of the main factors that trigger this geological hazard.

A landslide is a movement of material (rock, soil, and artificial fill) that includes different processes until it produces a slope failure. According to L. Highland (2004), the types of landslides can differentiate based on the kinds of material involved and the mode of movement. It can occur in bedrock or on soils, cultivated land, barren slopes, and natural forests; slope failures can affect dry areas and very humid areas. Moreover, some other landslides include more geophysical variables such as water, air, or ice content in the landslide material. The internal mechanics that describes the mass movement during a landslide is based on different types of movements, namely fall, topple, slide, spread or glow. (L. M. Highland & Bobrowsky, 2008)..

The sharp slopes in mountainous regions are often prone to landslides. But, landslides can occur in areas with a low relief due to cut and fill failures, river bluff failures, lateral spreading landslides, and the collapse of mine-waste piles (especially coal) (L. Highland, 2004). Moreover, the general geological aspects involved in landslides are weak or sensitive, weathered, sheared, jointed, or fissured materials which depend on their discontinuity and orientation (bedding, schistosity, fault, unconformity, contact, and so forth). On the other hand, the causes of landslides are mostly: geological, morphological, and human causes. Nevertheless, the most critical factor producing a landslide is the relationship between landslide and water because an increment in the amount of water might cause slope saturation. The main reason for such an increase might be intense rainfall (L. M. Highland & Bobrowsky, 2008).

The study region involves the Pichincha, Imbabura, and Carchi Provinces in the northern Andes of Ecuador. Physiographically, they are part of the Western Cordillera and lie on the hydrographic basin of the Mira and Esmeraldas Rivers. The study area has approximately 17878 Km^2 and presents a variety of reliefs from abrupt to heald. The area exhibits elevations between 71 and 5880 meters above sea level. This zone has variable slope ranges between 10-70 degrees, with large precipitation gradients. The climate in the study area is diverse, humid in the inter-Andean zone, hot and dry in the inter-Andean valleys, and cold in the high mountain paramos above 3000 m of altitude. This region is influenced by oceanic and Amazonian air masses and the oscillation of the Intertropical Convergence Zone, so it registers a bimodal rain distribution (two rainy seasons) between March-April and October-November (Varela & Ron, 2018). The annual rainfall fluctuates between 800 mm to 1400 mm, and the temperature varies between 8 and 30 degrees centigrade with a gradient of 5 degrees per 1000 m of height (Cedeño & Donoso, 2010).

Studies on landslide triggering are often based on empirical and physical models, where the empirical models are more useful at a regional scale provided a sufficient amount of information is available (Berti et al., 2012). On the other hand, the physical models use thresholds for spatially variable characteristics (slope gradient, soil depth, and soil resistance) and develop a hydrological model to predict the pore pressure with the input rainfall. Moreover, the studies for landslide-triggering are focused on precipitation because there are many empirical and probabilistic approaches to quantify the joint occurrence of rainfall involving landslides (Leonarduzzi, Molnar, & McArdell, 2017). The rainfall Intensity-Duration (ID) threshold curve is considered the most useful approach to quantify landslide triggering rainfall conditions by statistical methods. The Bayesian inference method is useful to determine a minimum threshold with probabilities of landslide occurrences conditional to characteristics of rainfall events (Berti et al., 2012). Furthermore, some studies approach different conditions such as topographic, geological, climatic, and meteorological for the rainfall-induced landslide on the basis of empirical cumulated event rainfall-rainfall duration (Peruccacci et al., 2017). The method for this

study is the rainfall threshold model that relates the cumulated event rainfall with the duration of rainfall that implements a frequentist approach to determine the intercept and the slope of the power-law curve selected to represent the rainfall threshold (Melillo et al., 2018).

In this work, we intend to unveil the rainfall thresholds that trigger landslides in the northern Andes of Ecuador through a statistical analysis that considers precipitation characteristics of different pluviometric regions. In addition, the area provides an interesting study case because recently an upgraded database of landslides is publicly available on the one hand (DesInventar, 2021); and, the use of satellite data to supply ground precipitation scarcity on the other hand. The importance of this work lies in the use of satellite information since this eliminates the errors that could be generated when assigning an event to the nearest rain gauge.

2 | Problem Statement

In Ecuador, landslide early-warning systems has just received attention and is still an area of emerging research for the National Meteorological and Hydrological Services and Civil Protection Agencies. On the way to landslide prediction, targeted to local users, the main challenges are 1) the lack of accurate high-resolution for characterization of land mass movement and 2) the institutional barriers to cope with this lack of information and turn it into readily available products for early warning. To create models that show the characteristics of landslide triggering factors, it is necessary to gather all available information from several sources (ground observations and satellite data). According to L. M. Highland and Bobrowsky (2008), the three major triggering mechanisms that cause landslides are water, seismic activity, and volcanic activity. Several factors like slope, morphology, soil type, underlying geology and weather are widely affected by these mechanisms. According to Rivera (2021), the main triggering mechanism in the study area is soil wetness (pore pressure), which is largely influenced by rainfall conditions. Currently, there is a large database with landslide information available for Ecuador (DesInventar) and a network of weather stations with daily rainfall information country-wide Instituto Nacional de Meteorología e Hidrología (INAMHI). Nevertheless, there are no-studies using this data together to get the antecedent rainfall characteristics that triggered landslides, but small-scale empirical studies in Southern Ecuador (Soto et al., 2017).

Based on Nerini et al. (2015), there are three methodologies to cope with the lack of rainfall information to analyze landslides, the first one uses direct measurements by rain gauges; however, the lack of description of spatial structures in tropical regions is

considered an important limitation for hydrological modeling applications. The second corresponds to indirect satellite measurements, which are considered an alternative source of forcing data because the satellite algorithms perform gauge correction using the mean-field bias correction (MBC). Finally, the combination of both consists of the analysis of signals from daily rain gauge and Multi-satellite Precipitation Analysis (TMPA) time series reconstructing the product and reducing the noise in the precipitation field.

Urgilez, Robles, Bakker, Guzman, and Bogaard (2020), analyzed the landslides of Guarumales located in the Cordillera Real using the meteorological information (rainfall and evaporation data) from the stations located at the south part of the landslide within 2013-2018. However, the limitation was the temporal resolution of geodetical data with just a few measurements per year, which are not enough to determine the groundwater response to rainfall. Therefore, a solution to improve the data on pore water fluctuations was to use the groundwater level data with higher Spatio-temporal accuracy. Also, the real-time kinematic global positioning systems (RTK-dGNSS) acquisitions are useful tools to yield better results.

For this reason, we propose to use satellite imagery and make a reconstruction of rainfall antecedent conditions for each event. In this way, it can be possible to determine the threshold for triggering a landslide. Moreover, through statistical modeling, we aim at determining the following characteristics of precipitation: rainfall duration, frequency, and intensity. Finally, the analysis of these parameters with physiographic aspects such as type of soil, land use, and lithology will yield disaggregated thresholds for different physiographic categories.

3 | Objectives

The objectives of this study are:

- To understand the relationship between rainfall antecedent conditions as triggering mechanisms for landslides in northern Ecuador.

- To estimate rainfall thresholds for landslide initiation as a function of physiographic characteristics (lithology, soil type, land cover, precipitation region, climatic season, and hydrographic units)

4 | Study Area

4.1 Location

The study area embraces the northern Andes of Ecuador, which limits with Colombia, specifically the provinces of Carchi, Imbabura, and Pichincha, where there are large agricultural areas. The landscape hetereogenety and the variability of precipitation in space and time can trigger landslides, risking productive agriculture areas and energy chains. Recently, some events happened on important roads of each province, such as San Lorenzo road on Imbabura, Monte Olivo (Carchi), and General Rumiñahui road on Pichincha.

Moreover, from a geographic point of view, the study area is part of the Inter-Andean Depression (IAD), which is separated by the Cordillera Occidental and Real. The total area between the three provinces corresponds to 17904 km². Another important aspect is rivers because most of the landslides happened nearby the Mira River (the boundary between Imbabura and Carchi) and the San Pedro River (Pichincha) (Figure 4.1).

4.1.1 Characterization of the study area

The characteristics of the study area are divided into morphological aspects, human activity, hydrographic units, and climatic regions. Firstly, the morphological features are important for environmental conditions, specifically information about soil type and lithology. They both give us information about porosity and permeability of soil particles playing an important role in the mobilization of water. On the other hand, the informa-
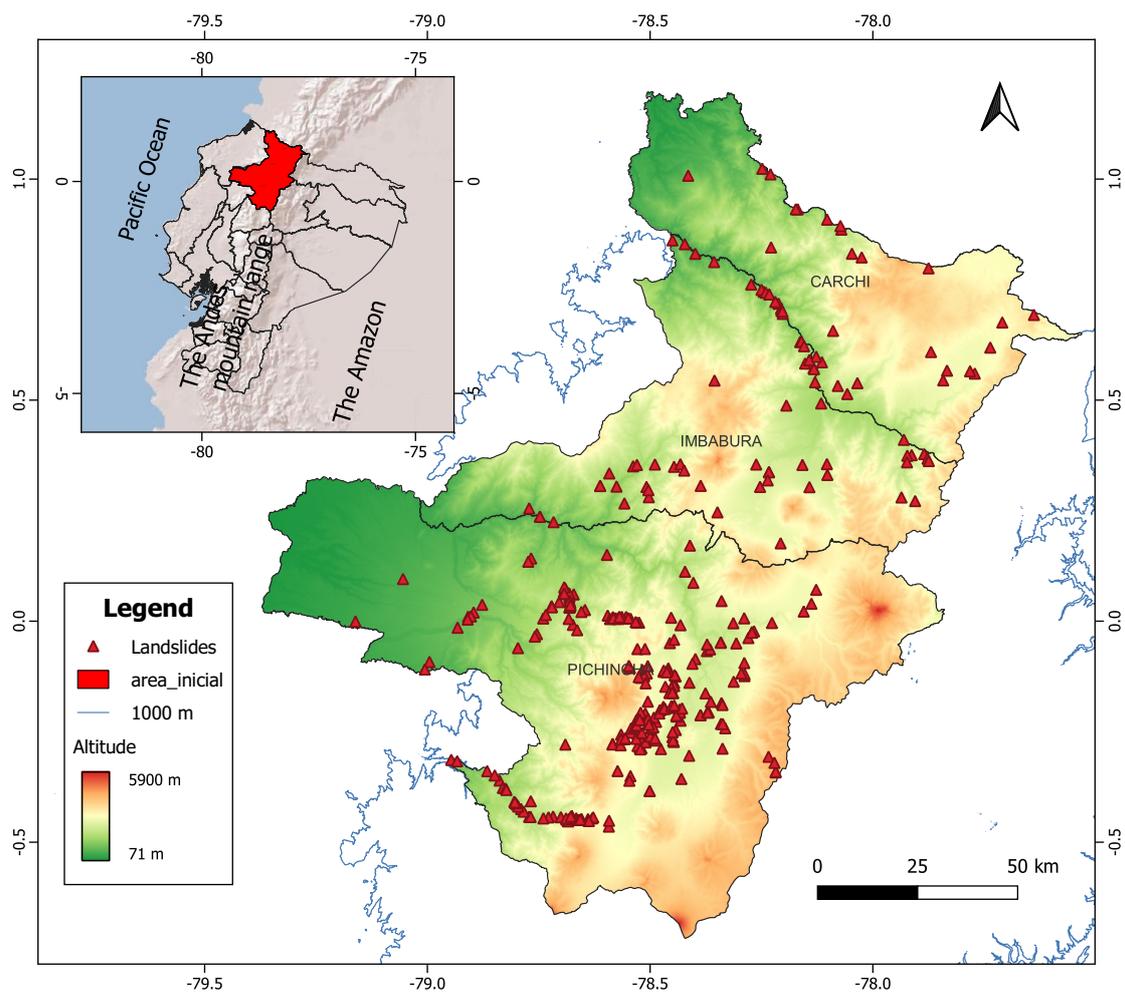
Figure 4.1: Map of the study area (area covered by the red line), with landslides (red triangles) and the main geographical factors that affect climate. The regions Coast, Sierra, and Oriente that are delimited by a contour line of 1000 m. (blue line).

tion related to human activity can influence land cover changes, and all ground alterations can play an important role in triggering a landslide (Brenning et al., 2015). Moreover, to understand the geographic context of the area is necessary to have information about the hydrographic basins. In Ecuador, this information is available from institutions such as the Ministerio de Agricultura y Ganaderia (MAG), Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (MAGAP), Ministerio del Ambiente, Agua y Transicion Ecológica (MAATE), and Geoportal del Instituto Geografico Militar (IGM). All the information used in this study is shown in Table 4.1.

Table 4.1: Cartographic information used in this study

Cartographic information						
Datasets	**Format**	**Year**	**Scale**	**Description**	**Access Date**	**Source**
SHP hydrological units N1-N2	Vector	2008	-	level 3 hydrographic unit	2022 /05 /30	MAATE
SHP hydrological units N3-N6	Vector	2020	-	level 3-6 hydrographic unit	2022 /05 /30	MAATE
SHP hidrogeologic map	Vector	2005	1:100000	Lithology and stratigraphic formations	2022 /05 /30	MAGAP
SHP land cover 2020	Vector	2020	1:25 000	It represents the coverage and Land cover of Continental Ecuador for the year 2020.	2022 /05 /30	MAATE
SHP geopedologic unit 2019	Vector	2019	1:25000	Homogeneous morphologic area, which characterize the soil properties.	2022 /05 /30	MAGAP
Rivers	Vector	2017	1:1000000		2022 /05 /30	Geoportal IGM
Provinces	Vector	2012			2022 /05 /30	INEC
22 Pluviometric Regions	Vector	2022		Precipitation Regions	2022 /05 /30	(Ilbay-Yupa et al., 2021)

4.1.1.1 Soil Taxonomy

The soil taxonomy is important to determine the permeability and the hydraulic properties. This information was obtained from MAGAP. In the study area, 30.51% corresponds to andisols, which from a general point of view are soils generated during a volcanic eruption, where the volcanic material deposited on upper layers cools down fast. For this reason, the crystallization of minerals does not happen, generating vitreous material and amorphous volcanic material (Moreno, Ibáñez, & Gisbert, 2011). In Ecuador, the source for this type of soil is volcanic deposits, specifically volcanic ash, pumice, and lava. Also, the pyroclastic material can be another source for the development of these dark soils, which have a content of amorphous clays and a low density of $< 0.90g/cm^3$ (SIGTIER-RAS, 2017). The coarse texture of the soil depends on whether a volcano is nearby. On the other hand, if the volcano is far, the texture is fine such as silty. Moreover, this type of soil works to develop grass and grow in reliefs with external slopes on the western cordillera. They are well-structured soil having good drainage and moisture retention characteristics. Finally, Janeau, Grellier, and Podwojewski (2015) suggests that the retention of water for

volcanic ash soil capacity varies > 500 g kg^{-1} at 1,500 kPa. There is an area with 24.5 %, corresponding to 87 events of landslides, which are unknown (urban-peri-urban areas).

Then, the mollisols represent 23.73% of the study area, which can grow in many climatic zones from warm to cold. The annual precipitation of the zone can vary from 200 to 800 mm (Ibañez et al., 2011). In Ecuador, the mollisols can be found in the IAD which is filled with volcanic sedimentary deposits. This type of soil is composed of organic matter due to the addition of plant roots, also is important the protolith can be sandstone and calcite. For this reason, organic matter is concentrated as a strong, thick, and dark boundary, and it's the most representative feature of this type of soil (SIGTIERRAS, 2017). According to Ibañez Asensio, Gisbert Blanquer, and Moreno Ramón (2011) in moist areas, the mollisols will generate an albic horizon for the presence of water due to the area's high permeability that allows water mobilization.

The last one is the entisols, which covered 9.3% of the study area. According to SIGTIERRAS (2017), the entisols are located in strong slopes (> 40 to 70%) or in areas susceptible to flooding, also the erosion process will occur faster than in other zones. For this reason, this type of soil is considered of low evolution, however, the crops of cacao and banana can be developed in entisols in the province of Guayas. The grain size of this soil varies from coarse to fine, allowing the mobilization of water, which means there are more porous between the soil particles (Ibáñez Asensio, Gisbert Blanquer, & Moreno Ramón, 2011). The remaining percentage of 11.86% corresponds to Inceptisols, miscellaneous lands, and Anfisols as shown in Table 4.2.

4.1.1.2   Land cover

Land cover is relevant to know the vegetation coverage of the area and determining its characteristics for water storage and retention. The information was obtained from MAATE as shown in Table 4.3, where the selected coverage classification is on the layer "cob lv2". In this way, in populated areas occurred 108 events (30.5%), which includes areas of the main roads. Next, the shrubby vegetation recorded the occurrence of 67 land-

Table 4.2: Number of landslides and percentage by soil type

Soil taxonomy	Count	Percentage (%)
Andisols (A)	108	30.51
Unknown:urban-peri-urban areas (Un)	87	24.58
Mollisols (M)	84	23.73
Entisols (E)	33	9.32
Inceptisols (In)	26	7.34
Mix types: shallow soil-rock outcrop (Mx)	15	4.24
Alfisols (Al)	1	0.28
	354	100

slides (18.9%), this land cover is characterized by a discontinuous cover of shrubs and small trees. It grows between 2000 and 3000 m of elevation (Jorgensen & Leon-Yanez, 1999).

The following information about the features of Land cover categories is from MAG (n.d.). The native forest has 54 events (15.3%), this type of coverage includes all the native trees of different ages and sizes. On the grassland, 40 events occurred (11.3%), which include herbaceous vegetation dominated by species of grasses and loaded legumes, used with fine livestock. The remaining 24% correspond to agricultural mosaic, herbaceous vegetation, crops, forestry plantations, moorland, uncultivated land, natural water bodies, and infrastructure.

4.1.1.3 Lithology

The lithology information was acquired from the MAGAP, where categories were classified into general four groups as shown in Table 4.4. The group recording 267 events of landslides (75.42%) is Volcanics group (VO), which relates to volcanic rock, pyroclastic rock, and lahars. According to Burbano, Becerra, and Pasquel (2015) the study area is characterized by the presence of plioquaternary sediments (volcanic origin), cangahua, lapilli pumice, and volcanic rocks (lavas, breccias, and stuff). Then, the alluvial deposit group (AD) has 45 (12.71%) events of landslides, which include specifically alluvial and colluvial deposits. The undifferentiated terraces contain 19 landslide events.

Table 4.3: Number of landslides and percentage by Land cover level 2

Land cover		
cover	**count**	**Percentage (%)**
Populated area (Pa)	108	30.5
Shrubby vegetation (Sv)	67	18.9
Native forest (Nf)	54	15.3
Grassland (G)	40	11.3
Agricultural mosaic (Am)	30	8.5
Herbaceous vegetation (Hv)	20	5.6
Crops (C)	17	4.8
Forestry plantations (Fp)	6	1.7
Moorland (M)	4	1.1
Uncultivated land (Ul)	4	1.1
Natural water body (Nwb)	2	0.6
Infrastructure (In)	2	0.6
	354	100.0%

Moreover, it constitutes a primary porosity or inter-granular, with medium to low permeability. The conglomerate group (CL) has 16 events (4.52%), which corresponds to sandstone, shale, limestone, conglomerate, dacite, and phyllite. In the zone of Imbabura, the lithological units are composed of no consolidated rocks which are permeable with inter-granular porosity (Burbano et al., 2015). Moreover, at the time of intense precipitation, these deposits can lose slope accumulations of fine-grained debris (Rivera, 2021). The metamorphic rocks group (MR), schists, amphibolites, quartzites, and quartz involve 4 landslides representing 1.13%. Finally, the flysch deposit group (FD) is a compound of low-grade metamorphic rock, sandstone, clay, and chaotic terrain; which has 3 landslides corresponding to 0.85%.

4.1.1.4 Hydrographic Units

The information related to hydrological basins of levels one to three were obtained from MAATE as shown in Table 4.5. Based on this information there are 353 events (99.7%) of landslides corresponding to the following basins; Hydrographic Region 1 (basin of level one) and Hydrographic Unit 15 (basin of level two). The remaining percentage (0.3%)

Table 4.4: Number of landslides and percentage by Lithology; VO: volcanic rock; AD: alluvial and colluvial deposits; TI: undifferentiated terraces; CL: conglomerates; FD: flysch deposit; MR: metamorphic rock, schists.

Lithology	count	Percentage (%)
VO	267	75.42
AD	45	12.71
TI	19	5.37
CL	16	4.52
MR	4	1.13
FD	3	0.85
	354	100

for each type of basin is for the Rio Amazonas Basin (Basin level 1) and Hydrographic Unit 49 (basin of level 2).

Moreover, the information about the basins of level three is relevant to understanding more specifically the geographic features of each area. In Esmeraldas´s River Basin occurred 281 events (79.4%), and it covers a total area of 12 450.7 km^2 where precipitation has a value of 2338 mm year^{-}1, and the volume of the total hydric resources (including superficial water and groundwater) is 21046 hm ^{3}. Here the runoff is concentrated from January to June representing 63.5% - 80.1% of the annual runoff (CISPDR, 2016a).

Finally, the last basin of level three is Mira River Basin with 71 events (20.1%). It covers an area of 6 537.5 km ^{2} where precipitation varies to 1484mm year^{-}1, and the total volume of the hydric resources is 7417 hm ^{3} (represents 2.0% of the country). Here the concentration of runoff varies depending on the location; in the northern and eastern parts runoff concentrates within December and May. In the southwest occurs from January to June, and in the middle regions happens from March to August (CISPDR, 2016b). The remaining percentage (0.6%) corresponds to the following basins of level three: Carchi River Basin and Hydrographic Unit 497.

Table 4.5: Number of landslides and percentage by Hydrographic Units levels 1-3

Hydrographic Units		
Basin level 1	**Count**	**Percentage (%)**
Hydrographic Region 1 (HR1)	353	99.7
Amazonas River Basin (ArB)	1	0.3
Basin level 2	**Count**	**Percentage (%)**
Hydrographic Unit 15 (HU15)	353	99.7
Hydrographic Unit 49 (HU49)	1	0.3
Basin level 3	**Count**	**Percentage (%)**
Esmeraldas River Basin (ErB)	281	79.4
Mira River Basin (MrB)	71	20.1
Carchi River Basin (CrB)	1	0.3
Hydrographic Unit 497 (HU497)	1	0.3

4.2 Geology overview

Ecuador is composed of six geological regions from east to west, and several fault systems affect them (Tamay, Galindo-Zaldívar, Martos, & Soto, 2018). The first region is the Oriente Basin, composed of Paleozoic marine sediments, Triassic-Jurassic marine, continental rift deposits, and Late Jurassic volcanoclastic sediments (Jaillard et al., 1997). Then, there is the Sub-Andean zone, which is characterized by a series of sedimentary, plutonic, and volcanic rocks formed during the Jurassic to Oligocene-early Miocene (Aspden & Litherland, 1992).

The third region is the Cordillera Real compound by several Paleozoic to Cretaceous metamorphic and volcanic terranes (Hungerbühler et al., 2002). Aspden and Litherland (1992) consider that from south to north the Cordillera Real is divided into three areas with distinctive types of rock, at the southern part is well-exposed and comprises granitic gneiss, cordierite gneiss, amphibolite, schist, phyllite, and quartzite. In the central part, the area is buried by volcanic deposits, but the presence of schist, quartz-feldspathic, andalusite gneiss, and amphibolites is higher in the west of Cuenca. Finally, in the northern part specifically in Pichincha, there are cordierite gneiss xenoliths suggesting that this

terrane could be part of the IAD.

The next geologcal region is IAD considered as a whole block compound of Pliocene to Quaternary volcano-sedimentary deposits (Tamay et al., 2018). Then, in the Cordillera Occidental, there is an oceanic plateau due to the accretion of oceanic terranes associated with a variety of sediments, which took place in the Campanian, Late Paleocene, or Eocene (Jaillard et al., 2004). Finally, in the coastal region, there are shallow marine sediments deposited during Oligocene-Miocene, also during the Miocene-Quaternary the flood plains and alluvial deposits were transported (Jaillard et al., 1997).

In addition, during the late Miocene, the formation of sedimentary basins took place between the Cordillera Real and Occidental. From north to south there are four important sub-basins recognized, the first one is the Chota basin located in the northern IAD between the cities of Ibarra and Tulcán; the second corresponds to the Quito - San Antonio - Guayllabamba basin; the third is the Ambato - Latacunga basin and finally the Riobamba-Alusí basin in the extreme south (Winkler, Villagómez, Spikings, Abegglen, & Egüez, 2005).

4.2.1 Geodynamics

In Ecuador, the Andes Cordillera is extended north to south and the origin is due to the subduction of the Nazca Plate eastward at 6 cm/yr relative to the South America Plate, while the Caribbean plate moves 1-2 cm/yr to the east to the southeast (Tamay et al., 2018). During the Quaternary, the volcanic activity in the northern zone of Ecuador increased due to the subduction of plate tectonics. According to Fairbridge (2006), the Dolores - Guayaquil Megashear (DGM) going northeast to southwest is a strike-slip fault system oblique to the continental margin of Ecuador. Moreover, most of the accretion events took place during the Campanian, the Late Paleocene, and the Eocene (Jaillard et al., 2004).

The Inter-Andean Depression (IAD) is located between the Calacalí - Pallatanga - Palenque Fault (CPF) in the western part and the Peltetec Fault (PF) in the eastern part,

also there is intensive volcanic activity and is affected by an active fault system (Villagómez, Egüez, Winkler, & Spikings, 2002). The eastern limit of PF may have formed during the accretion of the oceanic Pallatanga terrane and separates the IAD which is formed during the accretion of the terranes of the Cordillera Real (Winkler et al., 2005). On the other hand, the CCP limited the Cordillera Occidental and IAD until El Oro Provinces, because they disappear and are replaced by east-west-striking metamorphic rocks (Aspden & Litherland, 1992).

4.2.2 Lithotectonic Terranes

The major lithotectonic terranes of Ecuador may be compound of two different sources: (1) continental and (2) remnants of oceanic arcs (Aspden & Litherland, 1992; Litherland, Aspden, & Jemielita, 1994). In the study area, there are the following lithotectonic terranes, except for the Amotape terrane which is located in the southern part of Ecuador. From east to west, the first terrane is Salado limited by Llanganates and Cosanga-Mendez faults, it is composed of plutonic, mafic to intermediate volcanic rocks, and volcano-sedimentary rocks from the Jurassic (Noble, Aspden, & Jemielita, 1997).

Then, the Alao terrane consists of Jurassic metamorphosed volcanic sedimentary rocks and it is limited by the Baños front and the Peltetec fault (Aspden & Litherland, 1992). The Guamote terrane is located between the Chaucha and Alao terrane, also it is limited by the Peltetec fault, the composition consists of metamorphosed Jurassic to Lower Cretaceous quartzites and slates (Litherland et al., 1994).

Moreover, the oldest lithologies are in the Loja, Chaucha, and Amotape terranes which contain Paleozoic sedimentary rocks, and metamorphosed mafic igneous rocks mostly from Devonian, and Triassic granitoid (Noble et al., 1997). Moreover, the Loja terrane is limited by the Baños front and Llanganates Fault, where there are semi-pelitic rocks and meta-granitoid from the subdivision of Tres Lagunas (Litherland et al., 1994). Finally, the Piñon terrane is located in the western part of Ecuador and is composed of pillow basalts, mafic rocks, and intrusive rocks of the Macuchi terrane of the western

cordillera (Noble et al., 1997).

4.2.3   Carchi Geology

Falsini (1995) suggest that in Carchi province the accumulation of plateau happened inside the valleys, in which two alluvial terraces were recognized in the progressive elevations on the valley floor. Moreover, the accumulation of a thick sequence of sediments can be due to the activity of strato-volcanoes.

The deposits that form this plateau are pyroclastics, debris flows, breccias, agglomerates, and lava flows, where the basal part lies conformably over these sediments, and the elevation of these deposits can vary due to the activity of recent normal faults (Ficcarelli et al., 1997).

The Cangahua formation is a compound of a buried series of eolian sands, tephra layers, and paleosoils with a thickness over 100m (Sauer, 1965). According to Ficcarelli et al. (1997) the rubified argillic horizon is located at the top of the sequence and truncated by an unconformity, it represents the most developed paleosoils, which could correspond to the last Interglaciation.

Finally, the sequence of the deeper part of the Apaquí river is composed of conglomerates with thin layers of sand and pyroclast. Then, there is a layer of lava flows intercalated with breccias and agglomerates. In the upper part of the sequence, there are sandy layers, lahars, and mudflow (Coltorti & Ollier, 2000).

4.2.4   Imbabura Geology

According to Spikings, Seward, Winkler, and Ruiz (2000), the formation of the basement in northern Ecuador took place because Macuchi Terrane is sutured against the Chaucha Terrane. Nowadays the basement is saturated towards the Cordillera Real. This basement is considered the Macuchi Formation, it is located in the western sector of Imbabura, and is compounded by Cretaceous altered metapelites and metabasalts intercalated with volcanoclastic sediments and green andesite lavas.

Another important formation is the Silante, Hughes and Pilatasig (2002) suggests that it is characterized by the presence of allochthonous terranes, including ophiolitic and oceanic crustal fragments. In the southern part of Otavalo of Imbabura province, there is evidence of those events, there is a plateau deeply dissected and faulted (Coltorti & Ollier, 2000).

Moreover, based on geologic maps from the Geological and Energy Research Institute (IIGE) (1979) the most relevant formations are the following: Chota, the Angochagua, Yanahurco, Cotacachi, Negro Puño, and Imbabura volcanic, alluvial and colluvial deposits.

The Chota Unit is a compound of sedimentary and volcanic rocks, the main sedimentary sequences occur during the Miocene and Pleistocene epoch. These sequences have intercalations of volcanic conglomerates and breccias with volcanic sandstones and very fine limestones (Winkler et al., 2005). The deposits located in the center of Imbabura are from Miocene to Holocene and the geologic formations are Angochagua, Yanahurco, Cotacachi, Negro Puño, and Imbabura volcanic. Finally, the alluvial and colluvial deposits during the beginning of the Holocene, and are located within the boundaries of the main rivers: Mira, Lita, and San Juan.

4.2.5   Pichincha Geology

In Pichincha Province there are two main volcanic sequences, the older corresponds to the Rucu Pichincha volcano, and the younger from the Guagua Pichincha volcano (Samaniego, Robin, Chazot, Bourdon, & Cotten, 2010). Moreover, Hughes and Pilatasig (2002) identifies metamorphic xenoliths in lavas from the Pichincha volcano near Quito, suggesting the presence of metamorphic rocks. Moreover, north of Quito on the road to Cayambe, there are a series of small-scale folds, faults, and overthrusts. These sediments take place due to the emplacement of a slump deposit during Early Pliocene times (Coltorti & Ollier, 2000).

In this province, there is a geologic feature recognized as an active fault system,

where there are from north to south ridges: Calderón-Catequilla, Batán-La Bota, and Ilumbisí-Puengasí. These ridges divided three younger sub-basins in the central IAD (Quito, San Antonio, and Guayllabamba). During the Miocene, the first opening occurred in the northern part, more specifically in the Chota Basin. The central IAD may happen during the late Pliocene, and this opening was driven by strike-slip movements along the Calacali-Pallatanga Fault.

Due to these events, the sequence of this area is divided into two: the lower sequence, and the upper sequence. In the lower sequence, there is a Pisque formation composed of andesitic lavas and breccias. Another formation part of this sequence is San Miguel, which is a compound in general by lacustrine deposits. In the upper sequence, there are five formations. The first is the Guayllabamba formation compound by volcanic and lahar deposits in the southwest, more to the west this formation consists of alluvial fan systems. The next formation is the Chiche characterized by lacustrine and fluvial deposits. Then, the Machángara formation consists of primary volcanic deposits and lahars, which were recognized as two facies. The Mojanda Formation is located more to the north and it is the main compound with volcanic and volcanoclastic deposits from the Mojanda volcanic complex. Finally, the last formation in the upper sequence is the Cangahua during the Late Pleistocene, which is distributed in the entire IAD and considered as the younger sediments (Villagómez et al., 2002).

4.3   Climate Setting

The climate of Ecuador is mainly influenced by its geographical location (Figure 4.1), the presence of the Andes mountain range, the Amazon, and the Pacific Ocean determine different climatic regimes: Coast, Sierra, and Oriente (Cedeño & Donoso, 2010). According to Markgraf and Hastenrath (1982) in Ecuador, the weather of coastal areas and lower western slopes of the Andes are affected by air masses from the Pacific Ocean, while in the eastern part of the country the main factor is the moisture-bearing easterly trade winds

originating over the Amazon basin and tropical Atlantic.

In addition, an important phenomenon known as El Niño-Southern Oscillation (ENSO) affects the precipitation in the coastal areas, which has two phases: (1) the warm phase (El Niño) produces often torrential rains, high river runoff, and flooding due to strong positive sea surface temperature; (2) cold phase (La Niña) generates droughts (Vuille, Bradley, & Keimig, 2000).

Morán-Tejeda et al. (2016) suggest that the rainy season in Ecuador is influenced by the Inter-tropical Convergence Zone (ITCZ) and the equatorial front that bring air masses originating from the northwest to cover the coastal region and bring significant rainfall and raise air temperatures. On the other hand, the retreat of the ITCZ and the equatorial front result in the presence of cooler and dryer air masses descending from up-welling regions in the southwest, this generates the dry seasons from June to November. Therefore, the precipitation is not evenly distributed during the year, but 80% is considered a wet season from December to May.

The Inter-Andean Depression has two rainy seasons which are influenced by oceanic and continental air masses. The first one from February to May, and the second one from October to November. Moreover, the precipitation varies between 800 and 1500 mm yr^{-1} due to a loss of humidity in the air masses on both flanks of the Andes mountain. On the other hand, the dry period is from June to September, which is more drastic than the second period during December (Vuille et al., 2000). For this reason, with a heuristic approach in the study area, it is considered that the climate has two periods, the warm period from June to September and the rainy period from October to May.

4.3.1 Precipitation

The regionalization of precipitation is important to identify physical processes in each region; in this way it is possible to understand many aspects of the ecosystem, its development, and vulnerability (Badr, Zaitchik, & Dezfuli, 2015). In Ecuador, a recent study by Tobar and Wyseure (2018) classified the rain gauges into four regions: strong seasonal

Coastal, moderate seasonal Coast, Amazon with mild seasonality, and Sierra (Andes) with moderate seasonality.

In the Sierra region the average annual precipitation varies from 800 to 1500 mm with an exception in the Chota valley (300mm) and Jubones Valley (400mm), this zone is constantly influenced by the Intertropical Convergence Zone. On the other hand, the temperature has a strong relationship with height, between 1500 to 3000 m the temperature varies from 8 to 20 ° C. The temperature gradient is 5 ° C per 1000 m of height (Varela & Ron, 2018). Another relevant factor is solar radiation because the moist air masses rise and generate rain as they cool. As a result of that, the precipitation increased drastically, and the atmospheric pressure down.

There are several sources of information for the precipitation variable, these can be direct or indirect. Direct measurements refer to observations taken in the field, either by automatic sensors or by qualified personnel. While on the other hand, indirect measurements are obtained through remote sensors, which are installed in airplanes or satellites. In the case of Ecuador, both sources of information are available. Direct measurements are made using rain gauges that can be automatic or manual. All this information is provided by INAMHI, which has 761 stations distributed throughout the Ecuadorian territory. There are other data sources available from former projects such as CEDIA which are in charge of evaluating the annual records of each meteorological station for later use. In the INAMHI database, there are 30 years of information, from 1985 to 2015 corresponding to 24-hour totals.

On the other hand, indirect measurements are carried out using sensors installed on satellites. For this case study, one of the products of the GMP mission was used (Integrated Multi-satellite Retrievals for GPM (IMERG)). This project provides rain and snow observations based on an international satellite network. This project is the successor to the Tropical Rainfall Measuring Mission (TRMM) project and its main objective is to unify the information obtained from various satellites. The IMERG provides three runs (Global Precipitation Measurement (GPM) products) to accommodate different user

requirements of delay and accuracy, including IMERG Early with 4 h latency, IMERG Late with 14 h latency, and IMERG Final with 3.5 months (Li et al., 2021).

4.3.1.1 Precipitation regions

The regionalization of precipitation is relevant to identifying the physical processes responsible for the spatio-temporal variability in each region (Ballari, Giraldo, Campozano, & Samaniego, 2018; Ilbay-Yupa et al., 2021). Also, it's important for the development of risk-management plans and for hydrological models to understand the water availability of the area (Ilbay-Yupa et al., 2021).

For this study, the regionalization of precipitation from Ilbay-Yupa et al. (2021) will be used to know the start and the end of the warm season (season with the shortest rainfall period) for each time series in different zones. The regions have been classified based on the homogenization of precipitation for each natural region of Ecuador. In this way, the regions for our study area are the Pacific coast (the regions R6 and R7) and the Inter-Andean region (regions R8, R12, R16, and R17) as shown in Table 4.6.

The region R6 with 39 events of landslides is located in the coastal region, which indicates a marked seasonality of precipitation from January to April, the range of precipitation (2100-3300 mm year^{-1}) decreases with latitude from regions located toward the Pacific coast. Region R7 with 4 events is located more to the northern coast, it's characterized by abundant precipitation (2117 mm year^{-1}) throughout the year, but the period of higher concentration is from January to June.

For the Inter-Andean region, most of the landslides (195 events) occur in region R8, this region is located in the northwestern foothills of the mountain range, with an average annual rainfall of 1406 mm. Region R12 is located in the highest areas of the Andes range, where 15 events occurred, with an average annual rainfall of 1064 mm year^{-1}. Region R16 (64 events) and region R17 (37 events) are located in the inter-Andean valleys on the eastern and western slopes of the Andes mountain range, in central and northern Ecuador. The precipitation decreases from north to east to west and from south. It is influenced by

orographic rain and local warming.

Table 4.6: Number of landslides and percentage by Precipitation regions

Precipitation regions		
Region	**Count**	**%**
R8 (950-2000 mm $year^{-1}$)	195	55.1%
R16 (500-1000 mm $year^{-1}$)	64	18.1%
R6 (2100-3300 mm $year^{-1}$)	39	11.0%
R17 (450-1100 mm $year^{-1}$)	37	10.5%
R12 (550-1400 mm $year^{-1}$)	15	4.2%
R7 (1100-3250 mm $year^{-1}$)	4	1.1%

4.4 Landslides

Ecuador has had a record of 7717 landslides since 1970, of which 10.25% are in the Coastal region, 66.75 % in the Sierra region, and 23 % in the Amazon region. Our study area has a record of 936 landslides, but only 354 have enough information.

A landslide is the result of the complex interaction of several factors that cause a slope failure and due to it, the downhill movement of earth, rock, and organic material occurs. According to L. M. Highland and Bobrowsky (2008), there are two main categories of causes of landslides: the first is natural causes that are subdivided into three triggering mechanisms: water, earthquakes, and volcanic events, and the second one is man-driven landslides.

It is important to note that these factors can also be separated as factors that predispose an event (long time) and factors that trigger the event (short time) and non-apparent trigger factors (medium time). The predisposing factor is the zone characteristics like slope, morphology, soil type, and underlying geology. The triggering factor is an external stimulus such as intense rainfall, earthquakes, volcanic eruptions, and storm waves. Finally, the non-apparent trigger factors are the combinations of several causes such as chemical or physical weathering that steadily produce a landslide without an apparent trigger(Wieczorek, 1996).

4.4.1 Triggering Mechanisms

A triggering mechanism is a description process of the interactions of several factors that end in a sliding event. As mentioned above, landslides have several causes but just one triggering that causes a stress increment or a strength reduction in slope material. There are three natural triggering mechanisms: water (intense rainfall, rapid snow-melt, water level change), volcanic eruptions, and Earthquake shaking. Understanding the triggering mechanisms is necessary to create a threshold model that allows us to understand and predict possible events.

4.4.1.1 Water

The primary cause of landslides is the saturation of soil by water that produces an increment in the pore water pressure, the water sources could be intense rainfall, snow-melt, changes in groundwater levels, and changes in superficial water (L. M. Highland & Bobrowsky, 2008). According to several studies, Peruccacci et al. (2017); Rivera (2021), and others, there is a clear relationship between rainfall intensity and the triggering of landslides. This is also evident if we analyze most of the well-documented sliding events, most of them happen after an intense rainfall for short periods (hours) or after a moderate rainfall for long periods (days) (Wieczorek, 1996). The rapid infiltration of water in soil causes an increment in pore-water pressure as a consequence the shear resistance decreases, and adding the predisposing factors, causes the sliding event (Zaruba & Mencl, 2014).

4.4.1.2 Seismic Activity

Tremors produced by seismic activity affect the equilibrium of slopes and cause changes in stresses due to the oscillation of different frequencies. An important factor is the susceptibility of the soil, for instance in loose sand the vibration produces a change in the intergranular bonds and reduces the cohesion (Zaruba & Mencl, 2014). The dilatation

of the material allows a rapid infiltration of water that saturates the soil, in the water-saturated sand the displacement or rotation of grains reduces the load capacity of the layers and ends as a liquefaction event L. M. Highland and Bobrowsky (2008). Thus, water also plays an important role in landslides triggered by earthquakes. The interaction of these two mechanisms is complex because both act as a trigger or precondition of a landslide.

4.4.1.3 Volcanic Activity

The volcanic activity works as a triggering mechanism in different ways, causing a rapid snow melting, which can form a deluge of rock, soil, ash, and water that accelerates rapidly on the steep slopes of volcanoes. Volcanic edifices that are young and not well consolidated, are weak structures that can collapse easily causing rock slides, landslides, and debris avalanches, for instance, one extreme example is volcanic islands that periodically experiment failure of their structures. Huge masses of soil and rock slide into the ocean or other water bodies, which create massive sub-marine landslides (L. M. Highland & Bobrowsky, 2008).

5 | Data and Methodology

Landslides are events that could be triggered by several factors like tremors, intense rainfall, and others. Based on previous studies in the study area, precipitation is the variable that best explain the triggering of landslides (Rivera, 2021).

To identify the relation between rainfall and triggering landslides, we use the Calculation of Thresholds for Rainfall induced Landslides Tool (CTRL-T) algorithm by Melillo et al. (2017). This algorithm automatizes a systematic analysis that relates the rainfall parameters with the initiation of a landslide. For our study, the workflow is divided into 6 phases (Figure 5.1):

1. **Gathering data:** In this phase cartographic information, landslide data DesInventar, satellite, and field precipitation data were obtained.

2. **Data preparation:** Here the database was built up from a combination of landslides data, type of soil, Land cover, lithology, hydro-graphic units, and precipitation regions. It was linked to the download files through a code in R to obtain the rainfall time series.

3. **Landslide and Rain data analysis:** In this phase, all the attributes were analyzed to get the percentage of each feature relating to the events of landslides. Also, quality control is carried out on the time series and starts with the analysis proposed by Melillo et al. (2018). We use the CTRL-T code that requires specific format input files that were prepared in this phase (Melillo et al., 2017).

4. **Rainfall data processing:** Here we start with Block 1 of the CTRL-T code, which
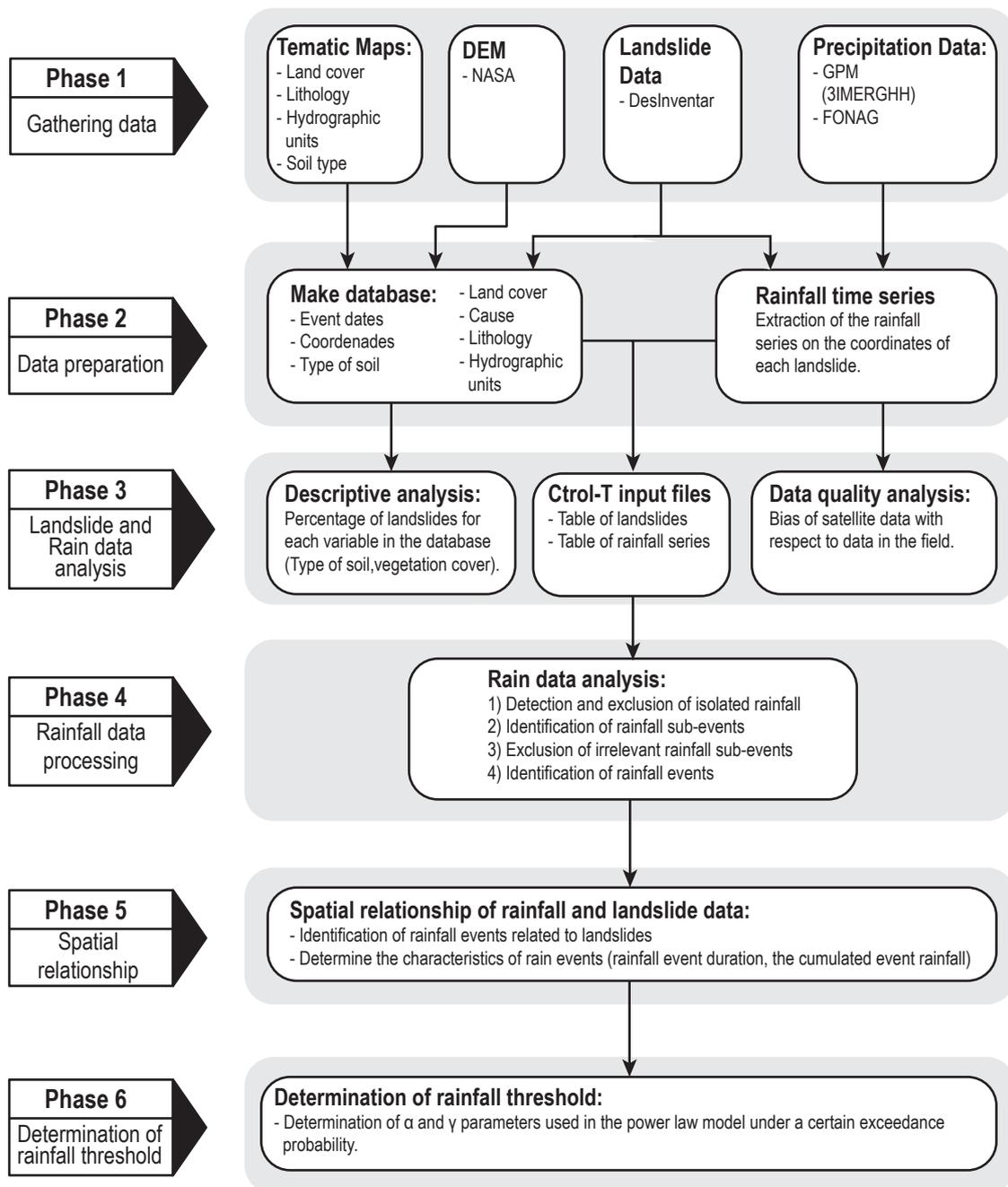
Figure 5.1: Workflow to estimate rainfall thresholds

analyzes the rainfall information to reconstruct the rainfall events based on pre-defined parameter settings and the rainfall series obtained by satellite data in the area, where the isolated events and the irrelevant sub-events are excluded.

5. **Spatial relationship of rainfall and landslide data:** In Block 2 of the CTRL-T

code through the input information (location of the rainfall series and landslide coordinates), it selects the nearest rainfall to each landslide event, then identifies the rainfall event associated with the landslide. The reconstruction of multiple rainfall conditions (MRC) is likely responsible for each slope failure.

6. **Determination of rainfall threshold:** In the last Block, the set of MRC is used to calculate the rainfall threshold. Moreover, the thematic maps will be based on the relationship between MRC and the following three main aspects:

   (a) Political division.

   (b) Attributes (precipitation regions, soil type, land cover, lithology, and Hydrographic Units level 3).

   (c) Climatic seasons.

5.1 Phase 1: Gathering data

5.1.1 Landslides data

Information on landslide occurrences was downloaded from DesInventar which is a systematic database of small, medium, and high-impact disasters. The download process can be found in Annex A.

The raw database from DesInventar has the following information: a serial number identifying the landslide events; longitude; latitude; province; canton; parrish; location, which gives the information of the specific place where the event occurred (roads, and neighborhood); event, which specifies the event date; time; cause, comments, source, victims (deaths, injured, missing), infrastructure losses (houses destroyed, and damaged), and others.

Here a first filter was applied to identify those events that record occurrence time and coordinates, the result showed that out of 936 landslides in the study area, only 354 events from 2013 to 2018 have enough information (coordinates). These landslide events are

distributed in the study area as follows: Pichincha 265 (74.85%), Imbabura 64 (18,07%), and Carchi 25 (7.06%).

5.1.2   Rainfall data

Rainfall data comes from 2 databases: satellite GPM and ground sensors from the Fondo para la protección del Agua (FONAG) hydro-meteorological network.

According to Ramadhan et al. (2022) the most accurate IMERG run is that with a greater delay (IMERG-Final) because this data is compared and adjusted to rain gauges. Therefore, the product of the GPM to be used in this study is 3IMERGHH (IMERG-Final). This product is the result of the inter-calibration of various precipitation-relevant satellite passive microwave (PMW) sensors comprising the GPM constellation using the 2017 version of the Goddard Profiling Algorithm to the GPM Combined Ku Radar-Radiometer Algorithm Combined Radar-Radiometer Precipitation Algorithms (CORRA) product, and merged into half-hourly 0.1°x0.1° fields. It has a frequency of half an hour and information from 2000 to the current date. Among the available estimates, the "precipitationCal" variable was chosen because it is a multi-satellite precipitation estimate with rain gauge calibration (Huffman, Stocker, Bolvin, Nelkin, & Tan, 2017). The information was downloaded from 2013-01-01- 00:00:00 UTC to 12-31-18 23:30:00 UTC, between -1.65 to 1.35 of latitude and -79.75 to -76.25 longitude, which comprised a total of 105168 files in ncf4 format, more information Annex B.

FONAG operates a hydro-meteorological network that provides information on key environmental variables for the hydrology of high Andean ecosystems. The network is made up of 89 hydro-meteorological stations (56 pluviometric, 23 meteorological, and 10 hydro-metric), all located in water-source areas of great importance for the city of Quito (Coronel, 2020). The temporal resolution of rainfall data is 1 hour. In this study, 24 stations were used, which had precipitation information in the pre and post period (500 hours before and 250 hours after) of the closest landslide. To get more information on this process see Annex C.

5.2   Phase 2: Data preparation

In this phase, the objective is to cross-reference geographical information to carry out a statistical analysis of the area; process the satellite-based rainfall data to form time series.

5.2.1   Building up the landslides database

Using geographic information systems and various thematic maps, an attribute event assignment by location was made. The selected attributes are shown in Table 5.1 which summarizes each element of the database. The first attributes (until comments) correspond to the landslide information, and the other corresponds to the selected descriptive attributes of the study area.

Table 5.1: Summary of database Information

Name	Description	Source
ID project	Index landslide LXXX	
longitude	Longitude coordinates in degrees	DesInventar
latitude	Latitude coordinates in degrees	
serial	Internal code from DesInventar	
date	Specific date and time of each event	
province		
canton	General location	
parish		
location	Specific location description, magnitude, and time range of occurrence	
Comments	Specific details of the event such as magnitude, general location, structures damages, and casualties	
rain_region	Precipitation Regions based on the homogenization of rainfall for each natural region of Ecuador (Ilbay-Yupa et al., 2021)	
soil_taxo	Soil taxonomy (classification of soil types according to the properties)	MAG
litho	Lithology according to MAGAP (detailed)	MAGAP
litho_gen	General classification of the major groups of lithology	
cob_lev1	Land cover Classification based on (MAG, n.d.)	MAG
cob_lev2		
cob_lev3		
hb_lev1	Identifies the hydro-graphic basins	MATE
hb_lev2		
hb_lev3		

5.2.2 Rainfall time series

To make the rainfall time series, the code C002 (Annex D Table D.1) was used, this code extracted the information for the "precipationCal" variable in ncf4 files and write time series on csv files for each pixel in the downloaded gridded-data domain. The data has 31 x 36 pixels of 0.1° with a time resolution of half-hour. To use these series in the CTRL-T code, it is necessary to accumulate values in 1-hour intervals. The code C001 yielded 1116 series in csv files with hourly precipitation. Out of the 1116 series, only those found in the study area were chosen. As a result, 220 series were obtained which were named P000 (Figure 5.2).
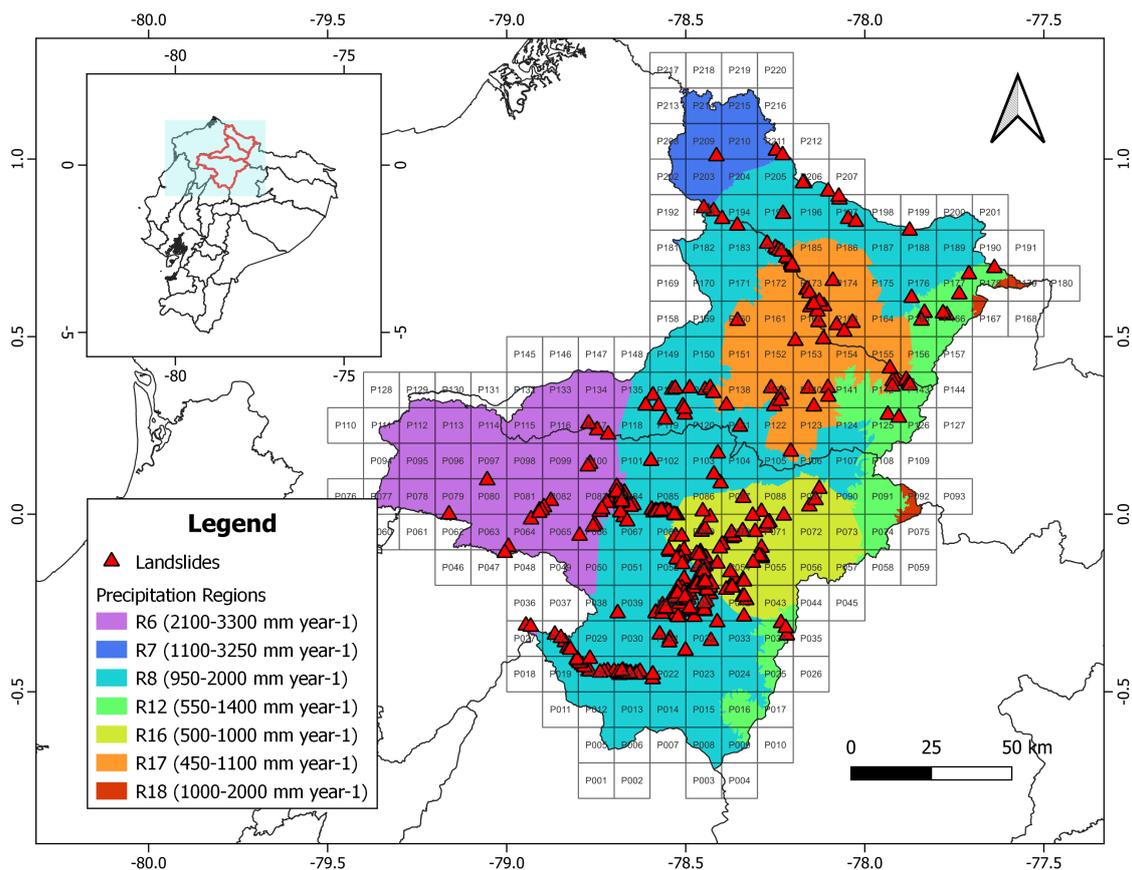
Figure 5.2: Precipitation gridded data (P001 - P220) over the study area showing precipitation regions from (Ilbay-Yupa et al., 2021)

5.3 Phase 3: Landslide and Rainfall data analysis

5.3.1 Descriptive analysis

The descriptive analysis of the study area was carried out by matching information (attributes) of thematic maps with landslide locations. Then, the area was estimated for each attribute. This information is detailed in Chapter 4. In this way we identify the most relevant attributes linked to event occurrences as shown in Table 5.2. All these attributes are used for the generation of thematic maps with the specific MRC of each area, which is done after phase 5.

Table 5.2: Summary of relevant attributes of each category

Category	Attribute	Landslides	Percentage (%)
Soil type	Andisols (A)	108	30.51
	Unknown:urban-peri-urban areas (Un)	87	24.58
	Mollisols (M)	84	23.73
	Entisols (E)	33	9.32
Land cover	Populated area (Pa)	108	30.5
	Shrubby vegetation (Sv)	67	18.9
	Native Forest (Nf)	54	15.3
	Grassland (G)	40	11.3
	Agricultural mosaic (Am)	30	8.5
	Herbaceous vegetation (Hv)	20	5.6
Lithology	VO	267	75.42
	AD	45	12.71
	TI	19	5.37
	CL	16	4.52
Hydrographic Units level 3	Esmeraldas River Basin (ErB)	281	79.4
	Mira River Basin (MrB)	71	20.1
Precipitation Regions	R8 (950-2000 mm $year^{-1}$)	195	55.1
	R16 (500-1000 mm $year^{-1}$)	64	18.1
	R6 (2100-3300 mm $year^{-1}$)	39	11
	R17(450-1100 mm $year^{-1}$)	37	10.5
	R12 (550-1400 mm $year^{-1}$)	15	4.2

5.3.2 Data quality analysis

The GPM data (IMERG-Final) involves adjustment using ground data available through the Global Telecommunication System (GTS). However, in Ecuador, the number of sta-

tions reporting to the GTS is still low, thus a first quality assessment is mandatory. To assess data quality, GPM satellite grid precipitation was compared to FONAG ground precipitation data. A short-time and pixel-to-point comparison were made between the FONAG time series with their respective GPM series (Figure 5.3).
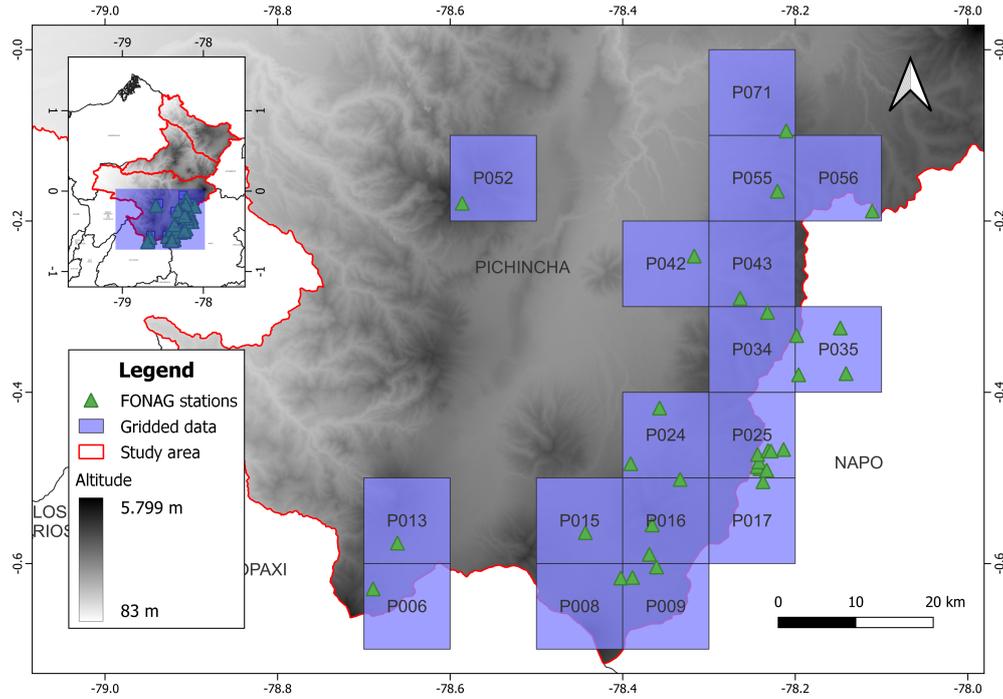
Figure 5.3: Location of FONAG stations and gridded data from GPM

A period of 750 hours was selected around the event date of the closest landslide to each FONAG station (500 hours before and 250 hours after). The following metrics were calculated: BIAS (Eq. 5.1), root mean square (RMSE) (Eq. 5.2) includes both systematic (bias) and non-systematic (random errors), and Pearson Correlation γ_{xy} (Eq. 5.3).

$$BIAS = \frac{1}{n}\sum_{i=1}^{n}(EG_i - EF_i) \tag{5.1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(EG_i - E_i)^2} \tag{5.2}$$

$$\gamma_{xy} = \frac{Cov(EG, EF)}{\sqrt{Var(EG)} \times \sqrt{Var(EF)}} \tag{5.3}$$

where EG is the GPM precipitation value, EF is the FONAG precipitation ground-data value, and n is the number of observations.

The statistical metrics are as follows: BIAS between 0.12 to 0.35 with a mean of 0.26, RMSE between 1.15 to 2.01 with a mean of 1.58, and Person´s correlation between -0.04 to 0.3 with a mean of 0.07. The low value of the Pearson correlation suggested that the GPM product is delayed with respect to the ground series. A detailed analysis of the causes for the lag is beyond this study, however, they might be related to the poor performance of the morphing time interpolation scheme due to the intrinsic limitations of the reanalysis model to simulate the vertically integrated vapor over complex terrain. Lag-correlation analysis was performed in order to identify the lag time that maximizes the correlation pattern among the GPM and ground precipitation estimates (Figure 5.4).

$Lag = -6$ showed the best correlation among all-time series. Using a lag equal to 6h the correlation values range between 0.19 to 0.55 with a mean value of 0.36. This Lag was introduced as a correction for the GPM time series and then the evaluation metrics were recalculated. The new metrics are BIAS between 0.13 to 0.35 with a mean value of 0.26 (Figure 5.5), and RMSE between 0.87 to 1.96 with a mean of 1.39 (Figure 5.6).

To compensate for bias, a bias correction factor was used. The factor was calculated with the mean relative bias equation (Equation 5.4 and used according to Equation 5.5). Then, with the corrected values the mean value of BIAS is 0.052, and for RMSE the mean value is equal to 0.77. Finally, the bias-corrected time series are then used as input for the CTRL-T code.

$$\text{Mean Relative BIAS} = \sum_{i=1}^{n.station} \frac{\frac{\sum_{i=1}^{n}(EG - EF)}{\sum_{i=1}^{n} EF}}{n.station} = 2.059 \tag{5.4}$$

$$\text{time serie} = GPM \times \frac{1}{\text{Mean Relative BIAS}} \tag{5.5}$$
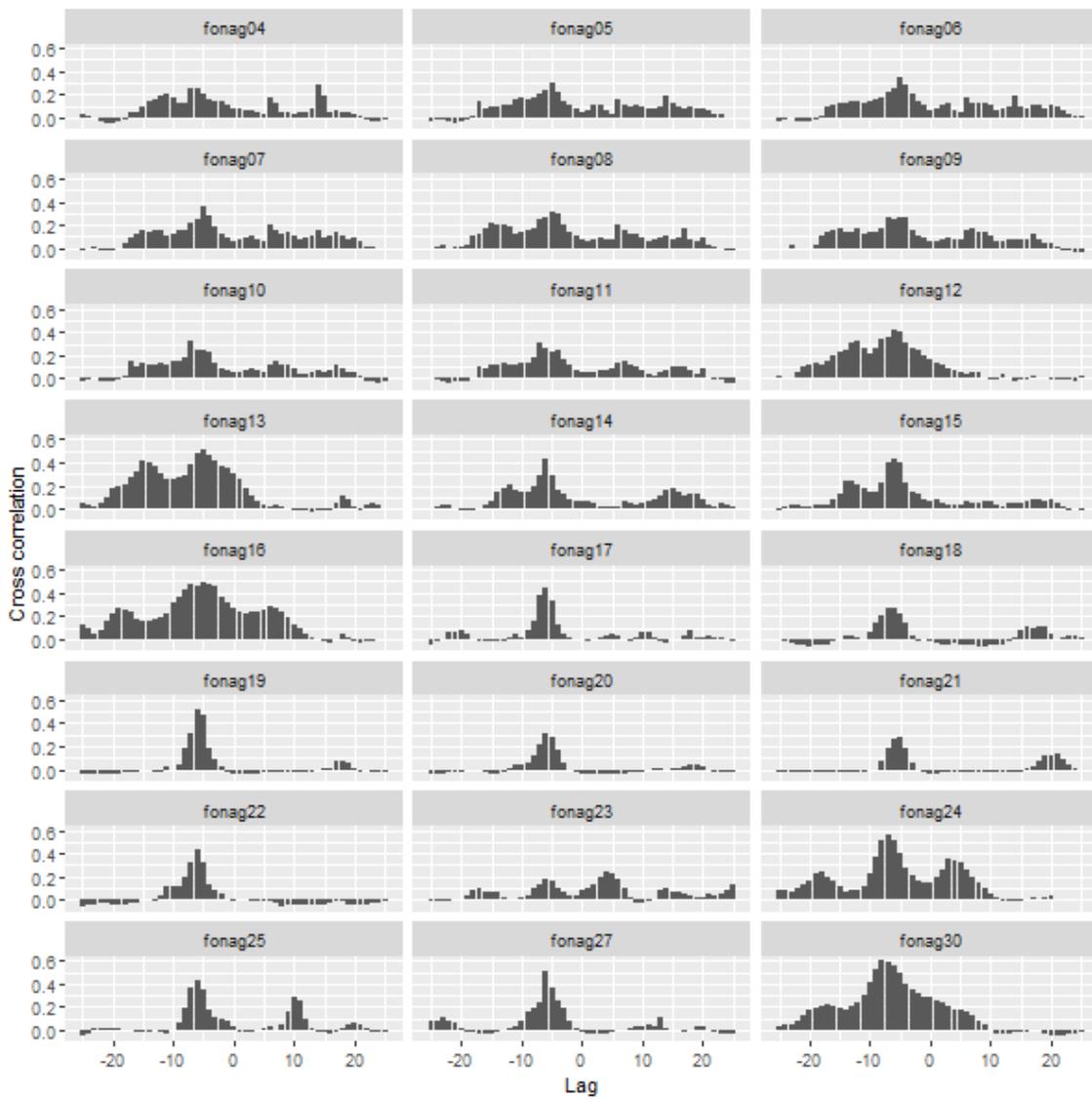
Figure 5.4: Lag cross-correlations for 24 FONAG stations versus their corresponding GPM precipitation data
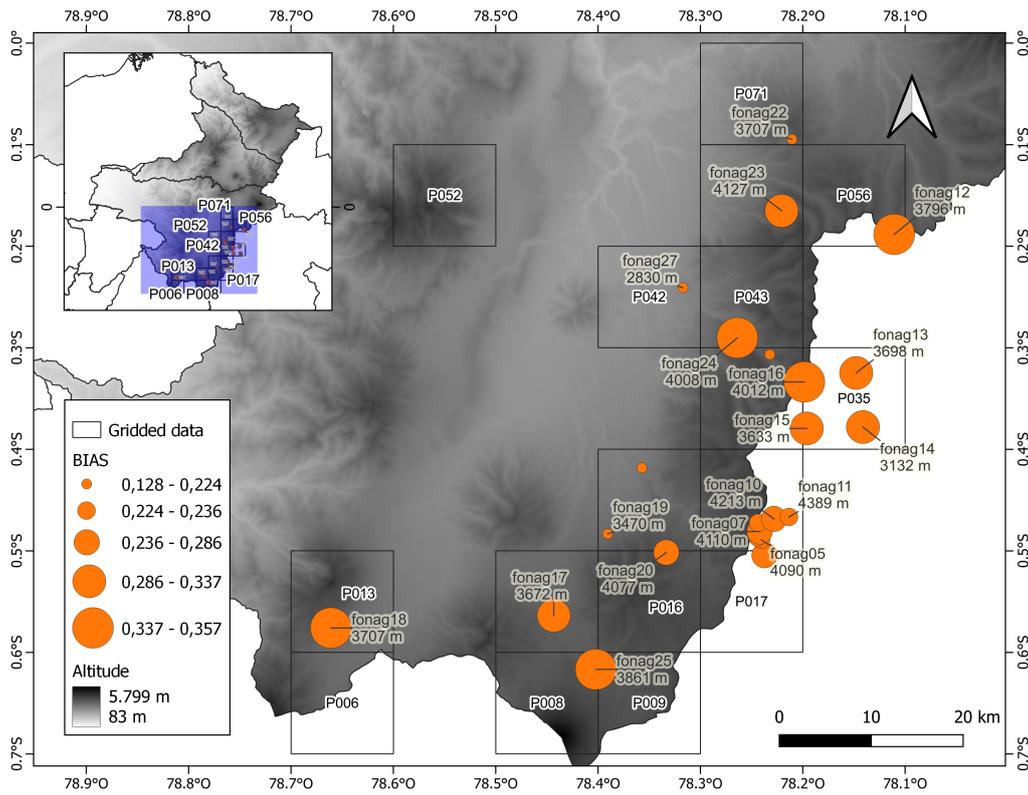
Figure 5.5: BIAS (graded circles) of FONAG stations and gridded data from GPM
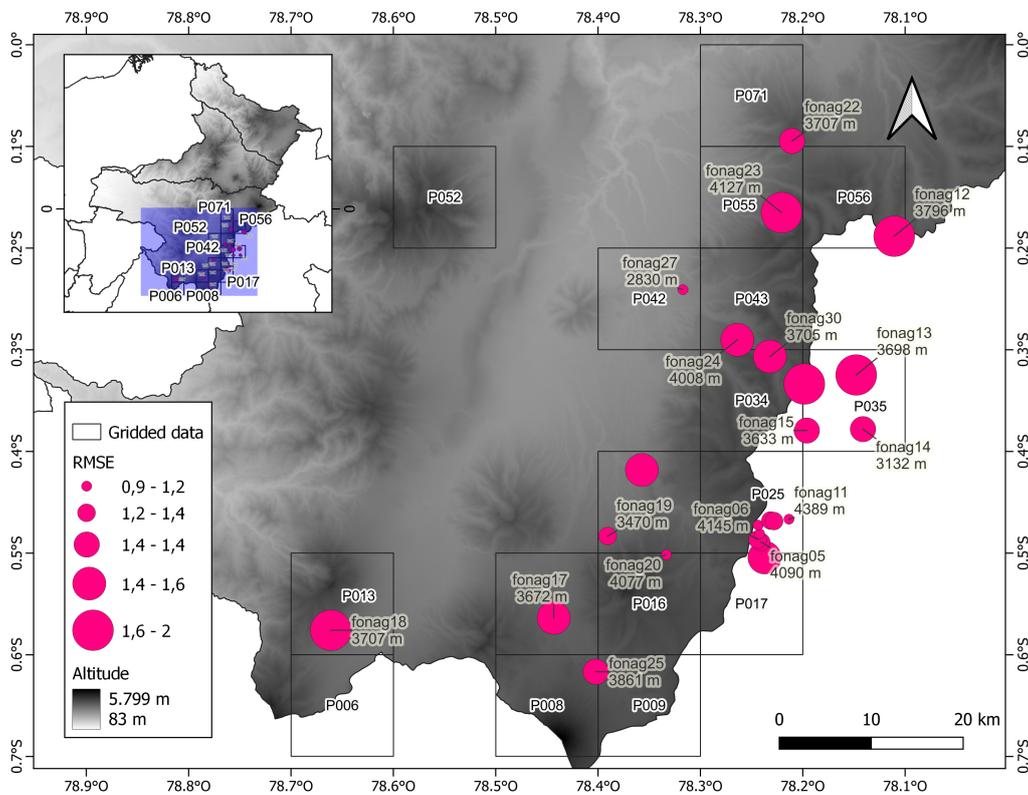
Figure 5.6: RMSE (graded circles) of FONAG stations and gridded data from GPM

5.3.3 CTRL-T input files

The CTRL-T is an algorithm that automates the process of reconstructing rainfall events, selecting the representative rain gauges, getting multiple failure rainfall conditions in terms of rainfal event Duration (D) and cumulated event rainfall (E), modeling the antecedent rainfall, and giving a probability to each rainfall condition, and calculating probabilistic rainfall thresholds and their associated uncertainties. This code reduces time analysis that usually is made by hand and reduces the probability to make a mistake or avoid some observation.

The inputs are a set of two csv files and one folder with rainfall series (220 pixels). The first csv file (Table of Landslides) has the landslide information and the second csv file (Table of rainfall series) has the rainfall series parameters used in CTRL-T code. And, finally, the folder has csv files with rainfall series for each pixel.

5.3.3.1 Landslide Information

The information items of the input file of landslides are shown in Table 5.3. First, the ID_project identifies each event according to the date (i.e. L001 is the 1st landslide that started in the year 2013, L002 is the 2nd, etc). Then, ID_lan describes the temporal order of an event during a single rainfall event (i.e. a is the 1st landslide, b is the 2nd, etc). The class_number labels landslides based on the place of occurrence (i.e. N_0 when there is no information, N_1 when the event occurred on a road, and N_2 when occurred in a different place).

Then, the class_type classifies the landslides according to their impact (i.e. C_0 is used for a gentle event with no obstructions, C_1 when the landslide has blocked a road from 1 - <10m, C_2 when the landslide cover a road from 10 - <20m, C_3 when the landslide has blocked the road from 20 - <30m and there are human injuries, C_4 when the landslide blocks the road from 30m to 200m, there are human injuries, and structures damaged. Finally, C_5 when the landslide blocks the road more than 200m, also there are structures

damaged, and human injuries).

The geo_acc indicates the mapping accuracy for each landslide (i.e. $P_1 < 1km$, and $P_2 < 10Km$). Date information indicates the occurrence date (day/month/year Hour: Minutes). Finally, the date_acc represents the accuracy of date occurrence (T_1 when the exact time of occurrence is known, T_2 is used when the event happened in a period, T_3 when the day of occurrence is known).

Table 5.3: Landslides input information

Landslides records								
ID project	**ID lan**	**class number**	**class type**	**longitude**	**latitude**	**geo acc**	**date**	**date acc**
L002	a	N1	C2	-78.54515	-0.105512	P1	6/12/ 2013 23:11	T1
L003	a	N1	C2	-78.50519	-0.232499	P1	7/12/ 2013 23:26	T1
L004	a	N2	C3	-78.41246	-0.303877	P1	28/12/ 2013 9:41	T1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
L352	a	N1	C2	-77.6385	0.6932	P1	4/12/ 2018 13:30	T1
L353	a	N1	C2	-78.44602	-0.190208	P1	18/12/ 2018 10:42	T1
L354	a	N2	C4	-78.41071	0.171466	P1	27/12/ 2018 9:37	T1

5.3.3.2   Rainfall series information

The input rainfall series information, necessary to the CTRL-T code, is shown in the Table 5.4. The "cod area" identifies rainfall series based on the political administration borders (Pichincha, Imbabura, and Carchi provinces, PIC). The pk sensor indicates the height of the rain gauge; but, in this study, this is not applicable as the information is obtained from satellites. Moreover, in this study, the value of the instrumental sensitivity (GS) depends on the Dual-frequency Precipitation Radar (DPR), which is a spaceborne precipitation radar capable of making accurate rainfall measurements with a 0.2 mm/h of sensitivity (Savtchenko, 2017).

The following aspects are parameters used for the reconstruction of the rainfall series. Firstly, the parameters p1_w (wet season), and p1_d (dry season) identify the isolated events in each season using a time range in hours. The p2_w and p2_d are parameters to determine the rainfall sub-events with a prescribed time range (hours). The p3 is a

parameter in mm used to identify the irrelevant rainfall sub-events and exclude them. The last parameters used for the reconstruction of the rainfall series are p4_w, and p4_d, which consider rainfall as one. They aggregate single or multiple sub-events to obtain single rainfall events. The values of p(1-4) parameters will be explained later.

Then, there are two items describing the start warm season (sws) and end warm season (ews) month of the dry season. To obtain this information we analyzed the in-year monthly precipitation distribution (Figure 5.7) for each rainfall series considering the different precipitation zones as defined by Ilbay-Yupa et al. (2021). Also, this analysis unveils the spatio-temporal features of precipitation seasonality in the study area, i.e. when the rainy season is onset the rainfall series tends to change. Finally, the last item of this table is ID rain which identifies each precipitation series (i.e. P001 for the first rainfall series, etc).

Table 5.4: Example of the input csv file with parameters for processing the rainfall series

Table of rainfall series

cod area	longitude	latitude	pk sensor	gs	p1_w	p1_d	p2_w	p2_d	p3	p4_w	p4_d	sws	ews	ID rain
PIC	-78.75	-0.75	—	0.2	6	3	6	3	1	48	24	6	9	P001
PIC	-78.65	-0.75	—	0.2	6	3	6	3	1	48	24	6	9	P002
PIC	-78.45	-0.75	—	0.2	6	3	6	3	1	48	24	6	9	P003
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
PIC	-78.45	1.25	—	0.2	6	3	6	3	1	48	24	7	9	P218
PIC	-78.35	1.25	—	0.2	6	3	6	3	1	48	24	7	9	P219
PIC	-78.25	1.25	—	0.2	6	3	6	3	1	48	24	7	9	P220
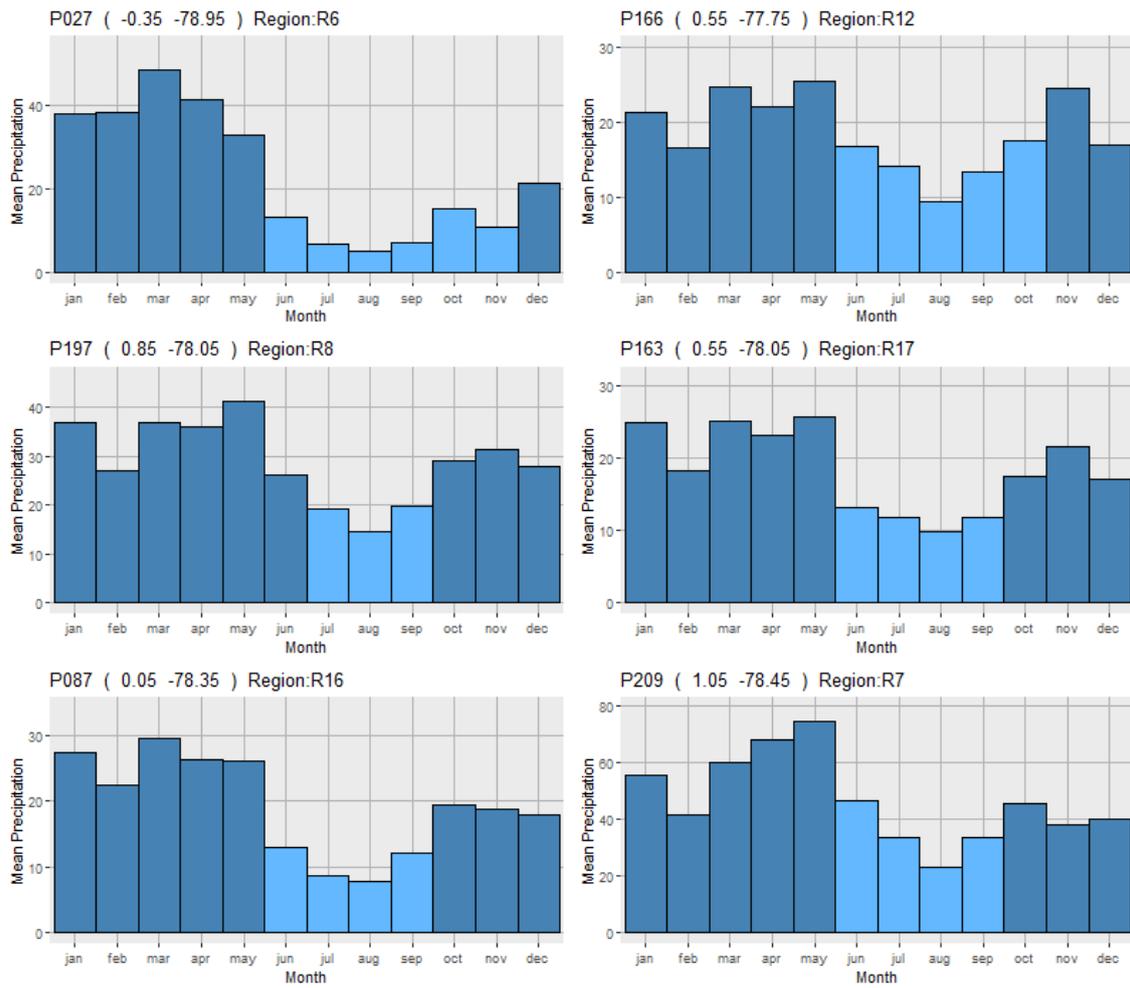
Figure 5.7: Rainy season in different precipitation regions: wet (blue) and dry (light blue).

5.4 Phase 4: Rainfall data processing

To perform the automatic reconstruction of the rainfall events associated to landslides (Figure 5.8), we make use of the tool developed by Melillo, Brunetti, Peruccacci, Gariano, and Guzzetti (2015). The algorithm first identifies individual rainfall events from a record of rainfall estimates. The rainfall information was obtained from the hourly bias-corrected satellite precipitation data for a period of five years (from 2013 to 2018). The reconstruction process involves five steps explained below. The set of parameters used in the analysis are shown in Table 5.5.

Table 5.5: Parameters for identification of rainfall events associated with landslide occurrences (phase 4)

Spatial and Rain data processing Parameters					
Steps	**Parameter**	**Description**	**Parameter**	**value**	**Unit**
S1	Gs	Instrumental sensitivity	0,2	0,2	mm
S2	P1	Required time to identify a isolated rainfall	3	6	h
S3	P2	Time to consider sub-events	6	12	h
S4	P3	Irrelevant sub-events	1	1	mm
S5	P4	Time to estimate rainfall events	48	24	h

5.4.1   Step 1: Pre-processing of rainfall data

The possible gaps in the record can have periods from a minimum of 1h to several days or weeks, marked by specific tags in the record. In this step, the main goal is to exclude the minimum values of rainfall as shown in Figure 5.8 a, in this case, a filter deletes all the values lower than 2 mm (Melillo et al., 2015).

The algorithm checks the continuity of the record and detects the gaps it searches the rainfall record for tagged and untagged missing measurements, and in the case of finding a measurement that is not available, it changes to the "na" tag. In addition, when the hourly measurements (E_h) are lower than the instrumental sensitivity (G_s) of 0.2 mm/h (Savtchenko, 2017). The value of E_h is considered as noise, and the algorithm sets the measurements to $E_h = 0.0mm$. Following these steps, the corrected rainfall is ready to be processed to reconstruct the rainfall events starting with the next step.

5.4.2   Step 2: Detection and exclusion of isolated rainfall

According to Melillo et al. (2015), the algorithm identifies the isolated rainfall as an hourly volume. First, it selects the index where the values are different from 0. Then, based on the following three conditions it identifies the isolated events:

- Once all the indexes was identified the algorithm makes a subtraction between $index - (index - 1) > p1$.

- Then, it makes a subtraction between $(index + 1) - index > p1$.

- Finally, the values must be less than the minimum value (val min = 1 mm), and the algorithm sets the isolated, irrelevant measurements to $E_h = 0.0mm$.

For each precipitation measurement of the rainfall series, it is analyzed to the previous and following values. To determine the isolated rainfall the time between each value should not pass the p1 value as shown in Figure 5.8b. In this case, the value for this parameter depends on the local climatic conditions (C). As we mentioned before, in our study area we consider two rainy season periods, the wet (C_w) and dry (C_d) seasons. In this way, the wet season considers a range until 6 hours ($P1_w$), and the dry season has a range of 3 hours ($P1_d$).

5.4.3 Step 3: Identification of rainfall sub-events

In this step, the individual rainfall sub-events must be identified, which is a period of continuous rainfall separated from the previous and the following sub-events by dry periods with no rain as shown in Figure 5.8c. Moreover, it is important to know the rainfall behavior in each season for the study area. In this case, during the dry period (C_w) the rainfall is mostly generated due to the north-south movement of the Inter Tropical Convergence Zone (ITCZ).

On the other hand, in the wet season (C_w) the precipitation is influenced by orographic rain, local warming (convection rain), and movement of the ITCZ (Ilbay-Yupa et al., 2021). As we know, based on the season we consider the range to determine a sub-event, for the wet season this value is 6 hours ($P2_w$), and the dry season has a range of 3 hours ($P2_d$).

When the algorithm reconstructs a rainfall sub-event, it must check for the continuity of the rainfall time series for a sub-event. An event is excluded if single or multiple "na" measurements, considered interruptions, are found in the rainfall record in the period covered by the sub-event (Melillo et al., 2015). On the other hand, when there is no "na"

the sub-event is defined and the following rainfall features are measured:

- The sub-event duration

- The sub-event total rainfall (E_s)

For both, the algorithm computes summing the hourly rainfall measurements in the sub-event, $E_s = \sum E_h$. Which are important for the next step.

5.4.4 Step 4: Exclusion of irrelevant rainfall sub-events

In this step, the algorithm finds the sub-events that can be considered irrelevant for the reconstruction of rainfall events responsible for landslides occurrence as shown in Figure 5.8 d. For this study, a sub-event is considered irrelevant when the total cumulated rainfall for the sub-event (calculated at the end of the previous step) is lower than a given threshold value $E_s < P_3$, regardless of the duration of the sub-event. In our case $P_3 = 1mm$, it is considered a reasonable threshold to exclude sub-events whose contribution can be considered irrelevant to the possible initiation of rainfall-induced landslides. In this way, the irrelevant sub-events are excluded from the next analysis(Melillo et al., 2015).

5.4.5 Step 5: Identification of rainfall events

The final step for the reconstruction of the rainfall events is the aggregation of singles or multiple sub-events to obtain a single rainfall event as shown in Figure 5.8. Here the rainfall event is considered a period of continuous rainfall or a set of periods of continuous rainfall, separated from the events during the dry periods. In this stage the minimum length (in hours) is P_4, in this case, the minimum warm stage during the dry period is $P4_d = 24h$. On the other hand, for the wet period, the value of $P4_w = 48h$. Until here the algorithm all the rainfall events are recorded and defined. Then, it calculates the rainfall metrics for each of the detected rainfall events, including the event duration D_E computed by summing the number of hours in the rainfall event, taking into account the

hours for $E_H = 0$, and the event total of cumulated rainfall E_E computed by summing the sub-event rainfall $E_E = \sum E_S$ (Melillo et al., 2015).
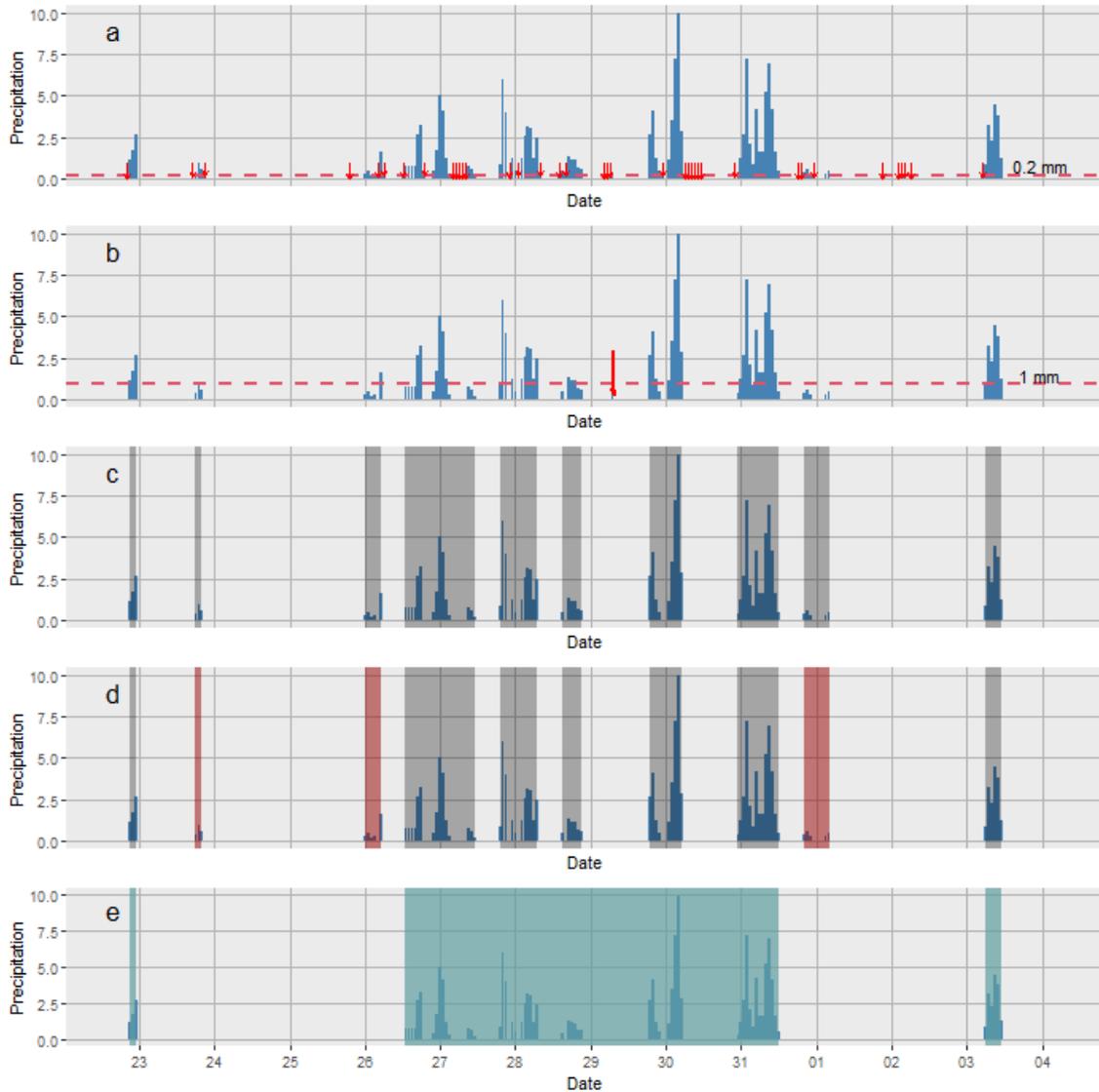
Figure 5.8: Pre-processing example of pixel 120. a) The first step excludes minimum values of rainfall using a filter of 0.2mm. b) The second step identifies isolated rainfall and excludes them. c) The third step identifies rainfall sub-events. d) The fourth step determines irrelevant sub-events, and e) The fifth step selects the single rainfall event.

5.4.6 Output from phase 4

This output is on the folder named "Reconstructed Rainfall events", where there is a csv file named Rainfall events, which has the following information as shown in the table 5.6. Therefore, for each ID_rain_gauge (rain grid code) there are two indexes index_pos1 and index_pos2, which relate each rainfall event detected on each gridded data. There are 68275 single events between 220 gridded data, so the grid precipitation code will repeat n times as rainfall events were detected based on the P_ parameters of the previous phase. The next two items are the RE_start_date and RE_end_date, which are the starting and ending dates of each rainfall event. The remaining items indicate some characteristics of rainfall events. The first one is the D_E representing the rainfall duration, E_E is for the cumulated rainfall, I_E is the rainfall mean intensity, IP_E is the maximum hourly rainfall, $Emax24_E$ is the maximum cumulated rainfall in 24 h, and the A_class is a rainfall event classification from 1 to 6 with the following ranges: 1 is the lowest precipitation during 24h (0-4 mm), 2 (4-16 mm), 3 moderate precipitation (16-32 mm), 4 (32-64 mm), 5 (64-128 mm) and 6 is the highest precipitation (128-1000 mm) (Melillo et al., 2018).

Table 5.6: Example of an output csv file that has the information obtained from phase 4

Table of rainfall series										
ID_rain_gauge	**index_pos1**	**index_pos2**	**RE_start_date**	**RE_end_date**	D_E	E_E	I_E	IP_E	$Emax24_E$	**A_class**
P001	34986	37334	28/12/2016 17:00	5/4/2017 13:00	2349	1636.05	0.696	19.62	75.53	5
P001	37411	38473	8/4/2017 18:00	23/5/2017 0:00	1063	783.68	0.737	44.6	100.86	5
P001	51434	52490	14/11/2018 1:00	28/12/2018 1:00	1057	434.36	0.411	19.85	84.48	5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
P220	48233	48233	3/7/2018 16:00	3/7/2018 16:00	1	1.11	1.11	1.11	1.11	1
P220	48330	48330	7/7/2018 17:00	7/7/2018 17:00	1	1.82	1.82	1.82	1.82	1
P220	49271	49271	15/8/2018 22:00	15/8/2018 22:00	1	1.1	1.1	1.1	1.1	1

5.5 Phase 5: Spatial relationship of rainfall and landslide data

5.5.1 Selection of rainfall events

In this phase, the algorithm analyzed the location of the grid precipitation data and landslides provided on the "input". According to Melillo et al. (2018) the algorithm automatically selects the relevant rainfall series, and it makes a buffer to determine the gridded data nearest to each event (Figure 5.9). The radius of this buffer depends on two parameters, the morphological settings of the study area (large radius for flat areas and short radius for mountain regions) and the density of gridded data. As the selected area is located in the Andes region, the altitude changes abruptly, so the radius used for the buffer was 10 km. On the other hand, the density for the gridded data is uniform (1 each 0.1°), so this was not taken in the count to determine the radius. Then, the algorithm identifies the rainfall event associated with the landslide based on the modeling of the cumulated event rainfall, it is explained detailed in the next section. Finally, the algorithm calculates the MRC for each relevant gridded data with the before information.

5.5.2 Modeling the cumulated rainfall responsible for the failure

According to Melillo et al. (2018), it is important to take into count the effect of soil water saturation as decay antecedent rainfall conditions. The model used by the code is the Antecedent water status model (Eq. 5.6). This model uses a k-decay factor that depends on the regolith storage capacity (porosity and depth), evapotranspiration rates, and the drainage rate of excess precipitation (Crozier, 1999). For the study, the conventional k-factor was used ($k = 0.84$), this value was proposed by Crozier (1999), and this value can be improved depending on the zone of the study (Glade, Crozier, & Smith, 2000).
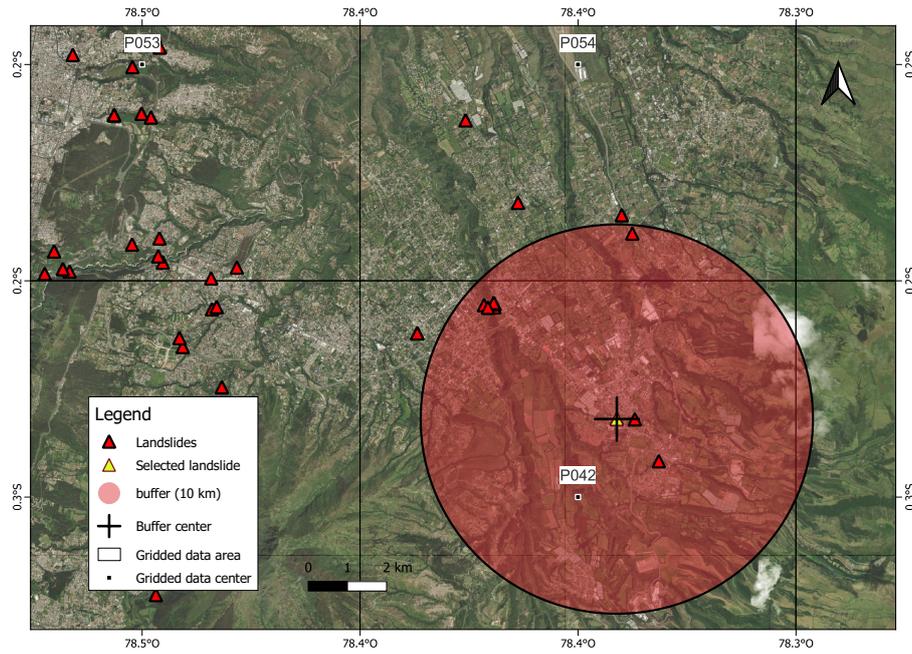
Figure 5.9: Example of selection of gridded data (white dots and squares) next to a selected landslide (yellow triangle), inside of a buffer (red circle) of 10km radius.

$$E_{\mathrm{L}} = E_{\mathrm{L}}(0) + k E_{\mathrm{L}}(1) + k^2 E_{\mathrm{L}}(2) + \ldots + k^N E_{\mathrm{L}}(N) = \sum_{i=0}^{N} k^i \left[E_{\mathrm{L}}(i) \right] \qquad (5.6)$$

Where:

$E_{\mathrm{L}}(0)$:   is the cumulated rainfall in the 24h before the landslide occurrence time (t_L).

$E_{\mathrm{L}}(i)$:   is the cumulated rainfall in the 24h of the $i - th$ day before t_L

N :   is the duration of the rainfall event in days (steps of 24h).

5.5.3   Selection of maximum probability

To know the most probable MRC of a grid precipitation data is necessary to consider the geographical and rainfall features. First, it selects the nearest gridded data inside the buffer and determines the w of each MRC. The MRC is a pair or set of pairs of the rainy characteristics (rainfall event duration (D_L) and cumulate event rainfall (E_L) with respect

to the landslide. In the begging, all of these combinations are considered equally probable for triggering a landslide, which means with the same weight $w = 1$.

Then, to determine the exact probability of each pairMRC the Eq.5.7 was used, which is proportional to the inverse square distance between the gridded precipitation data and the landslide (d^{-2}), the cumulated rainfall (E_L), and the rainfall mean intensity ($E_L D_L^{-1}$) (Melillo et al., 2018).

$$w = \mathrm{f}\left(\mathrm{d}, E_L, D_L\right) = d^{-2} E_L^2 D_L^{-1} \tag{5.7}$$

Once the algorithm assigns a weight to each pair of MRC (D_L, E_L), it evaluates the information and discards pairs ($w = 0$) based on the following two conditions: the first, if the difference between one pair and the subsequent is less than 10% of cumulated rainfall E_L. Second, if the delay between the rainfall ending time and the landslide occurrence time is more than 48h, this study only considered shallow landslides, so it is expected a short delay (Melillo et al., 2018). Then, it selects the grid with the most probable MRC (highest w). Finally, the algorithm calculates the threshold for all MRC and maximum probability rainfall conditions (MPRC)(which is the MRC with the highest w).

5.5.4 Outputs from phase 5

The output for this phase is the folder (Reconstructed rainfall conditions), which contains: one sub-folder (Individual files) with the analysis of rainfall conditions for each landslide, and 3 files: Processing Summary report.txt, MPRC.csv, and MRC.csv. All of this information is in a Drive folder. For an explanation of this output, the following sections indicate available the information for each output.

5.5.4.1 Individual files folder

For each landslide event, a single file is generated, and it is named Landslide_ID_project _ID_lan.pdf (i.e. Landslide_L002_a). This file contains a map with the nearest grid

precipitation data (inside the buffer) for the landslide event (Figure 5.10). To explain the information reported on the individual files the case of Landslide L009 was used, in the Appendix E the completed pdf file is shown, where the information of the two selected gridded data is reported.
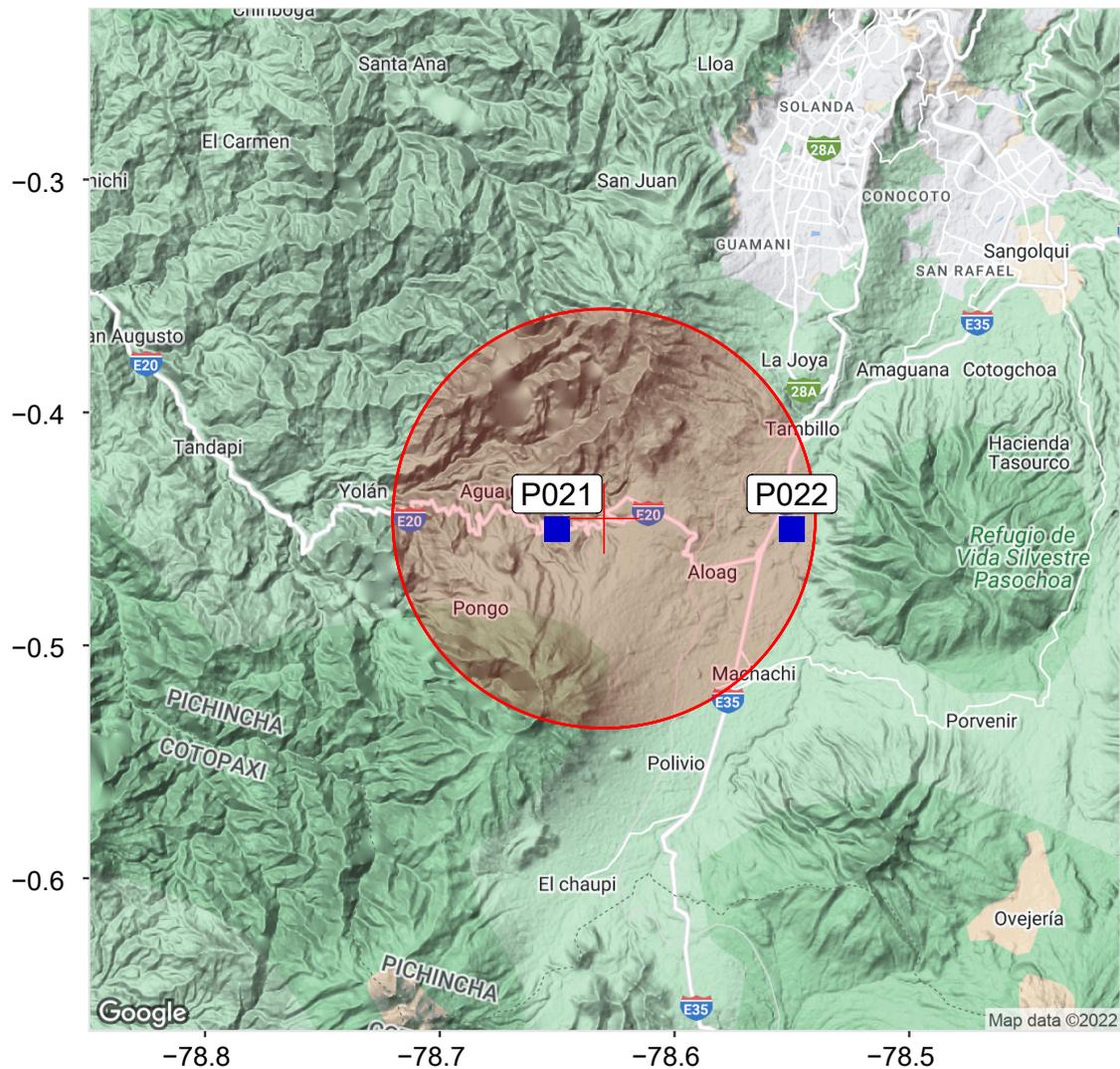
Figure 5.10: Example output phase 5: Selection of gridded data for the landslide L009 through a buffer, the nearest gridded data are two (P021 and P022)

To understand the reported file the grid P021 is used as an example, the following items are presented in this order:

1. information about rainfall data (Figure 5.11).

- Name of grid precipitation

- Location of gridded data (distance to landslide event)

- Total time (start and stop Date)

- Resolution (hours)

- Rainfall events indicate the number of rainfall events for the selected grid reported on the CSV file "Rainfall events".

- Finally, the statistics of rainfall data; for the rainfall duration event (D_E)) and cumulated rainfall event (E_E) are the values of Min, Max, Mean, Median, and SD.

RAIN GAUGE : P021
Distance from the landslide: 2.28 km
Temporal coverage
Start Date: 2013−01−01 07:00+0000
Stop Date: 2018−12−31 23:00+0000
Data Resolution
Temporal Resolution: hourly
Rainfall events
#: 248
Statistics:

	Min	Max	Mean	Median	SD
D_E (h)	1	1052	96.67	31.5	158.65
E_E (mm)	1.02	358.96	32.57	10.41	52.8

Figure 5.11: Example output phase 5: Rainfall information of the gridded data P021

2. Information related to the landslide and the selected grid precipitation data (Figure 5.12).

   - Identification of landslide event (ID_project_ID_lan)

   - Delay indicates the difference in time (hours) with failure occurrence

   - Date of landslide event

   - Rainfall event associated with the landslide indicates the position number in the CSV file "Rainfall event" for the gridded data.

- Date of the selected rainfall event (start and Stop).

- Rainfall information (D_E, E_E)

- Number of rainfall conditions responsible for the landslide

- Number of discarded rainfall conditions (threshold value:10%).

<u>ID LANDSLIDE</u> : **L009_a**   **DELAY:0 h**

Landslide date: 14−01−08 12:00
Rainfall event associated with the landslide
#: 42
Start Date: 2014−01−05 22:00+0000
Stop Date: 2014−01−15 12:00+0000
D_E **:** 231 h
E_E **:** 106.87 mm
Rainfall conditions responsible for the landslide
#: 3
Discarded rainfall conditions (threshold value:10%)
#: 1

Figure 5.12: Example output phase 5: Information of the grid P021 for the landslide L009

3. Distribution of the reconstructed rainfall pairs (D_E, E_E) in the logarithmic plane, all of these green squares represent the pair of combinations(MRC)(Figure 5.13).

4. Hourly rainfall measurement of the selected rainfall event, which is a compound of sub-events considered based on P-parameters (all of these pulses indicate the pairs of MRC) (Figure5.14).

5. Multiple rainfall conditions (D_E, E_E) responsible for the landslides in the logarithmic plane, which indicates the considered and discarded pairs (Figure 5.15).
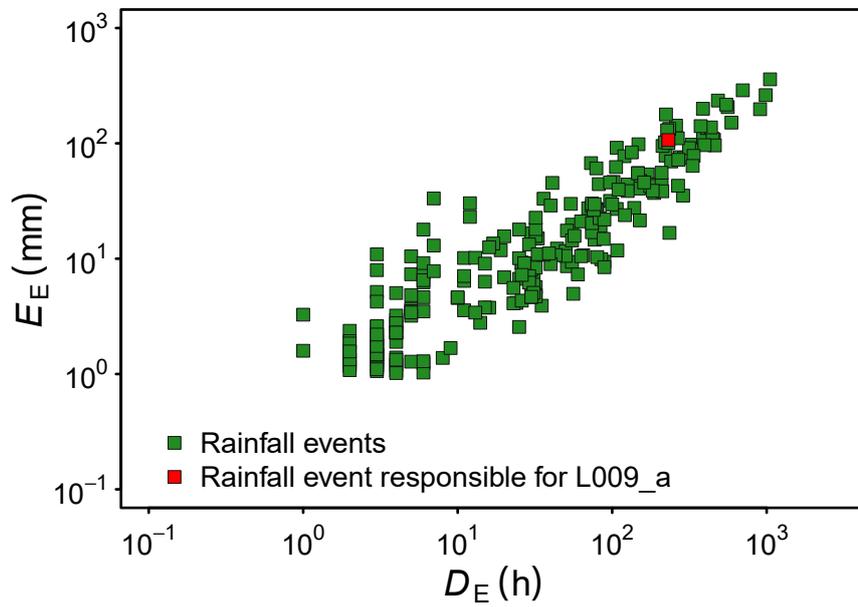
Figure 5.13: Example output phase 5: Logarithmic plane of E_E, D_E, where the green square represents the rainfall events, and the red square is the rainfall event responsible for L009.
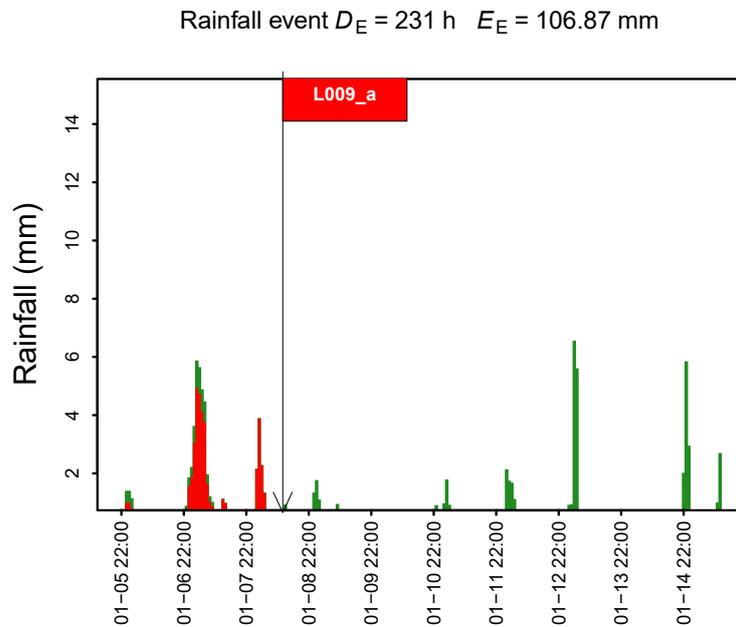
Figure 5.14: Example output phase 5: Rainfall event of the grid P021 for landslide L009
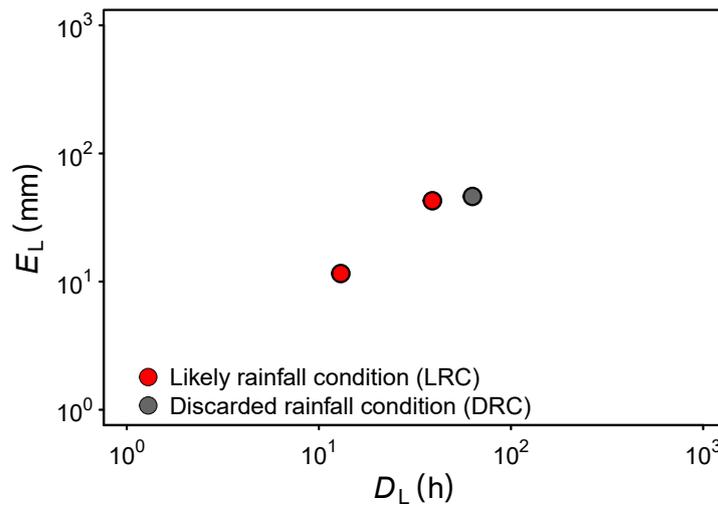
Figure 5.15: Example output phase 5: The red dots are the MRC selected for the landslide L009, and the gray dots represent the discarded MRC

5.5.4.2 Processing Summary report txt file

This file describes the total of landslide events, the number of events used for the determination of MRC (reconstructed, 301), and the number of discarded events (53). Also, it indicates each discarded event as shown in the Table 5.7.

Table 5.7: Summary Report

Total landslide #: 354
Total landslides reconstructed #: 301
Total landslides discarded #: 53
List landslides discarded:
L001
L004
L005
L071
⋮
L345
L349
L353
L354

5.5.4.3 MRC and MPRC csv files

Table 5.8 contains the information of MRC and MPRC files, which have the following items for each landslide event. The first three items correspond to landslide information explained in the input section, where the ID_project is for the identification of a landslide based on the date, the ID_lan describes the temporal order of an event during a single rainfall event, and the date indicates the time of failure occurrence. Then, the RRG_select identifies the relevant grid precipitation data based on the buffer. The RRG_distance indicates the distance between the landslide and grid location. From the next item until the last one, it contains the rainfall information related to a landslide in the following order: D_L, E_L, I_L, which indicates the rainfall duration, the cumulated event rainfall, and the rainfall mean intensity respectively. The ID_rain_gauge is the gridded data code, the number of rainfall events associated with a grid is reported on REN. Then, IP_L indicates the maximum hourly rainfall, the Emax24_L is the maximum cumulated rainfall in 24 h, and a class is a rainfall event classification described in the input section. The ms_flag is a binary value that indicates if the current rainfall condition has a maximum value of the score w (n_MRC = 1), and the n_MRC is the number of multiple rainfall conditions. This information is the same for both tables, but the difference is on the MPRC table which contains only the information for the MRC with a w equal to one for each landslide.

Table 5.8: Tables of MRC and MPRC output csv files

Table of MRC.csv output

ID_project	ID_lan	date	RRG_select	RRG_distance	D_L	E_L	I_L	ID_rain_gauge	REN	IP_L	Emax24_L	A_class	ms_flag	n_MRC
L002	a	13-12-07 00:00:00	1	4.98	5	5.5	1.1	P052	50	1.82	5.5	2	1	5
L002	a	13-12-07 00:00:00	1	4.98	55	14.7	266	P052	50	1.88	6.94	2	0	5
L002	a	13-12-07 00:00:00	1	4.98	121	16.8	139	P052	50	1.88	6.94	2	0	5
L003	a	13-12-08 00:00:00	2	6.44	14	9.3	667	P041	49	3.54	9.34	2	0	5
⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
L351	a	18-12-03 13:00:00	1	3.73	228	68.6	301	P177	328	7.83	32.74	4	0	4
L352	a	18-12-04 14:00:00	1	4.97	31	57.9	1 868	P178	320	7.44	53.67	4	1	3
L352	a	18-12-04 14:00:00	1	4.97	104	77.2	742	P178	320	7.44	53.67	4	0	3
L352	a	18-12-04 14:00:00	1	4.97	158	85.9	544	P178	320	7.44	53.67	4	0	3

Table of MPRC.csv output

ID_project	ID_lan	date	RRG_select	RRG_distance	D_L	E_L	I_L	ID_rain_gauge	REN	IP_L	Emax24_L	A_class	ms_flag	n_MRC
L002	a	13-12-07 00:00:00	1	4.98	5	5.5	1.1	P052	50	1.82	5.5	2	1	5
L003	a	13-12-08 00:00:00	2	6.44	87	26.1	0.3	P041	49	3.71	11.78	2	1	5
L006	a	14-01-07 00:00:00	1	2.41	3	15.7	5.23	P066	56	9.73	15.69	2	1	1
L007	a	14-01-07 01:00:00	1	5.33	4	24.6	6.152	P066	56	9.73	24.61	3	1	1
⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
L348	a	18-11-25 18:00:00	1	4.63	151	42	0.278	P041	315	4.76	17.21	3	1	3
L350	a	18-11-27 18:00:00	1	5.78	38	50.3	1.325	P194	261	4.45	38.86	4	1	1
L351	a	18-12-03 13:00:00	1	3.73	12	32.7	2.728	P177	328	7.83	32.74	4	1	4
L352	a	18-12-04 14:00:00	1	4.97	31	57.9	1.868	P178	320	7.44	53.67	4	1	3

5.6   Phase 6: Determination of rainfall threshold

According to (Brunetti et al., 2010), there are two statistical methods for the definition of rainfall threshold. The first one is the Bayesian inference method, which used a probability approach (Bernoulli probability) to estimate the intercept and the slope. The Bernoulli probability is determined for a data point occurring at a given value of cumulated rainfall and duration event. This method worked better for small data sets because it analyzed each data point of rainfall features.

The other method is frequency analysis, which determines the intercept and slopes with the power law curve selected to represent the rainfall threshold. It analyzed the empirical rainfall conditions of known landslides. To solve the problems due to adjusting data, the empirical data is log-transformed and plotted in a graph log (E) vs log (D). To fit this information the least square method is used with the linear equation ($log(E) = log(\alpha) - \gamma log(D)$), which is entirely equivalent to the power law (Equation 5.9) in linear coordinates. For each rainfall event, the difference $\delta(D)$ between the cumulated rainfall event and the corresponding cumulated rainfall value ($\delta(D) = log[E(D)] - log[E_f(D)]$) is calculated. The Kernel Density Estimation is used to determine the probability density function pdf of the distribution of $\delta(D)$. Then, the result is fitted with a Gaussian function (Equation 5.8)

$$f(x) = a \exp\left(-\frac{(x-b)^2}{2c^2}\right) \tag{5.8}$$

Where $a > 0, c > 0$ and $a, b, c \in \mathbb{R}$

Finally, the threshold corresponds to different exceedance probabilities (50 to 0.005%) based on the fitted distribution $\delta(D)$. For the threshold at 5%, the rainfall conditions below the curve have a probability of less than 5% for triggering a landslide. The determination of multiple thresholds at different exceedance levels allows the prediction of possible landslide occurrence based on rainfall measurements.

Moreover, to determine the mean values of the intercept (α) and slope (γ), and the

associated uncertainties ($\Delta\alpha$, $\Delta\gamma$) the bootstrap technique is used. It is a non-parametric technique for determining the mean sample distribution of a population from an empirical data set (Peruccacci, Brunetti, Luciani, Vennari, & Guzzetti, 2012). This technique generates k series of m randomly selected from the empirical data set (n events). Then, with the analysis of each k series is possible the calculation of mean value and the uncertainty associated with a parameter, including the standard deviation. The rainfall threshold is defined using the mean values of α and γ, and the respective uncertainties are calculated $\Delta\alpha$ and $\Delta\gamma$ for each one is determine the minimum ($E_{min\Delta\alpha}$ and $E_{min\Delta\gamma}$) and maximum ($E_{max\Delta\alpha}$ and $E_{max\Delta\gamma}$) curves around the 5% threshold. Then, the equations for min and max are in the E-D system as explained Peruccacci et al. (2012), and determine that for longer rainfall duration the cumulated rainfall (E) is sensitive to $\Delta\gamma$ because γ is the exponent of the power law curve.

In this phase, the CTRL-T used the power law (Equation 5.9) to determine the threshold for the following categories: political division (provinces Pichincha, Imbabura, and Carchi), attributes (precipitation regions, land cover, soil type, lithology, and hydrographic units level 3), and climatic seasons.

$$E = (\alpha \pm \Delta\alpha) \times D^{(\gamma \pm \Delta\gamma)} \tag{5.9}$$

Where:

E	:	is the cumulated event rainfall (in mm)
D	:	is the rainfall event duration (in h)
α	:	is a scaling constant (the intercept)
γ	:	is the shape parameter (that defines the slope of the power law curve)
$\Delta\alpha, \Delta\gamma$	:	represent the uncertainties of α, γ, respectively; which measure the variation of the threshold around the tendency line.

As mentioned before the frequentist method allows the calculation of threshold at different probabilities for the α (50% to 0.005%). According to Peruccacci et al. (2012),

the 5% exceedance probability of the E-D threshold is selected, which means that the empirical data points below the curve have not been able to trigger a landslide. On the other hand, the data points above the curve are responsible for triggering a landslide. These data (set of E-D) fit in the logarithmic coordinates. The quality of the thresholds depends on the abundance and the distribution of the empirical data points, so better results are obtained when this method is applied to a large data set covering the range of the examined rainfall duration.

5.6.1 Outputs from phase 6

The output for this phase is the folder (Rainfall thresholds), which contains 4 files: boot MPRC.pdf, boot MRC.pdf, boot MPRC.csv, and boot MRC.csv. All of this information is in an Annex F and detailed as follows.

5.6.1.1 boot MRC and boot MPRC pdf files

These two files summarize the calculation of the rainfall thresholds for the MRC and MPRC conditions defined in Section 5.5. Below is an example of the MRC file for the entire study area. Figure 5.16 a,b,c are the empirical Cumulative Distribution Function (ECDF) for duration, cumulative rainfall, and RRG_distance respectively. These three graphs allow us to observe the distribution of these three variables with respect to the entire data set. Figure 5.16 d is a scatter plot that allows to observe the distribution of the pairs of conditions E-D that are going to be used in the analysis. Figure 5.16 e,f are histograms that show the density of the number of gridded data analyzed (RRG_select), and the number of multiple rainfall conditions (n_MRC) by each landslide. Next, in the file, the result of thresholds (16 threshold curves) for various α exceedance probabilities is presented (50% to 0.005%).

Figure 5.17 is the example of pages two to seventeen that has the summary of the percentage of exceedance, specifically this represents the 50% for the MRC data set for all study area, in sub-figure a) and b) the threshold is observed in linear coordinates and in
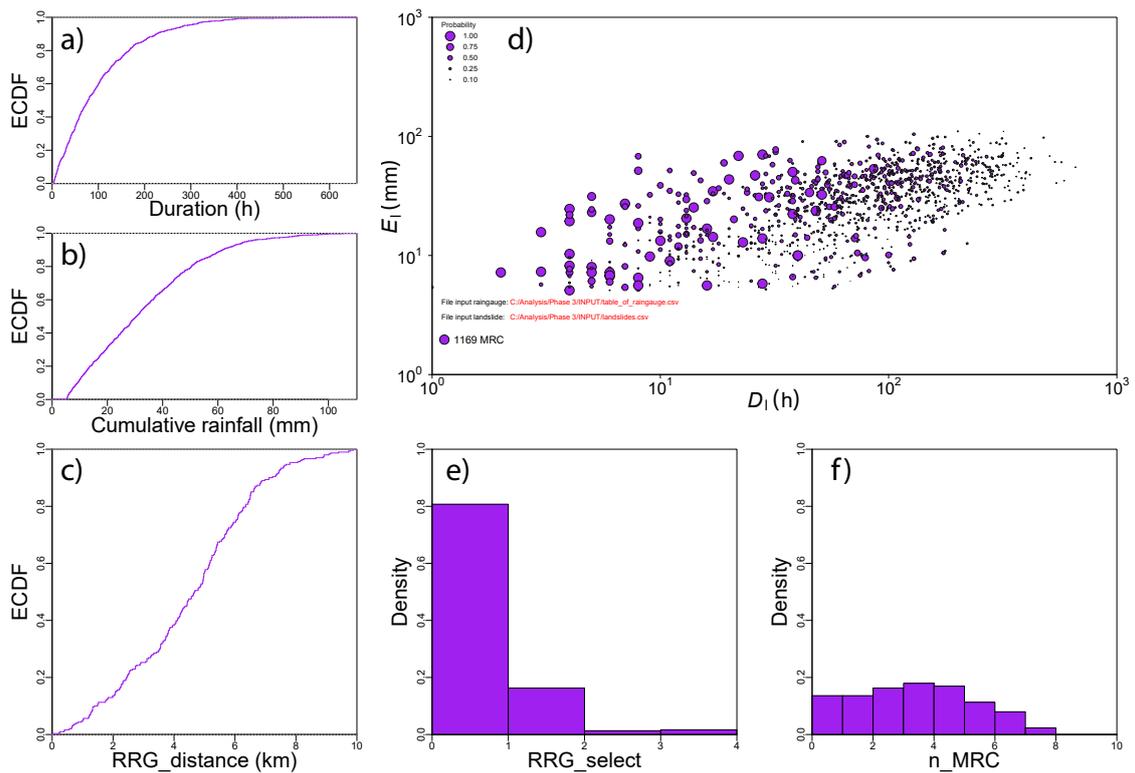
Figure 5.16: Description of data used in the calculation of the rainfall thresholds for the MRC data set (first page). a),b) and c) are the empirical Cumulative Distribution Function (ECDF) for Duration, cumulative rainfall, and RRG_distance respectively. d) Distribution of the pairs of conditions E-D in log-log coordinates. e) Density of the number of gridded data analyzed by the station. f) Density of the number of multiple rainfall conditions by each landslide.

logarithmic coordinates, respectively. These two graphs show the applied model (power law) and the transformation used to perform the analysis. Sub-figure c) and d) show the result of mean α with its error from the bootstrapping and the relative uncertainty $\Delta\alpha$ according to the number of samples respectively. These two graphs show how the α and error varies (decreases) and how the value stabilizes as the sample increases.

Figure 5.18 is the example of the last page of the boot file, this has the threshold of 5% and 1% for the MRC data set for all study area, in sub-figure a) and b) the threshold is observed in linear coordinates and in logarithmic coordinates, respectively. Sub-figure c) and d) show the result of mean γ with its error from the bootstrapping and the relative uncertainty $\Delta\gamma$ according to the number of samples respectively. These two graphs show how the α and error varies (decreases) and how the value stabilizes as the sample increases.
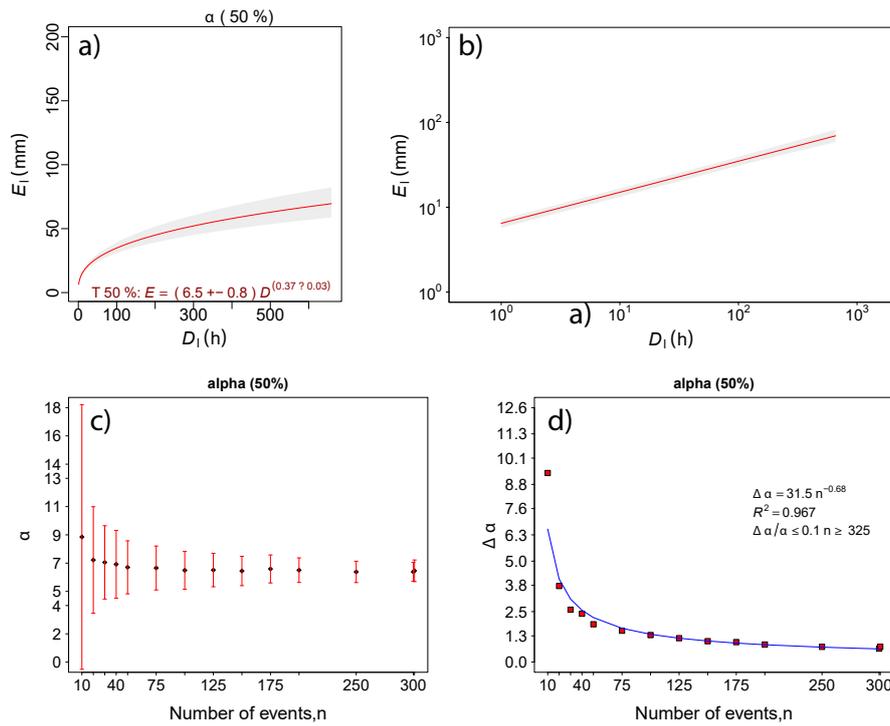
Figure 5.17: Summary of the 50% probability of ascendance (second page). a) Power law in linear coordinates. b)Power law in log-log coordinates. c) the α variable and its error of bootstrapping according to a number of samples. d) relative uncertainty $\Delta\alpha$
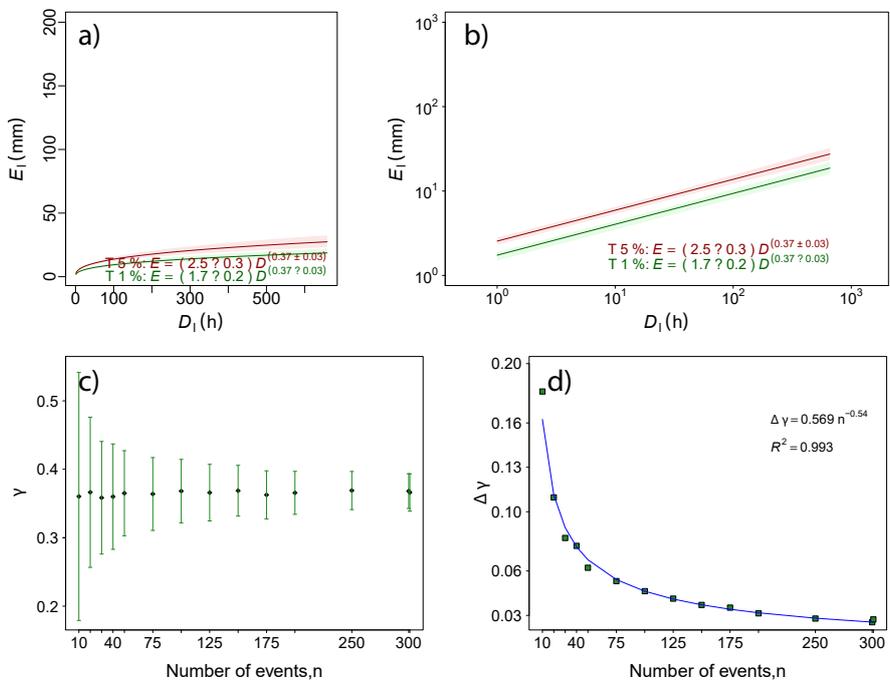
Figure 5.18: 1% and 5% probability of ascendance (last page). a) Power law in linear coordinates. b) Power law in log-log coordinates. c) the γ variable and its error of bootstrapping according to a number of samples. d) relative uncertainty $\Delta\alpha$

5.6.1.2 boot_MRC and boot_MPRC csv files

The csv files (boot_MRC.csv or boot_MPRC.csv) have the values of α and γ that define 16 threshold curves, at different exceedance probabilities (from 0.005 to 50%).

The table has 6 variables: variable name, probability mean, sigma, min, and max (Table 5.9). The probability variable has the corresponding exceedance probabilities for each threshold curve, the mean has the mean value resulting from the bootstrap process (Peruccacci et al., 2012), sigma is the standard deviation value, and min, and max are the extreme values.

Table 5.9: Example of boot_MRC.csv

variable	probability	mean	sigma	min	max
beta	50	0.63394	0.02729	0.61	0.66
alfa	50	6.4606	0.762	5.7	7.2
alfa	35	5.1986	0.605	4.6	5.8
alfa	20	4.0154	0.462	3.6	4.5
alfa	10	3.133	0.359	2.8	3.5
alfa	9	3.0316	0.349	2.7	3.4
alfa	8	2.9232	0.337	2.6	3.3
alfa	7	2.8094	0.324	2.5	3.1
alfa	6	2.684	0.31	2.4	3
alfa	5	2.5492	0.295	2.3	2.8
alfa	4	2.4024	0.279	2.1	2.7
alfa	3	2.2318	0.263	2	2.5
alfa	2	2.0242	0.24	1.8	2.3
alfa	1.5	1.894	0.225	1.7	2.1
alfa	1	1.7374	0.209	1.5	1.9
alfa	0.5	1.5062	0.184	1.3	1.7
alfa	0.005	0.7126	0.098	0.6	0.8

6 | Results

Using the catalog of landslides (354 events), GPM rainfall information (220 series of hourly precipitation from 2013 to 2018), and the method proposed by Brunetti et al. (2010); Peruccacci et al. (2017), implemented in CTRL-T and adjusted for the study area, we obtain 1119 ED conditions (Figure 6.1) that were used to calculate rainfall thresholds for shallow landslides in the northern Andes of Ecuador.

We determine the thresholds according to the following classification.

- One for the study area and one for each province: Pichincha, Imbabura, and Carchi.

- For regional characteristics:

  - One threshold for each one of the five precipitation regions (R8, R16, R6, R17, and R12).

  - One threshold for each one of the six land covers (populated area, shrubby vegetation, native forest, grassland, agricultural mosaic, and herbaceous vegetation)

  - One threshold for each one of the five soil types (andisols, unknown:urban-peri urban areas, mollisols, and Inceptisols)

  - One threshold for each one of the four main lithological domains in the study area (volcanic rocks, alluvial and colluvial deposits, conglomerates, and no data).

  - One threshold for each one of the two Hydrographic Units of level 3 (Esmeraldas and Mira River Basins).

- One threshold for each climatic season: dry and wet

The thresholds are reported in Table 6.1 with the following information: one label that identifies each E-D threshold, the area indicates the specific category for threshold, the events show the number of MRC obtained as output from phase 6 (MRC pdf file), the threshold that shows the formula of E-D threshold for each category, the range of duration, and the uncertainties for α and γ.

Table 6.1: Rainfal ED thresholds for the possible triggering of landslide in North of Ecuador (Pichincha, Imbabura, and Carchi).

#	Label	Area	Events	Threshold	Range	Uncertainty
1	T5, PIC	Pichincha, Imbabura, Carchi	1169	$E = 2.5D^{0.37}$	1<D<650	$\Delta\alpha = 0.3$ $\Delta\gamma = 0.03$
by political division						
2	T5, P	Pichincha	888	$E = 2.3D^{0.38}$	1<D<600	$\Delta\alpha = 0.3$ $\Delta\gamma = 0.03$
3	T5, I	Imbabura	184	$E = 3.4D^{0.37}$	1<D<650	$\Delta\alpha = 1.0$ $\Delta\gamma = 0.07$
4	T5, C	Carchi	97	$E = 4.8D^{0.27}$	1<D<450	$\Delta\alpha = 2.3$ $\Delta\gamma = 0.1$
by climatic season						
5	T5, Wet	Wet	1150	$E = 2.7D^{0.36}$	1<D<650	$\Delta\alpha = 0.3$ $\Delta\gamma = 0.03$
6	T5, Dry	Dry	19	$E = 1.6D^{0.53}$	1<D<75	$\Delta\alpha = 0.6$ $\Delta\gamma = 0.2$
by precipitation Regions						

Continuation of Table 6.1

#	Label	Area	Events	Threshold	Range	Uncertainty
7	T5, R8	R8 (950-2000 mm $year^{-1}$)	705	$E = 2.5D^{0.37}$	1<D<650	$\Delta\alpha = 0.4$ $\Delta\gamma = 0.03$
8	T5, R16	R16 (500-1000 mm $year^{-1}$)	157	$E = 2.1D^{0.39}$	1<D<350	$\Delta\alpha = 0.5$ $\Delta\gamma = 0.06$
9	T5, R6	R6 (2100-3300 mm $year^{-1}$)	164	$E = 3.9D^{0.33}$	1<D<550	$\Delta\alpha = 1.0$ $\Delta\gamma = 0.06$
10	T5, R17	R17 (450-1100 mm $year^{-1}$)	73	$E = 3.8D^{0.34}$	1<D<350	$\Delta\alpha = 1.8$ $\Delta\gamma = 0.1$
11	T5, R12	R12 (550-1400 mm $year^{-1}$)	55	$E = 3.1D^{0.38}$	1<D<370	$\Delta\alpha = 3.3$ $\Delta\gamma = 0.2$
by Land cover						
12	T5, Pa	Populated area	377	$E = 2.1D^{0.41}$	1<D<600	$\Delta\alpha = 0.4$ $\Delta\gamma = 0.04$
13	T5, Sv	Shrubby vegetation	219	$E = 2.4D^{0.36}$	1<D<650	$\Delta\alpha = 0.6$ $\Delta\gamma = 0.05$
14	T5, Nf	Native Forest	184	$E = 3.5D^{0.3}$	1<D<550	$\Delta\alpha = 1.4$ $\Delta\gamma = 0.08$
15	T5, G	Grassland	136	$E = 2.9D^{0.45}$	1<D<450	$\Delta\alpha = 1.3$ $\Delta\gamma = 0.09$
16	T5, Am	Agricultural mosaic	107	$E = 2.9D^{0.35}$	1<D<400	$\Delta\alpha = 1.1$ $\Delta\gamma = 0.09$
17	T5, Hv	Herbaceous vegetation	42	$E = 4.5D^{0.35}$	1<D<300	$\Delta\alpha = 1.8$ $\Delta\gamma = 0.09$

Table 6.1 continue on next page

Continuation of Table 6.1

#	Label	Area	Events	Threshold	Range	Uncertainty
by soil type						
18	T5, A	Andisol	398	$E = 3.1D^{0.35}$	1<D<550	$\Delta\alpha = 0.5$
						$\Delta\gamma = 0.04$
19	T5, Un	Unknown: urban -peri-urban areas	326	$E = 1.8D^{0.43}$	1<D<600	$\Delta\alpha = 0.5$
						$\Delta\gamma = 0.05$
20	T5, M	Mollisol	233	$E = 2.2D^{0.4}$	1<D<375	$\Delta\alpha = 0.6$
						$\Delta\gamma = 0.06$
21	T5, E	Entisol	73	$E = 5.9D^{0.21}$	1<D<350	$\Delta\alpha = 2.7$
						$\Delta\gamma = 0.1$
by Lithology						
22	T5, VO	Volcanic rock	917	$E = 2.8D^{0.36}$	1<D<650	$\Delta\alpha = 0.4$
						$\Delta\gamma = 0.03$
23	T5, AD	Alluvial and Colluvial Deposits	157	$E = 2.3D^{0.38}$	1<D<500	$\Delta\alpha = 0.6$
						$\Delta\gamma = 0.07$
24	T5, TI	undifferentiated terraces	33	$E = 4.8D^{0.24}$	1<D<350	$\Delta\alpha = 3.0$
						$\Delta\gamma = 0.2$
25	T5, CL	Conglomerates	47	$E = 4.2D^{0.37}$	1<D<300	$\Delta\alpha = 5.8$
						$\Delta\gamma = 0.2$
by hydrographic Units level 3						
26	T5, Erb	Esmeraldas River Basin	952	$E = 2.4D^{0.37}$	1<D<650	$\Delta\alpha = 0.3$
						$\Delta\gamma = 0.03$
27	T5, Mrb	Mira River Basin	206	$E = 3.4D^{0.34}$	1<D<475	$\Delta\alpha = 0.9$
						$\Delta\gamma = 0.07$

End of Table

6.1 Rainfall thresholds for north of Ecuador

Knowledge of landslide thresholds in the study area (three provinces) is an important step towards understanding the rainfall conditions necessary to trigger this hydro-geomorphological phenomenon. Figure 6.1 shows the distribution of multiple rainfall conditions (D-E) responsible for slope failure for landslides in Northern Ecuador (black dots). The higher concentration of (D, E) combinations is in the range of $1 \geq E \leq 50(mm)$ and $1 \geq D \leq 200(h)$.
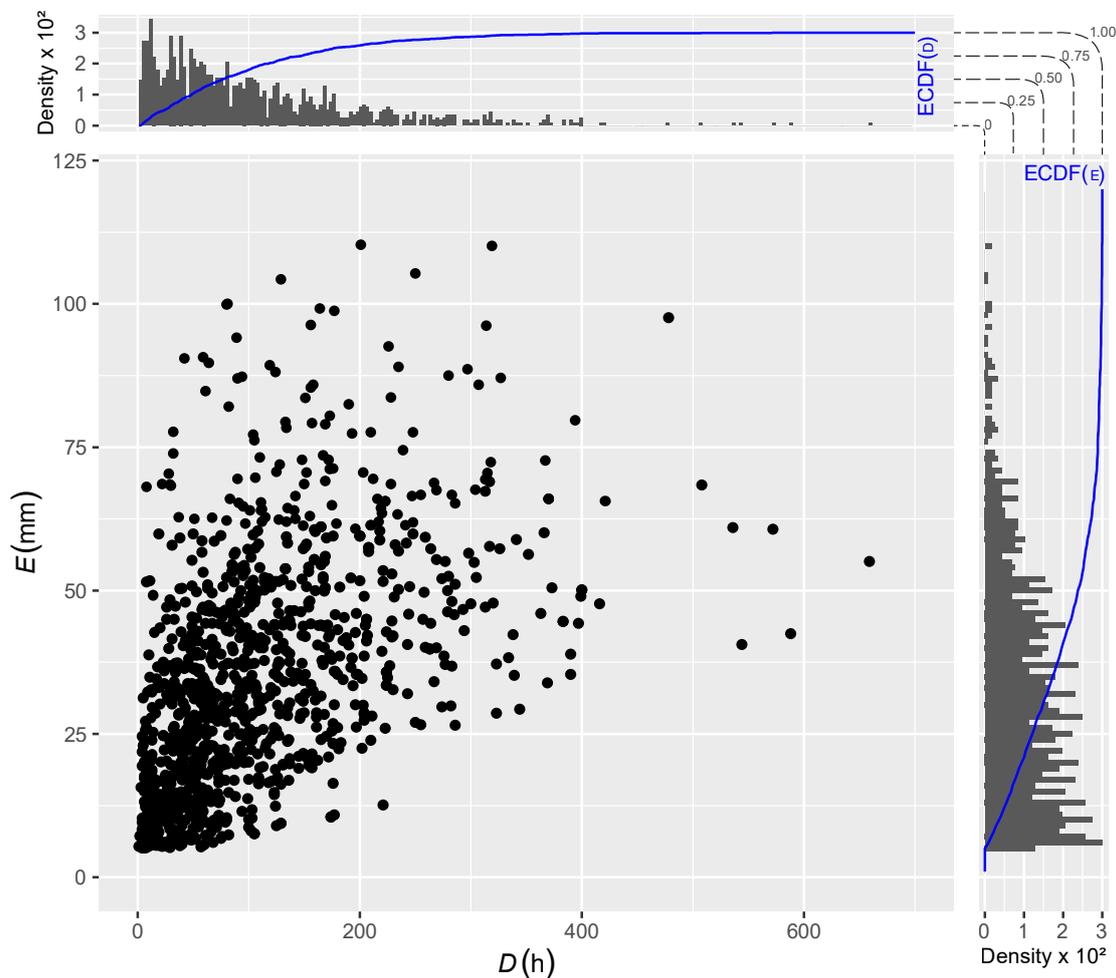
Figure 6.1: Rainfall duration D (x-axis) in hours and cumulated event rainfall E (y-axis) in mm conditions that have resulted in landslides in the area of study from January 2013 to February 2018 (1169 back dots). Data is shown in linear coordinates. The upper and right (histograms) show marginal distributions of D and E and their empirical cumulative distribution functions ECDF (blue lines).

Moreover, the histograms in the upper part indicate the density of rainfall duration (D). It represents the occurrence of each condition. When the duration decreases the density also decreases (density $\times 10^2$). On the other hand, on the right side, the histogram indicates the density $\times 10^2$ of cumulated rainfall events (E). For large values of cumulated event rainfall the count of combinations decrease (density $\times 10^2$). The empirical cumulated distribution functions (ECDF, blue line) and their percentage (dashed lines) are also shown.

The distribution of the entire set of multiple rainfall conditions (D, E) associated with landslides (1169 events, gray dots) in the northern part of Ecuador (Pichincha, Imbabura, and Carchi) for five years is shown in Figure 6.2a, where the black line represents the 5% E-D threshold for PIC and the shaded area indicates the uncertainties $(\Delta\alpha = 0.3, \Delta\gamma = 0.03)$ associated with the threshold reported in the Table 6.1. The threshold value of $T5, PIC$ has an intercept value of $\alpha = 2.5 \pm 0.3$ and the slope value is $\gamma = 0.37 \pm 0.03$. Around 58 events are below the threshold, which represents the 5% in the range of $1 \geq D \leq 200$ (vertical dashed lines).

Moreover, in the logarithmic coordinates (Figure 6.2b) the range of duration where the combination of MRC is responsible for triggering a landslide (above the black line) is concentrated between $10^{1.5} \geq D \leq 10^{2.5}$ (between the dashed lines). The values of uncertainties are shown in Figures 6.2c, d ($\Delta\alpha$, red bars, and $\Delta\gamma$, dark red bars) in relation to the number of events, where the black dots represent the mean values of α and γ. In general, as the number of events increases the values of α and γ decrease to 2.5 ± 0.3 and 0.37 ± 0.03 respectively.
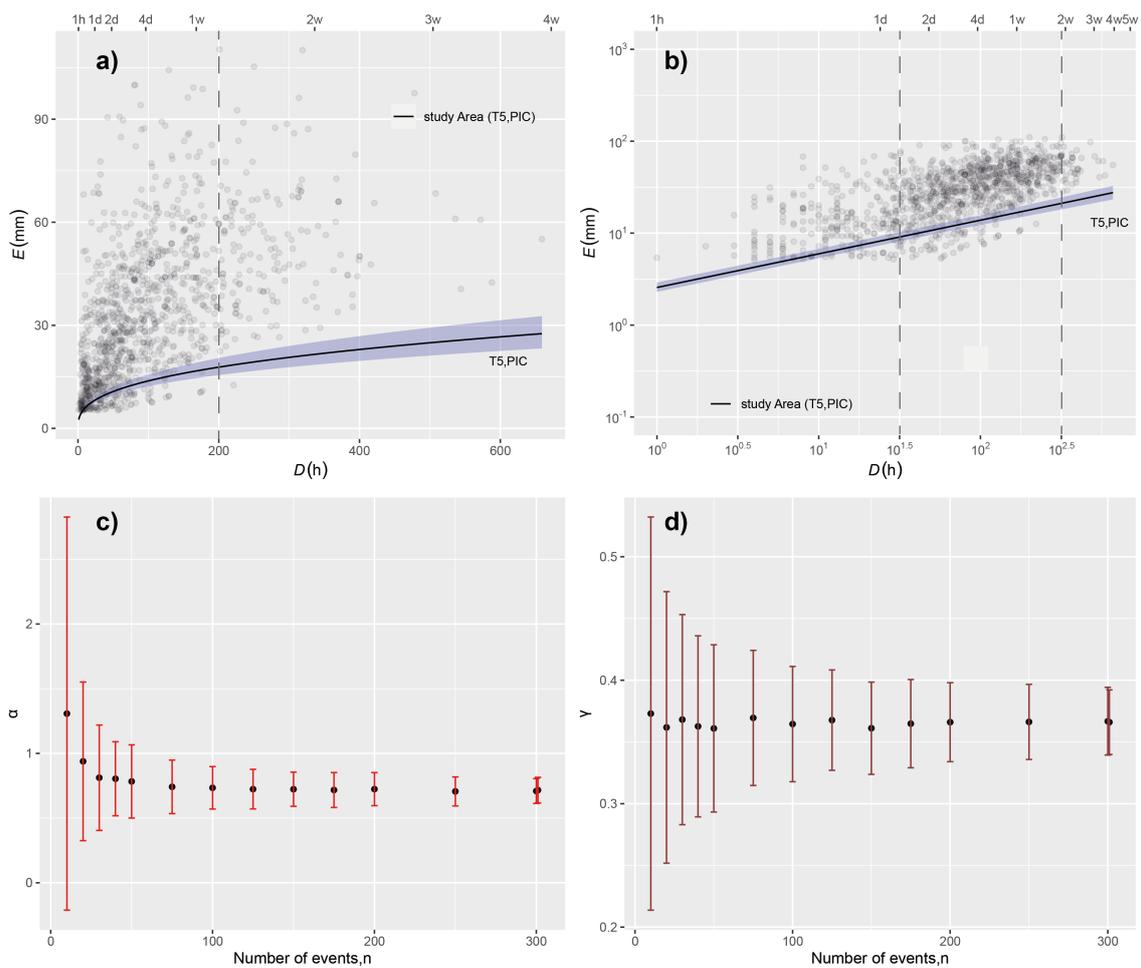
Figure 6.2: a) Rainfall duration D (x-axis) and cumulated event rainfall E (y-axis) conditions that have resulted in landslides (301 events) in the area of study, Pichincha, Imbabura, and Carchi (PIC). The black line is the 5% E-D threshold for the PIC area (T5, PIC). Data is shown in linear coordinates, b) The T5 threshold (black line) with a shaded area (light blue) showing uncertainty around the threshold curve. Data is shown in log-log coordinates. c) Represents the variation of α (intercept) and its associated uncertainty $\Delta\alpha$ (red bars) based on the number of events. d) Indicates the variation of γ (scaling exponent) and the associated uncertainty $\Delta\gamma$ (dark red bars) based on a number of events.

6.2 Rainfall thresholds by provinces

It is important to analyze the E-D threshold for each province, Figure 6.3b,c indicates the location of landslide events with respect to elevation. The highest incidence of landslide events is concentrated in altitudes lower than 2500 m (56% of landslide events), and most of the landslides from Carchi occur in the boundary with Imbabura. As we mentioned before the province with more landslide events is Pichincha (228), then Imbabura (51),

and finally Carchi (22).

The threshold information is in Table 6.1 where the province with small uncertainties values is Pichincha ($\Delta\alpha = 0.3$ and $\Delta\gamma = 0.03$), then Imbabura has $\Delta\alpha = 1.0$ and $\Delta\gamma = 0.07$ and finally Carchi with the highest values of $\Delta\alpha = 2.3$ and $\Delta\gamma = 0.1$.
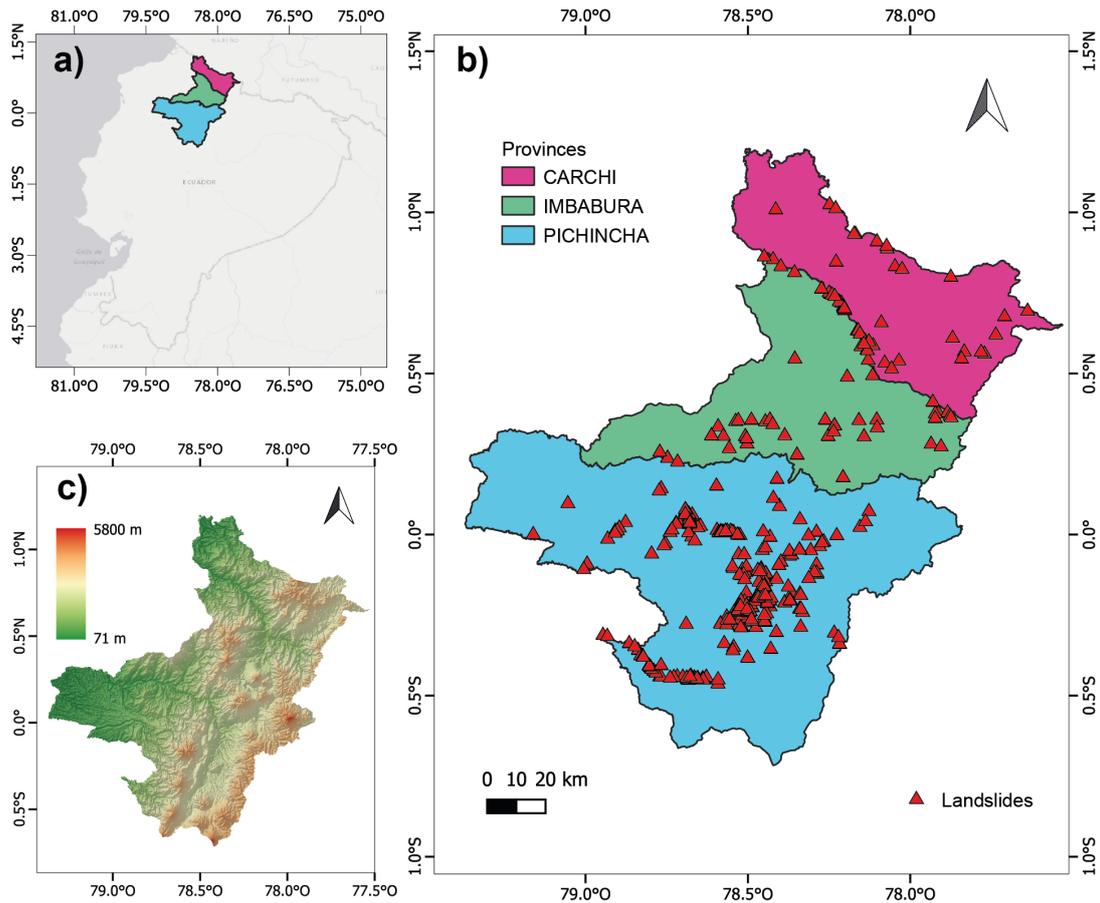
Figure 6.3: a) Map showing the location of the study area, Pichincha, Imbabura, and Carchi(PIC). b) Map showing the location of landslides for each province. The legend shows the name of the province in northern Ecuador. Red triangles show landslides in the area. c) Altitude in the study area with precipitation regions.

In Figure 6.4a each threshold is plotted in log space, where the threshold for Pichincha (T5, P - light blue line) has a trend similar to the threshold T5, PIC (black dashed line). On the other hand, the Imbabura (T5, I - green line) and Carchi (T5, C - red line) thresholds have a pretty similar tendency to exceed the threshold T5, PIC, which means that in these provinces the rainfall conditions need to be higher to trigger a landslide.

However, in the end, the red line coincides with T5, PIC.

In Figure 6.4b, the thresholds are plotted in linear coordinates, and the uncertainties are represented by the shaded area. As we mentioned before Pichincha (T5, P) is the province with the smallest values of $\Delta\alpha$ and $\Delta\gamma$ and the range of rainfall duration is $1 \geq D \leq 600(h)$. The Imbabura threshold (T5, I) has a large shaded area (green) with the same rainfall duration range as T5, PIC ($1 \geq D \leq 650(h)$), but it exceeds 30 mm of cumulated rainfall (E). Finally, the Carchi threshold has a large shaded area (red zone) with a lower range of rainfall duration ($1 \geq D \leq 450(h)$), but in the end, it tends to be similar to T5, PIC.
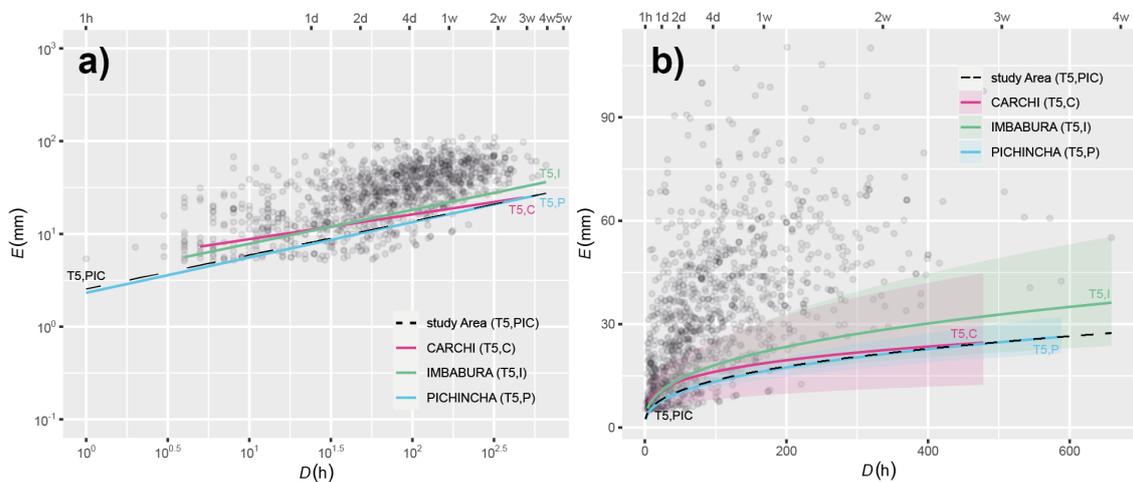
Figure 6.4: a) Rainfall duration D (in hours) and cumulated event rainfall E (in mm) conditions have resulted in landslides in the PIC area (gray dots). Data is shown in log-log coordinates. The lines represent 5% E-D thresholds for the three provinces. The study area thresholds T5, PIC (black line). The Pichincha T5,P (blue line).The Imbabura T5,I (green line).The Carchi T5,C (red line). b) E-D conditions have resulted in landslides in the PIC area (gray dots), and 5% E-D threshold lines for the three provinces and their uncertainties (shaded areas). Data is shown in linear coordinates.

6.3   Rainfall thresholds by climatic season

The determination of the climatic season for the study area is explained in Section 5, where for each precipitation region, the year-monthly precipitation distribution of the gridded data was analyzed to determine the wet and dry seasons. In Figure 6.5a the

threshold is plotted in linear coordinates, where the wet season (T5, Wet - orange curve) matches the T5, PIC (black dashed curve), also the rainfall duration is similar to the threshold T5, PIC. The orange shaded area indicates the uncertainty zone for the wet season and the values ($\Delta\alpha$ and $\Delta\gamma$) are reported in Table 6.1 0.3 and 0.03 respectively. On the other hand, for the dry season, the uncertainty values are higher $\Delta\alpha = 0.6$ and $\Delta\gamma = 0.2$, and the rainfall duration range is shorter (1<D< 70). In Figure 5.7b the log-log plot shows the trend of threshold T5, Dry (red line), it starts at rainfall duration of $10^{0.5}$ to $10^{1.75}$ and increases until 10 mm of cumulated rainfall (it crosses the threshold T5, PIC - black dashed line).
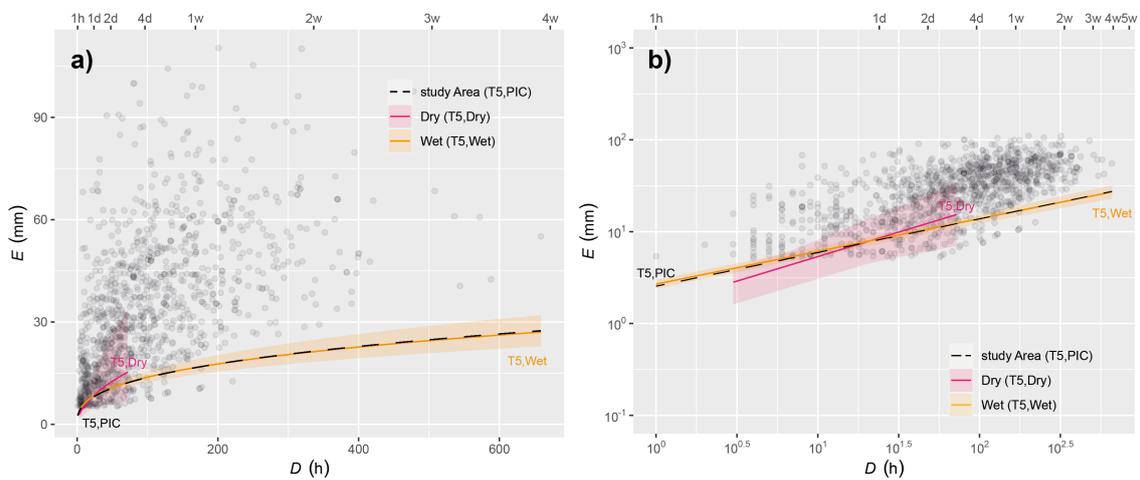
Figure 6.5: a) Rainfall duration D (in hours) and cumulated event rainfall E (in mm) conditions have resulted in landslides in the PIC area (gray dots). Data is shown in linear coordinates. The lines represent 5% E-D thresholds for the two climatic seasons (dry and wet). The study area thresholds T5, PIC (black line). The Dry season T5, Dry (pink line). The Wet season T5, Wet (orange line). And their uncertainties (shaded areas). b) E-D conditions that have resulted in landslides in the PIC area (gray dots), and 5% E-D threshold lines for the two seasons and their uncertainties (shaded areas). Data is shown in log-log coordinates.

6.4 Rainfall thresholds by precipitation Regions

The identification of homogeneous precipitation regions in Ecuador proposed by Ilbay-Yupa et al. (2021) is necessary to know the behavior of precipitation in the study area. Here, six precipitation regions cover the area as shown in Figure 6.6a. The precipitation

regions are reported on the map legend according to the incidence of landslides events from higher to lower, so the R8 (950-2000 mm $year^{-1}$) contains the largest number of landslide events (173) all the way to R7 (500-1000 mm $year^{-1}$) with the lowest number of events reported, for this reason, the determination of rainfall threshold were done up to R12 (550-1400 mm $year^{-1}$), which has 12 events of landslides and is above the threshold line in Figure 6.6b. Moreover, based on Figure 6.6c in each precipitation region most of the landslides events occur in areas with low elevation ($\sim 71m$), also in the regions with more events (R8, and R16) the landslides are concentrated on administrative boundaries and occurred in altitudes lower than 2500 m (56% of landslide events).
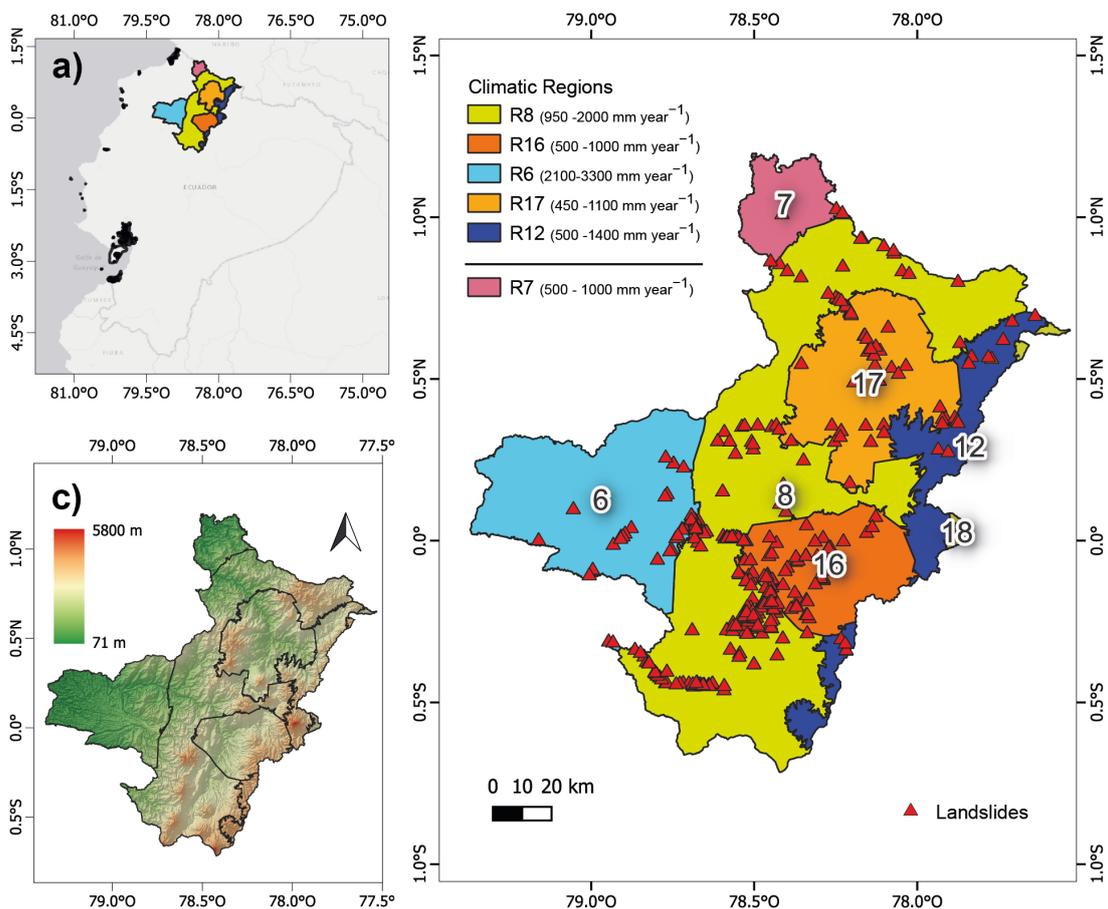
Figure 6.6: a) Map showing the location of the study area, Pichincha, Imbabura, and Carchi(PIC). b) Map showing the subdivision of PIC in six climatic regions, modified from a map published by Ilbay-Yupa et al. (2021). The legend shows the name of the climatic region and the annual precipitation range. Red triangles show landslides in the area. c) Altitude in the study area with precipitation regions.

The E-D threshold for each precipitation region is reported in Table 6.1, which indicates the threshold values and the respective uncertainties. Moreover, the region with small uncertainties values ($\Delta\alpha = 0.4, \Delta\gamma = 0.03$) is R8 (yellow line) and adjusts to the threshold of the entire zone (T5, PIC - black dashed line) as shown in Figure 6.7a the log-log coordinates allow to visualize in a better way the distribution of landslides. Another region with a similar tendency to T5, PIC is T5, R16 (orange line) with uncertainties values equal to $\Delta\alpha = 0.5, \Delta\gamma = 0.06$ (narrow zone). On the other hand, the other regions (R6, R17, and R12) have a similar tendency and are above T5, PIC, which means that the rainfall conditions to trigger a landslide in these regions need to be higher.

In Figure 6.7b the threshold curve of each precipitation region is plotted with the respective shaded areas (uncertainty zone). The variation of rainfall duration of the threshold depends on the grid precipitation data associated with each landslide. Furthermore, the regions (R6, R17, and R12) that need higher rainfall conditions have a high-shaded area with respect to the others. From a general point of view, the shaded area rises when the landslide events decrease, for this reason, the R12 with 12 events has the largest uncertainty zone (blue shaded area). On the other hand, the threshold T5, R8 ($1 \geq D \leq 650(h)$) has a small shaded area and is mostly fitted to T5, PIC. The difference to the T5, R16 is the rainfall duration ($1 \geq D \leq\sim 400(h)$).
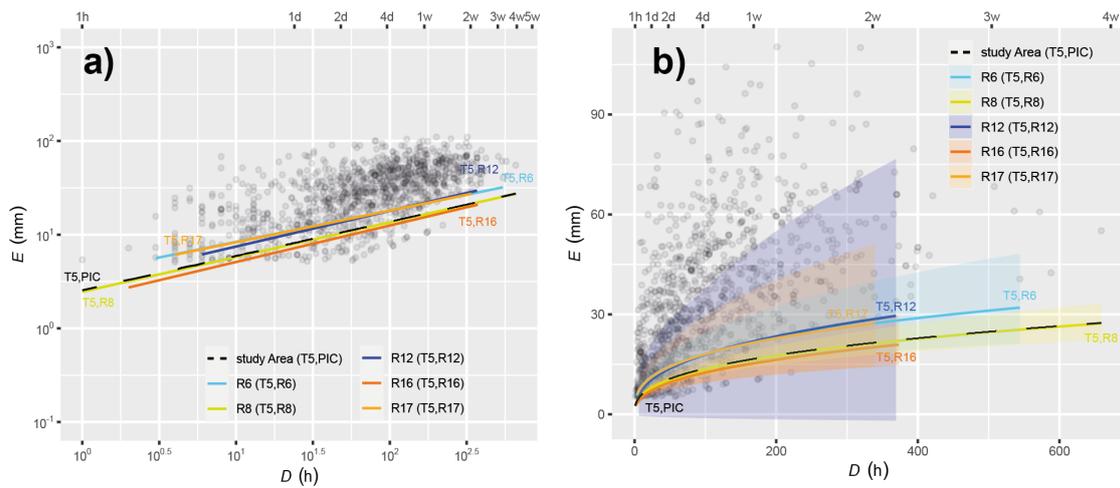
Figure 6.7: a) Rainfall duration D (in hours) and cumulated event rainfall E (in mm) conditions have resulted in landslides in the PIC area (gray dots). Data is shown in log-log coordinates. The lines represent 5% E-D thresholds for the six climatic regions. Study area thresholds T5, PIC (black line). R6 region T5, R6 (light blue line). R8 region T5, R8 (yellow line). R12 region T5, R12 (blue line). R16 region T5, R16 (orange line). R17 region T5, R17 (light orange line). b) E-D conditions have resulted in landslides in the PIC area (gray dots), and 5% E-D threshold lines for the six climatic regions and their uncertainties (shaded areas). Data is shown in linear coordinates.

6.5   Rainfall thresholds by Land cover

The land cover is relevant to determine specific characteristics related to the water storage capacity of each vegetation cover in the study area. The location of landslide events related to land cover is in Figure 6.8a,b; where the populated area (purple zone) is at the lowest elevation and contains most of the landslide events (95 events), which was analyzed to determine the E-D threshold. Then, the shrubby vegetation (light brown zone) located in the Western cordillera contains a high incidence of landslide events (58 events). The native forest (green zone) is located along the western part of the study area. The grassland (blue zone) and agricultural mosaic (red zone) are placed along the Eastern Cordillera and Inter Andean Valley. Another part is located near the neighborhood of Pedro Vicente Maldonado (Pichincha). The herbaceous vegetation (yellow zone) is located along the lower elevations (Figure 6.8c) of the Inter Andean Valley. The remaining categories (below the threshold line) were not used to determine the threshold due to the

lack of landslide events.
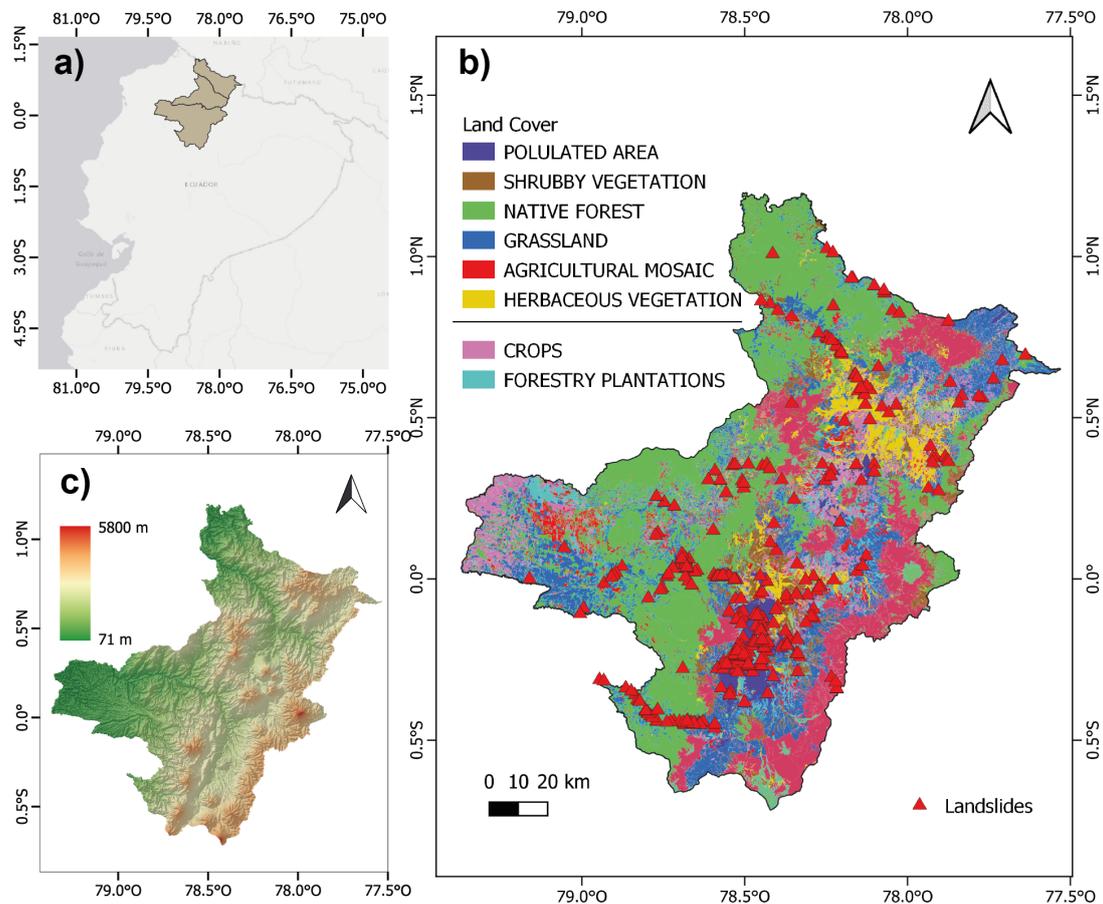
Figure 6.8: a) Map showing the location of the study area, Pichincha, Imbabura, and Carchi(PIC). b) Map showing the subdivision of PIC in eight land cover domains, modified from a map published by MAG 2020. The legend shows the name of the types of soil domains. Red triangles show landslides in the area. c) Altitude in the study area.

In Figure 6.9a the log-log plot shows the E-D threshold tendency of each land cover. The threshold of populated area (T5, Pa - purple line) and shrubby vegetation (T5, Sv - brown line) have a similar tendency to T5, PIC (black dashed line), but they begin at different values of rainfall duration, and the cumulated rainfall, increasing to match the black dashed line. On the other hand, the threshold for native forest (T5, Nf -), grassland (T5, G - blue line), and herbaceous vegetation (yellow line) start at higher values than T5, PIC. The rainfall duration is smallest in each case and only the native forest matches the black dashed line, the other (blue and yellow line) increases to obtain higher values

of cumulated rainfall and rainfall duration. Therefore, the necessary rainfall conditions to trigger a landslide needs to be higher than in the other zones.

In Figure 6.9b, in the linear coordinates the E-D threshold curve for each land cover is plotted with the respective shaded area (uncertainty zone), the values of $\Delta\alpha$ and $\Delta\gamma$ for each one are reported on Table 6.1. The lower shaded area of the E-D threshold for land cover is the purple one (T5, Pa), in which the uncertainties values are $\Delta\alpha = 0.4, \Delta\gamma = 0.04$, and the range of rainfall duration up to 600 (h). T5, Sv (brown curve) is the other with lower uncertainties values ($\Delta\alpha = 0.6, \Delta\gamma = 0.05$) and the rainfall duration range is similar to T5, PIC ($1 \geq D \leq\sim 650(h)$). The uncertainties values of the remaining land cover are higher, in the case of $\Delta\alpha$ the values are > 1 and for $\Delta\gamma > 0.08$.
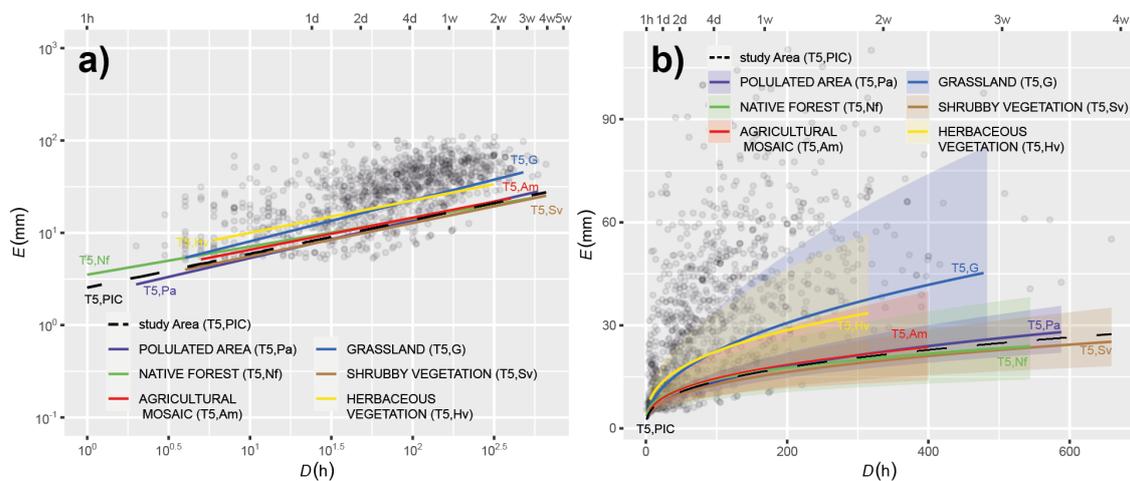
Figure 6.9: a) Rainfall duration D (in hours) and cumulated event rainfall E (in mm) conditions that have resulted in landslides in the PIC area (gray dots). Data is shown in log-log coordinates. The lines represent 5% E-D thresholds for the six land cover domains (with more than 10 landslide events). Study area thresholds T5, PIC (dashed black line). Populated area T5, Pa (purple line). Native forest T5, Nf (green line). Agricultural mosaic T5, Am (red line). Grassland T5, G (blue line). Shrubby vegetation T5, Sv (brown line). Herbaceous vegetation T5, Hv (yellow line). b) E-D conditions that have resulted in landslides in the PIC area (gray dots), and 5% E-D power law threshold lines for the six land cover domains and their uncertainties (shaded areas). Data is shown in linear coordinates.

6.6 Rainfall thresholds by Type of soil

The identification of the type of soil in the study area is essential to know the features of the regolith that might be prone to mechanical failure. In Figure 6.10a and b the map indicate all types of soil belonging to the study area. Where andisols (pink area) are located around the entire area from lower elevations (at the western part) to higher elevations (going to the east). Moreover, the unknown: urban peri-urban areas (red) are located in areas with lower elevations, and the incidence of landslide events is huge. Then, mollisols are located in the eastern part of the study area. Finally, entisols are scattered across the entire study area. For the remaining soil types, the E-D threshold was not calculated due to the poor landslide population (incidence of events).
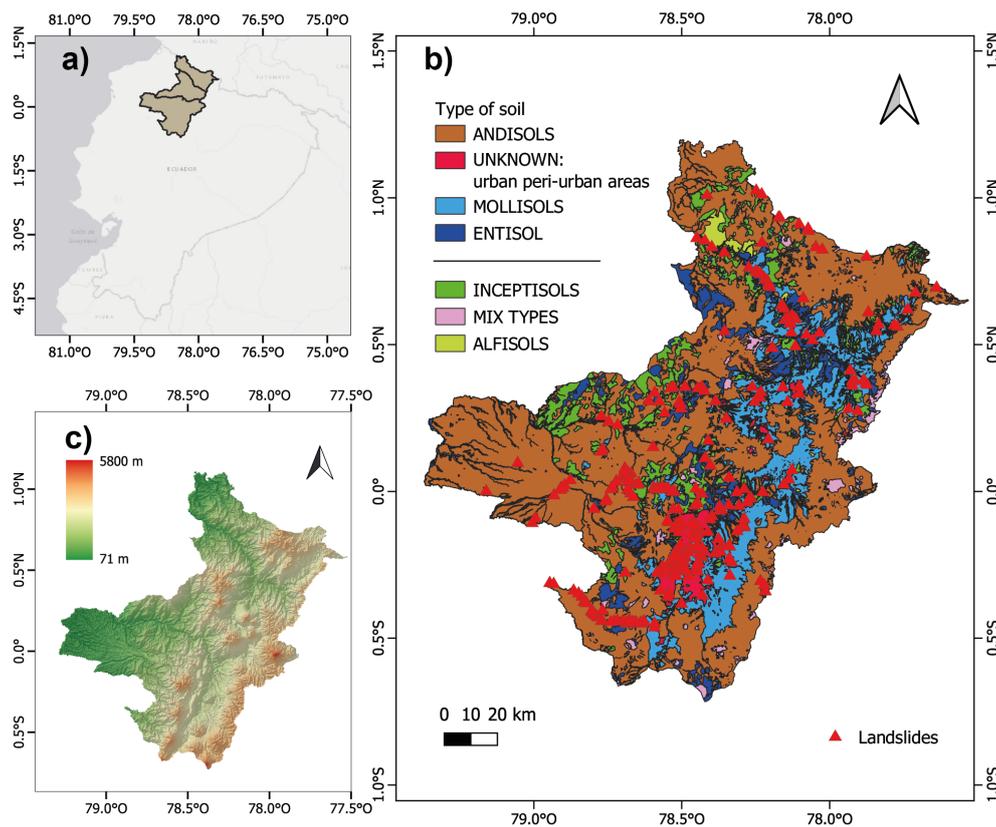
Figure 6.10: a) Map showing the location of the study area, Pichincha, Imbabura, and Carchi(PIC). b) Map showing the subdivision of PIC in nine types of soil domains, modified from a map published by MAGAP 2019. The legend shows the name of the types of soil domains. Red triangles show landslides in the area. c) Altitude in the study area.

In Figure 6.11a,b the thresholds of each type of soil are plotted in the log-log space, where the T5, A (Andisols - brown line) has a pretty similar tendency to the T5, PIC (black dashed line), but the Andisols threshold starts with 5 mm of cumulated rainfall (E). Moreover, the thresholds for Unknown: urban peri-urban areas (T5, Un - red line) and Mollisols (T5, M - light blue) start at $5(mm)$ and $10^{0.5}(h)$ and increase until it matches the black dashed line. The threshold of entisols (T5, E - blue line) starts with $\sim 10mm$ and $\sim 10^{0.5}(h)$ and increases until $\sim 35mm$ and $\sim 10^{2.5}(h)$

Moreover, in Figure 6.11b the plot with linear coordinates indicates the thresholds with the shaded area (uncertainty), where the Andisols (T5, An - brown curve), Unknown: urban peri-urban areas (T5, Un - red curve) and Mollisols (T5, M - light blue curve) have lower values of $\Delta\gamma = 0.5$, $\Delta\alpha = 0.04$; $\Delta\gamma = 0.5$, $\Delta\alpha = 0.05$ and $\Delta\gamma = 0.6$, $\Delta\alpha = 0.06$ (smaller shaded zone), respectively. On the other hand, the Entisol (T5, E - blue curve) has higher values of uncertainties ($\Delta\gamma = 2.7$, $\Delta\alpha = 0.1$).
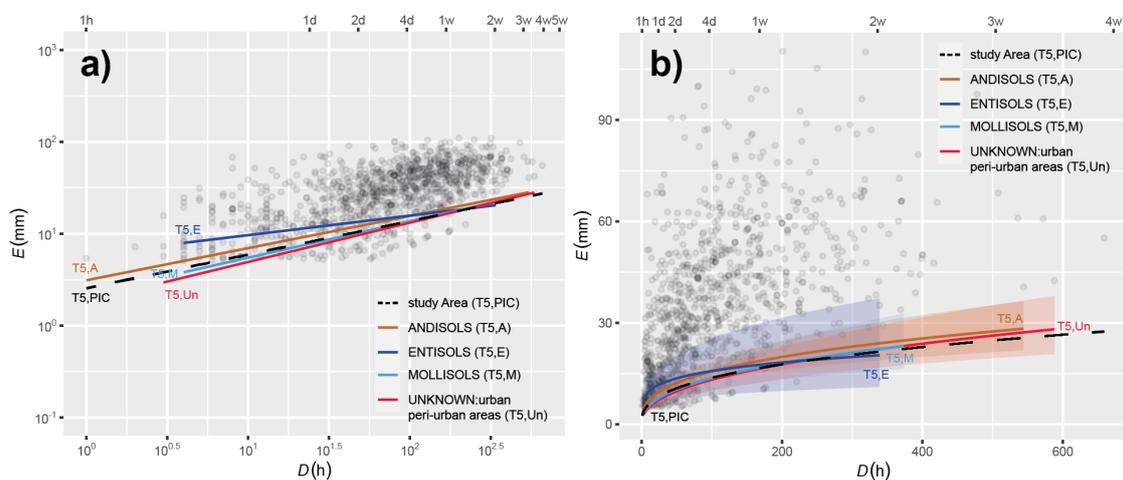
Figure 6.11: a) Rainfall duration D (in hours) and cumulated event rainfall E (in mm) conditions that have resulted in landslides in the PIC area (gray dots). Data is shown in log-log coordinates. The lines represent 5% E-D thresholds for the five types of soil domain (with more than 10 landslide events). Study area thresholds T5,PIC (dashed black line). Andisol T5,A (brown line). Entisol T5,E (blue line). Mollisols T5,M (light blue line). Unknown: urban, peri-urban areas T5,RNa (red line). b) E-D conditions that have resulted in landslides in the PIC area (gray dots), and 5% E-D power law thresholds lines for the five types of soil domains and their uncertainties (shaded areas). Data is shown in linear coordinates.

6.7   Rainfall thresholds by Lithology

The determination of the E-D threshold about the lithological information of the study area is important to identify the rainfall conditions to trigger a landslide based on the geological context. In Figure 6.12a,b the area of the main lithology classification is shown with the location of the landslide. Where the volcanic rock (VO-blue area) is located mainly in the entire study area with the highest incidence of landslide events (230 landslides for the determination of the E-D threshold). Then, alluvial and colluvial deposits (AD - violet area) in the western part of Pichincha with 40 landslide events were analyzed. The undifferentiated terraces (TI - purple area) are located in the Inter Andean Valley. Then, the conglomerates group (CL - green zone) is located in the northwest part of the study area. The remaining groups (below the line) were not used for threshold identification due to the lack of landslide events.
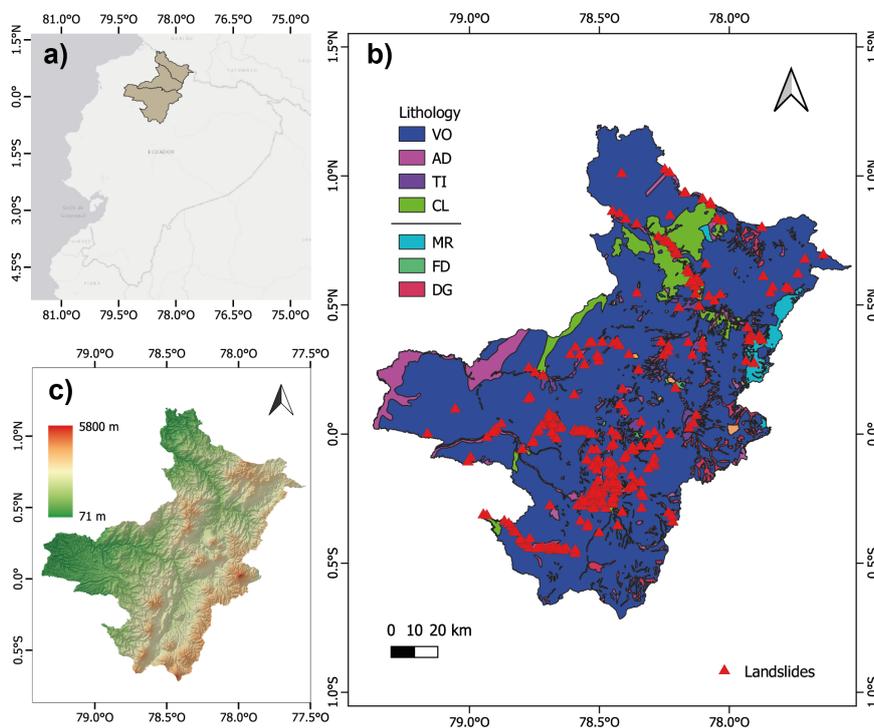
Figure 6.12: a) Map showing the location of the study area, Pichincha, Imbabura, and Carchi(PIC). b) Map showing the subdivision of PIC in nine types of soil domains, modified from a map published by MAGAP 2019. The legend shows the name of the types of soil domains. Red triangles show landslides in the area. c) Altitude in the study area.

In Figure 6.13a the E-D thresholds for lithology are plotted in log-log space, where the threshold for volcanic rocks (T5, VO - blue line) has a similar tendency to T5, PIC (black dashed line). Then, alluvial and colluvial deposits (T5, AD - violet line) start at $(10^{0.5}(mm), 10^{0.5}(h))$ increasing to matches the black dashed line. The undifferentiated terraces (T5, TI - purple line) start at $(\sim 10^{0.8}(mm), \sim 10^{0.7}(h))$ increasing to cross the black dashed line. The conglomerate (T5, CL - green line) requires higher rainfall conditions for triggering a landslide (above the black dashed line).

Moreover, in Figure 6.13b the E-D thresholds are plotted in linear coordinates showing the curves with the uncertainty zone (shaded area), the uncertainties values for each lithological division are reported in Table 6.1. Where the T5, VO (blue curve) has smaller uncertainty values ($\Delta\alpha = 0.4$ and $\Delta\gamma = 0.03$). Then, the alluvial and colluvial deposits (T5, AD - violet curve) have the uncertainties values $\Delta\alpha = 0.6$ and $\Delta\gamma = 0.07$. On the other hand, the undifferentiated terraces (T5, TI - purple curve) are below the black dashed curve and the uncertainties values are $\Delta\alpha = 3.0$ and $\Delta\gamma = 0.2$. Finally, the E-D threshold for conglomerates (T5, CL - green curve) is above the black dashed curve, and the uncertainties values $\Delta\alpha = 5.8$ and $\Delta\gamma = 0.2$ (big shaded area).
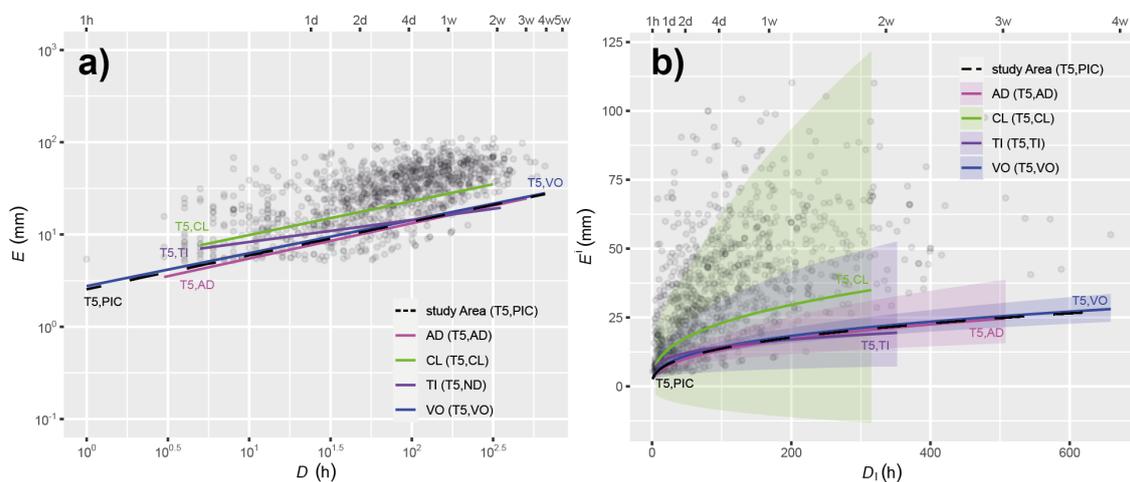
Figure 6.13: a) Rainfall duration D (in hours) and cumulated event rainfall E (in mm) conditions that have resulted in landslides in the PIC area (gray dots). Data is shown in log-log coordinates. The lines represent 5% E-D thresholds at 5% for the four lithology domains; VO: volcanic rock (blue line); AD: alluvial and colluvial deposits (violet); Ti: undifferentiated terraces (purple line); CL: conglomerates (green line).

6.8   Rainfall thresholds by hydrographic unit level 3

In Figure 6.14a,b there are two mainly hydrographic units of level 3 (Esmeralda and Mira River Basin) that cover mostly the entire study area and the incidence of landslides events are high in altitudes lower than 2500 m (56% of landslide events) of the hydrographic units. The hydrographic units with smaller values of landslide events were not analyzed.
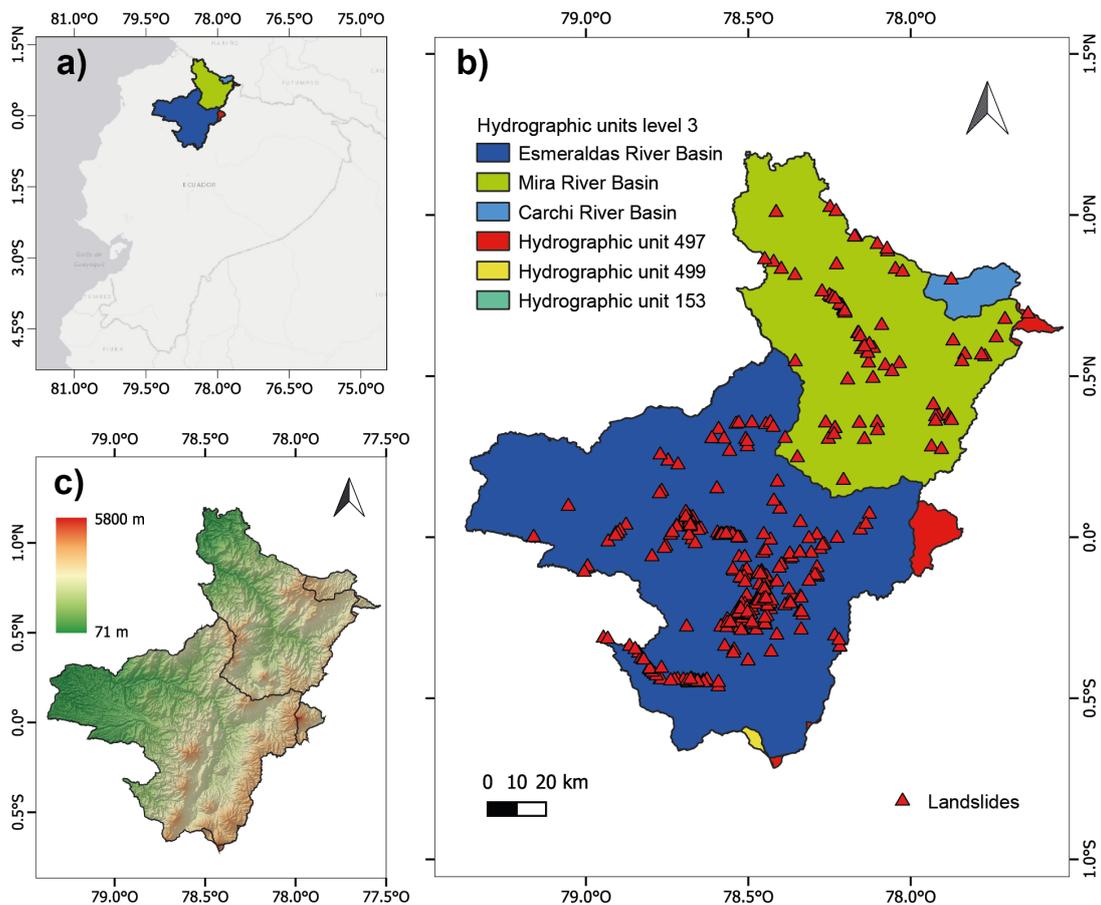
Figure 6.14:   a) Map showing the location of the study area, Pichincha, Imbabura, and Carchi(PIC). b) Map showing the subdivision of PIC into six hydrographic units, modified from a map published by MAATE (2020). The legend shows the name of the six hydrographic units. Red triangles show landslides in the area. c) Altitude in the study area.

For this reason, the determination of E-D threshold was determined for Esmeraldas and Mira River Basin (T5, Erb, and T5, Mrb respectively). In Figure 6.15a, the threshold lines are plotted in log-log space, where the blue line (T5, Erb) matches threshold T5,

PIC (black dashed line). The threshold T5, Mrb (light green line) starts with $\sim 7(mm)$ and $\sim 10^{0.5}(h)$ of the cumulated rainfall (E) increasing until $\sim 45mm$ it differs smoothly from the other. Furthermore, in Figure 6.15b the linear coordinates are useful to know the shorter rainfall duration range to T5, Mrb ($1 \geq D \leq\sim 500(h)$), the shaded area (uncertainty) is higher than T5, Erb and the values are reported in Table 6.1 where $\Delta\alpha = 0.9$ and $\Delta\gamma = 0.07$ and the rainfall duration range is also similar ($1 \geq D \leq\sim 650(h)$).
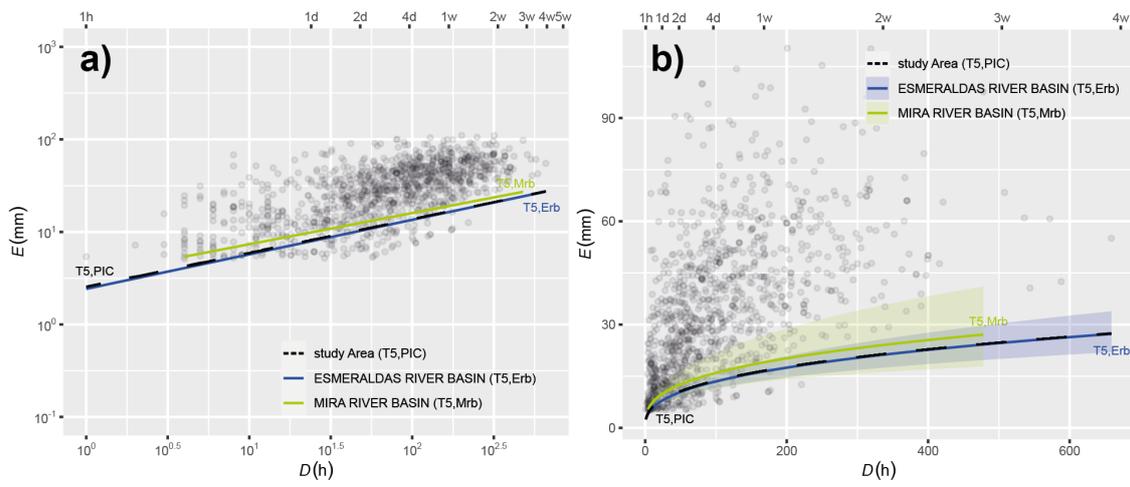
Figure 6.15: a) Rainfall duration D (in hours) and cumulated event rainfall E (in mm) conditions have resulted in landslides in the PIC area (gray dots). Data is shown in log-log coordinates. The lines represent 5% E-D thresholds for the 2 hydrographic units (with more than 10 landslide events). The study area thresholds T5,PIC (dashed black line). The Esmeraldas river basin T5,Erb (blue line). Mira river basin T5,Mrb (green line). b) E-D conditions that have resulted in landslides in the PIC area (gray dots), and 5% E-D threshold lines for the five types of soil domains and their uncertainties (shaded areas). Data is shown in linear coordinates.

7 | Discussion

7.1 Data

The landslide data in this study have various limitations due to the lack of detailed information from the database (DesInventar). This database does not provide a specific format to report the physical characteristics of slope failure or if they are related to each other. The main difference in this work compared to the literature (Peruccacci et al., 2017) is that all landslide events were used without classifying them by time of failure, which is important to know in order to unveil if the cause is the accumulation of rain and not a failure caused by a previous landslide. From the database, two fields were used to discriminate the data: coordinates and comments. From the comments field, it was possible to extract information on the time of occurrence and the magnitude of the event. With the aim of determining rainfall thresholds, the catalog used in this work was created from those two fields.

7.2 Threshold model

In the literature, the most common approach to determine the rainfall thresholds for possible landslide occurrence is the study of intensity–duration (ID) curves (Caine, 1980; Guzzetti et al., 2008). The main reason for this is that this was the model used by Caine (1980) in the first paper that determined global precipitation thresholds for landslide triggering (Peruccacci et al., 2017). To carry out a statistical analysis, one part is the indepen-

dence of the variables, when using the ID model, this independence does not exist since one variable depends on the other. when using the ED model, this problem does not exist (Peruccacci et al., 2017).

7.3   Analysis of the thresholds

In this study, for 27 of the 45 category divisions, the E-D threshold was determined, representing the $\sim 58\%$ reported in Table 6.1, in which the multiple rainfall conditions (dots) vary from 1169 for T5, PIC (maximum) to 19 pairs of rainfall conditions for the T5, Dry (minimum). To understand the E-D thresholds calculated for the study area, they were assigned to a different group with the same tendency, where the T5, PIC (black curve) is the general E-D threshold for the entire study area. In Figure 7.1a,b the different groups of E-D thresholds are plotted on the normal coordinates, but in the second graph, the y-axis is until 125h to visualize better the trend of each group.

The first group has one E-D threshold from land cover (T5, G - blue dashed curve), which increases drastically to high amounts of cumulated rainfall (45mm) and rainfall duration ($\sim 500\ h$). In this way, it has a different trend than other groups. The amount of landslide events analyzed in this group is 33 representing 9.3%. It has higher rainfall conditions to trigger a slope failure. The power law is equal to $E = 2.9D^{0.45}$, and uncertainty values $\Delta\alpha = 1.3\ \Delta\gamma = 0.09$. Geographically, the precipitation regions, which match the area of grassland are R8, R17, and R16; for this reason, rainfall conditions vary. Nevertheless, the precipitation region with a pretty similar power law is R8 ($E = 2.5D^{0.37}$ and $\Delta\alpha = 0.4\ \Delta\gamma = 0.03$). On the other hand, the soil type representative is the mollisols (T5, M) with the power law equal to $E = 2.2D^{0.4}$ and $\Delta\alpha = 0.6\ \Delta\gamma = 0.06$.

Group II has the E-D threshold of different categories, specifically three thresholds for precipitation regions (R12, R6, and R17); one threshold for land cover (HV, herbaceous vegetation); one for lithology (CL, conglomerates). This group increases faster and exceeds the following groups. As there is a large amount of E-D thresholds, the area
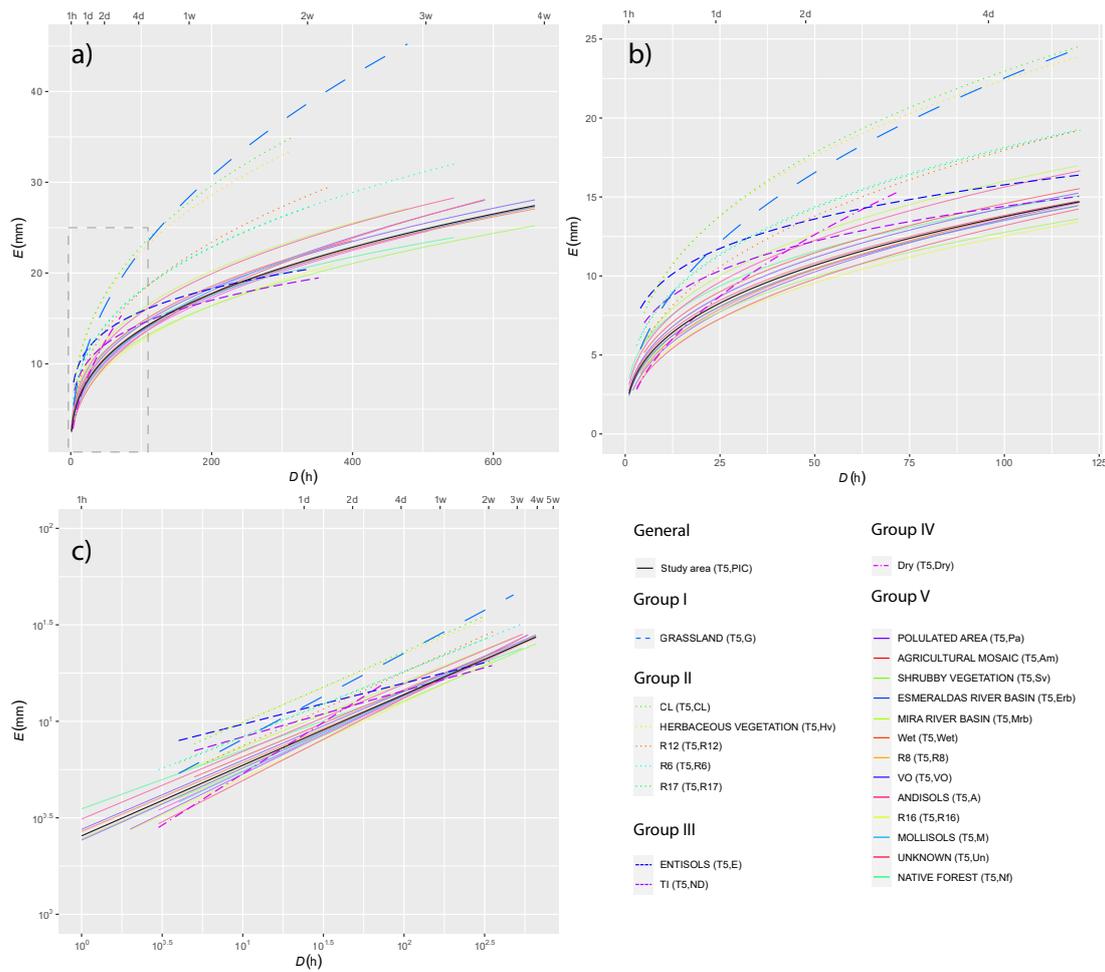
Figure 7.1: Rain threshold curves for 27 representatives of the 45 initial categories and the study area grouped by their trend.

coverage is larger, so the mean value of landslide events is equal to 20 representing 5.7%. Hence, in Figure 7.1a is possible to visualize that the E-D thresholds for precipitation regions are together, where the T5, R12 end with a cumulated rainfall of $\sim 30\ mm$ at earlier rainfall duration ($\sim 370\ h$) than others. Geographically, the R12 is located in the eastern part of the study area, and the annual precipitation varies from 500 - 1400 mm $year^{-1}$. It exceeds the others due to a low amount of events (12 landslides, and 55 MRC) to determine the power law, which has the following values $E = 3.1D^{0.38}$ and uncertainties values of $\Delta\alpha = 3.3\ \Delta\gamma = 0.2$. Then, the precipitation region of R17 has a similar trend to R6 but the range of rainfall duration is until $\sim 350\ h$ and the cumulated rainfall is equal to $\sim 27\ mm$ and the power law is $E = 3.8D^{0.34}$ with the uncertainties

values $\Delta\alpha = 1.8$ $\Delta\gamma = 0.1$. On the other hand, the R6 has the largest range of rainfall duration ($\sim 550\ h$) from this group because this region lies on the Esmeraldas river basin (the western part of the study area), and the cumulated rainfall arrives at $\sim 32\ mm$. The power law $E = 3.9D^{0.33}$ and uncertainties values of $\Delta\alpha = 1.0$ $\Delta\gamma = 0.06$. Finally, the last two E-D thresholds for this group exceed the other faster and have higher values of cumulated rainfall ($\sim 35\ mm$) in a shorter time ($\sim 300\ h$). The power law equations are equal to $E = 4.5D^{0.35}$ with uncertainties values equal to $\Delta\alpha = 1.8$ $\Delta\gamma = 0.09$ for the herbaceous vegetation; for conglomerates the equation is equal to $E = 4.2D^{0.37}$ and $\Delta\alpha = 5.8$ $\Delta\gamma = 0.2$. Moreover, geographically the location of herbaceous vegetation matches with R17, and conglomerates are in the area of R8 and R17; but the power law equations differ drastically from R8. In the case of R17, the power law can be considered similar if the uncertainties values are added.

Group III is compounded by two E-D thresholds from different categories, which are T5, E from types of soils, and T5, TI from lithology. These start with high rainfall conditions (MRC), increasing and crossing group IV as shown clearly in Figure 7.1c. The power law values for T5, E is $E = 5.9D^{0.21}$, the range of rainfall duration is 1<D<350, and the uncertainties values are $\Delta\alpha = 2.7$ $\Delta\gamma = 0.1$. On the other hand, the power law values for T5, TI is $E = 4.8D^{0.24}$, the range of rainfall duration is equal to T5, E (1<D<350), and the uncertainties values are $\Delta\alpha = 3.0$ $\Delta\gamma = 0.2$. Moreover, the two E-D thresholds match geographically, however, the entisols cover more area than the undifferentiated terraces. This group has higher uncertainty values due to the smaller amount of rainfall conditions responsible for the slope failure, also lower landslide events in the area. As the thresholds differ in the cover area the number of landslide events for this group is 33 representing 11%.

Group IV has one category of E-D threshold about the climatic season, in which the T5, Dry is recognized with a fast increase and short rainfall duration (1<D<70) than other groups. It represents 3% of the landslide events because the incidence of landslides during the dry season is 10. The power law for this threshold is equal to $E = 1.6D^{0.53}$, and the

uncertainties values are $\Delta\alpha = 0.6 \ \Delta\gamma = 0.2$. Moreover, in this case, the analysis is done with the information of precipitation regions of the study area with shorter regimes of annual precipitation suggested by Ilbay-Yupa et al. (2021). In this case, the R16 (500-1000 mm $year^{-1}$) is the precipitation region with the lowest values in the annual precipitation; being part of a different group (IV), Where the E-D threshold has the range of rainfall duration 1<D<350, which is a clearly different indicator to T5, Dry. Nevertheless, the T5, R16 is similar to the power law $E = 2.1D^{0.39} \ \Delta\alpha = 0.5 \ \Delta\gamma = 0.06$. These power law equations can be considered similar if the uncertainties values are added.

Group V has the largest group of E-D thresholds (13 thresholds), and the difference between them is not high, it can be considered a similar trend, where 30.7% of group V are categories from the land cover and 23% correspond to soil type. There are two categories (precipitation regions and hydrographic units) with the same percentage of 15.4% (for each one there are two E-D thresholds). Finally, there are two other categories (climatic season and lithology) with the same percentage of 7,7 % (for each one there is one E-D threshold). Moreover, from a general point of view, the different classifications of E-D thresholds match geographically and cover mostly the entire study area from north to south. The mean of landslide events analyzed in this group is ~ 251 representing 71%, for this reason, the E-D thresholds are related between them and the trend are similar.

Therefore, the categories of E-D thresholds with the same percentages are analyzed. The classifications with only one threshold (7.7%); for the climatic season analysis, the relevant threshold is for the wet season (T5, Wet) with the following power law $E = 2.7D^{0.36}$ with a range of rainfall duration 1<D<650, and uncertainties values $\Delta\alpha = 0.3 \ \Delta\gamma = 0.03$. Then, the E-D threshold for lithology (T5, VO - volcanic rocks) has the following power law value $E = 2.8D^{0.36}$, the range of rainfall duration is identical to T5, VO, and uncertainties values are $\Delta\alpha = 0.4 \ \Delta\gamma = 0.03$.

To understand the relationship between the rest of E-D thresholds of this group, the coverage area is taken into account. The precipitation region of R16 is located in the Esmeraldas river basin (T5, Erb) for this reason the power law equations are pretty

similar but differ with rainfall duration due to Erb having a larger area with more MRC (952 rainfall conditions). On the other hand, the R8 precipitation region matches the two hydrographic units and the power law equations are similar between Erb ($E = 2.4D^{0.37}$, $\Delta\alpha = 0.3\ \Delta\gamma = 0.03$) and R8 ($E = 2.5D^{0.37}$, $\Delta\alpha = 0.4\ \Delta\gamma = 0.03$) due to the number of rainfall conditions, the power law for the Mira river basin $E = 3.4D^{0.34}$, $\Delta\alpha = 0.9$ $\Delta\gamma = 0.07$ can be considered similar if the uncertainties values are subtracted.

Furthermore, the categories of land cover and soil types have strong similarities due to the location of each one. Firstly, the big area is covered by native forest (T5, Nf) and the andisols (T5, A). The range of rainfall duration is equal to 1<D<550 and the power law equations match $E = 3.5D^{0.3}$, $\Delta\alpha = 1.4\ \Delta\gamma = 0.08$ (T5, Nf) and $E = 3.1D^{0.35}$, $\Delta\alpha = 0.5\ \Delta\gamma = 0.04$ (T5, A). It differs in the uncertainties values due to the number of rainfall conditions for each one, the native forest has less (184 MRC), and the Andisols has 398 MRC due to more coverage in the study area.

Then, the agricultural mosaic (T5, Am), and shrubby vegetation (T5, Sv) match with the mollisols area (T5, M), in which the T5, Sv needs more rainfall duration ($\sim 650\ h$) to obtain the same amount of cumulated rainfall of T5, Am ($\sim 30\ mm$). The mollisols have the same range of rainfall duration to T5, Am (until $\sim 400\ h$). The power law equations are the following $E = 2.9D^{0.35}$, $\Delta\alpha = 1.1\ \Delta\gamma = 0.09$ (T5, Am); $E = 2.4D^{0.36}$, $\Delta\alpha = 0.6\ \Delta\gamma = 0.05$ (T5, Sv); $E = 2.2D^{0.4}$, $\Delta\alpha = 0.6\ \Delta\gamma = 0.06$ (T5, M), in which the agricultural mosaic has bigger uncertainties values due to the lower value of MRC (107 rainfall conditions).

Finally, the populated area (Pa from land cover) with the unknown: urban peri-urban areas (Un from soil type) matches. The range of rainfall duration for both is equal to 1<D<600, and the power law equations are $E = 2.1D^{0.41}$, $\Delta\alpha = 0.4\ \Delta\gamma = 0.04$ (T5, Pa) and $E = 1.8D^{0.43}$, $\Delta\alpha = 0.5\ \Delta\gamma = 0.05$ (T5, Un). This group is considered to have better values for the E-D thresholds because do not vary strongly between them. Therefore, the rainfall conditions for triggering a landslide can be easy to estimate in this group for each E-D threshold from different categories, and the tendency match with the

general threshold (T5, PIC).

7.4 Comparison with existing global and site-specific thresholds

PIC thresholds could be compared with other E-D thresholds proposed in the literature for the failure initiation of landslides. Figure 7.2 shows the threshold curve for the study area (PIC) and compares it with five published global ED thresholds for shallow landslides detailed in Table 7.1.

Table 7.1: Global and site-specific rainfall threshold equations. Source: 1,Caine (1980); 2-4,Guzzetti et al. (2008); 5,Peruccacci et al. (2017)

#	Location	Equation	Range
1	Global	$E = 14.82\,D^{0.61}$	$0.167h \leq D \leq 500h$
2	Global	$E = 2.20\,D^{0.56}$	$0.1h \leq D \leq 1000h$
3	Global	$E = 2.28\,D^{0.8}$	$0.1h \leq D \leq 48h$
4	Global	$E = 0.48\,D^{0.89}$	$48h \leq D \leq 1000h$
5	Italy	$E = 7.70\,D^{0.39}$	$1h \leq D \leq 1200h$
6	PIC	$E = 2.50\,D^{0.37}$	$1h \leq D \leq 650h$

The global PIC threshold inferred from this study data is lower than the other E-D threshold curves, for most rainfall durations.

Comparing the ED threshold curve PIC in Figure 7.2 with the global E-D curve obtained by Caine (1980) (Global #1 in Figure 7.2) shows a large difference that can be attributed to the number of landslides in each data set (79 for Global #1 vs 354 for PIC), another difference is that the landslides used by Caine are catastrophic events, thus the curve may not represent a minimum boundary (Guzzetti et al., 2008), while the data set used to infer the PIC does not differentiate the events by their magnitude.

Comparing the ED threshold curve PIC in Figure 7.2 with Global #2 in Figure 7.2 obtained by Guzzetti et al. (2008) shows a large difference for any given rainfall duration, for any given cumulated rainfall. The large difference is the result of a number of events in a set of rainfall and landslides (354 vs 2626). Comparing the ED threshold curve PIC

in Figure 7.2 with composite threshold #3 and #4 in Figure 7.2 obtained by Guzzetti et al. (2008) shows a large difference for large rainfall duration, and for much shorter rainfall periods, for large cumulated rainfall.

The ED threshold curve PIC in Figure 7.2 has a slope similar to the slope of the threshold curves established by Peruccacci et al. (2017) for shallow landslides in Italy that use a similar approach of this work (Italy curve in Figure 7.2); $\gamma = 0.39\ vs\ 0.37$, but predicts the likely occurrence of shallow landslide ĩwo times lower at average cumulated rainfall, for any given rainfall duration. The difference in these results is due to a large difference in the number of landslides in each data set (354 vs 2819) and also the period of time used (5 years vs 19 years).
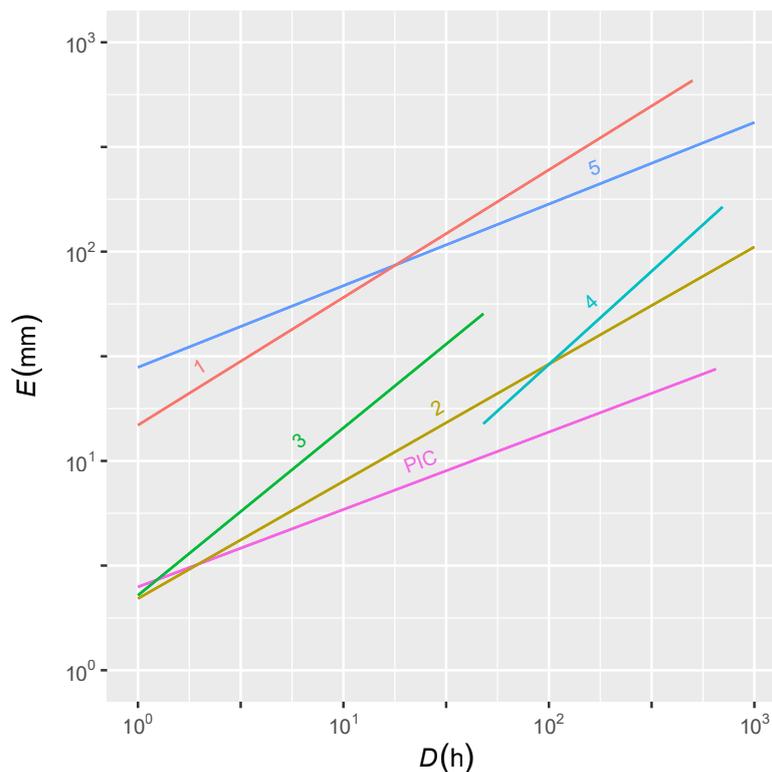
Figure 7.2: Comparison between PIC and Global, site-specific, rainfall threshold curves in log-log coordinates. Source: 1,Caine (1980); 2, threshold inferred from the entire set of ID rainfall Guzzetti et al. (2008); 3-4, thresholds inferred from estimates of the rainfall conditions, for two different rainfall periods ($D \leq 48h$, and $D \geq 48h$) Guzzetti et al. (2008); 4,Peruccacci et al. (2017); 5,Peruccacci et al. (2017).

8 | Conclusions

In the northern Andes of Ecuador (Pichincha, Imbabura, and Carchi provinces) during the period from 2013 to 2018 (5 years) occurred 354 shallow landslide events according to DesInventar (2021). The area with more incidence of landslides occurs in altitudes lower than 2500 m (56% of landslide events) and 72.8 % occurred on roads. The analysis of E-D thresholds at different categories for the study area was useful to establish actual threshold for each specific area. Moreover, the relation with other threshold for different categories allows to understand the combination effects of different features of the zone.

Therefore, 27 of the 45 category divisions were analyzed allowing the identification of different groups due to the trend of power law curves. We found that group V has an E-D threshold for all select categories, specifically: climatic season (wet), precipitation regions (R8, and R16), land cover (Pa, Am, Sv, and Nf), soil type (A, Un, and M), lithology (VO), and hydrographic units (Erb, and Mrb) to analyze the features of the study area. This is because the coverage area of this group contains 71% of landslide events. The trend of the corresponding E-D thresholds does not vary strongly between them and adjusts with the power law curve for the entire study area, generating the occurrence of shallow landslides due to small rainfall conditions in the area of their respective categories.

Then, group IV has one category (climatic season), the E-D threshold shows the shortest power law curve because it has a range of rainfall duration that reaches 75h due to a low number of shallow landslide occurrences (10 events), and 19 multiple rainfall conditions (MRC) during the dry season. Group III contains information on two cate-

gories (soil type and lithology) which have intermediate rainfall conditions to trigger a shallow landslide and the E-D thresholds have a short range of rainfall duration (350 h). Group II has three categories: precipitation regions (R12, R17, and R6), land cover (Hv), and lithology (CL) which need higher rainfall conditions to trigger a landslide. Group I has one E-D threshold from the land cover category which needs higher rainfall conditions to trigger a landslide. The last two groups (I, and II) have higher rainfall conditions due to a small number of landslide events and MRC.

The PIC thresholds curves have scaling exponents similar to the Italy curve that use the same approach. The PIC shows a much lower cumulated rainfall than the other published rainfall ED thresholds for the initiation of shallow landslides, for any given rainfall duration.

Overall we have shown that despite landslide database limitations and the lack of high-resolution ground-rainfall measurements it was possible to perform a first order approximation to rainfall thresholds for shallow landslides in the Andean landscape of northern Ecuador. This should provide reference values for further detailed studies in the region.

9 | Recomendations

- Based on the DesInvertar information most of the landslides (258 events representing the 72.8%) occurred on roads. The 14% (49 events) matched occurrences on other places, and the remaining 13.2% (47 events) did not have information. To improve the trustability of the threshold curve obtained by this work it is necessary to use a large data set of landslides. Therefore it is recommended to carry out the analysis for a larger area or to increase the data through remote landslide detection.

- To improve the accuracy of the thresholds estimation a better identification of some parameters (e.g. k-factor) is necessary. For this purpose is important to conduct studies about soil hydrology through samples taken on fieldwork such information is very useful, among other things, to model rainfall accumulation on the ground.

References

Aspden, J. A., & Litherland, M. (1992). The geology and mesozoic collisional history of the cordillera real, ecuador. *Tectonophysics*, *205*(1-3), 187–204. doi: 10.1016/0040-1951(92)90426-7

Badr, H. S., Zaitchik, B. F., & Dezfuli, A. K. (2015). A tool for hierarchical climate regionalization. *Earth Science Informatics*, *8*(4), 949–958. doi: 10.1007/s12145-015-0221-7

Ballari, D., Giraldo, R., Campozano, L., & Samaniego, E. (2018). Spatial functional data analysis for regionalizing precipitation seasonality and intensity in a sparsely monitored region: Unveiling the spatio-temporal dependencies of precipitation in ecuador. *International Journal of Climatology*, *38*(8), 3337–3354. doi: 10.1002/joc.5504

Berti, M., Martina, M., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a bayesian approach. *Journal of Geophysical Research: Earth Surface*, *117*(F4). doi: 10.1029/2012jf002367

Brenning, A., Schwinn, M., Ruiz-Páez, A., & Muenchow, J. (2015). Landslide susceptibility near highways is increased by 1 order of magnitude in the andes of southern ecuador, loja province. *Natural Hazards and Earth System Sciences*, *15*(1), 45–57. doi: 10.5194/nhess-15-45-2015

Brunetti, M., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., & Guzzetti, F. (2010). Rainfall thresholds for the possible occurrence of landslides in italy. *Natural Haz-*

ards and Earth System Sciences, *10*(3), 447–458.

Burbano, N., Becerra, S., & Pasquel, E. (2015). Introducción a la hidrogeología del ecuador. *Quito, Ecuador: INAHMI*.

Caine, N. (1980). The rainfall intensity-duration control of shallow landslides and debris flows. *Geografiska annaler: series A, physical geography*, *62*(1-2), 23–27.

Cedeño, J., & Donoso, M. C. (2010). *Atlas pluviométrico del ecuador.* PHI-LAC. Retrieved from `https://unesdoc.unesco.org/ark:/48223/pf0000216357.locale=en`

CISPDR. (2016a). Plan hidráulico regional de la demarcaión hidrográfica esmeraldas.. Retrieved from `http://suia.ambiente.gob.ec/files/MEMORIA%20DH%20ESMERALDAS.pdf`

CISPDR. (2016b). Plan hidráulico regional de la demarcaión hidrográfica mira.. Retrieved from `http://suia.ambiente.gob.ec/files/MEMORIA%20DH%20MIRA.pdf`

Coltorti, M., & Ollier, C. (2000). Geomorphic and tectonic evolution of the ecuadorian andes. *Geomorphology*, *32*(1-2), 1–19. doi: 10.1016/s0169-555x(99)00036-7

Coronel, L. (2020). *Fonag. 2020. plan estratÉgico 2021 -2025.* Retrieved from `https://www.fonag.org.ec/web/wp-content/uploads/2022/01/Plan-Estrategico-Fonag-2021-2025-WEB.pdf`

Crozier, M. J. (1999). Prediction of rainfall-triggered landslides: A test of the antecedent water status model. *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, *24*(9), 825–833. doi: 10.1002/(sici)1096-9837(199908)24:9<825::aid-esp14>3.0.co;2-m

DesInventar. (2021). *Desinventar data base of landslide.* Retrieved from `https://db.desinventar.org/`

Fairbridge, R. W. (2006). *Encyclopedia of world regional geology.*

Falsini, F. (1995). *Rilevamento geologico e geomorfologico dell'area compresa tra tumbaco e guayallabamba (depressione interandina, ecuador centrale)* (Unpublished

doctoral dissertation). Tesi di Laurea, Universitfi di Siena.

Ficcarelli, G., Azzaroli, A., Bertini, A., Coltorti, M., Mazza, P., Mezzabotta, C., . . . Torre, D. (1997). Hypothesis on the cause of extinction of the south american mastodonts. *Journal of South American Earth Sciences*, *10*(1), 29–38. doi: 10.1016/s0895-9811(97)00003-5

Glade, T., Crozier, M. J., & Smith, P. J. (2000). Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical "antecedent daily rainfall model". *pure and applied geophysics*, *157*, 1059-1079. doi: 10.1007/s000240050017

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

Highland, L. (2004). *Landslide types and processes, a fact sheet.* Retrieved from `http://pubs.usgs.gov/fs/2004/3072/`

Highland, L. M., & Bobrowsky, P. (2008). *The landslide handbook - a guide to understanding landslides.* US Geological Survey. doi: 10.3133/cir1325

Huffman, G., Stocker, E., Bolvin, D., Nelkin, E., & Tan, J. (2017). Algorithm theoretical basis document (atbd)version 06. In *Algorithm theoretical basis document (atbd) version 06.* Retrieved from `https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHHE_06/summary`

Hughes, R. A., & Pilatasig, L. F. (2002). Cretaceous and tertiary terrane accretion in the cordillera occidental of the andes of ecuador. *Tectonophysics*, *345*(1-4), 29–48. doi: 10.1016/s0040-1951(01)00205-0

Hungerbühler, D., Steinmann, M., Winkler, W., Seward, D., Egüez, A., Peterson, D. E., . . . Hammer, C. (2002). Neogene stratigraphy and andean geodynamics of southern ecuador. *Earth-Science Reviews*, *57*(1-2), 75–124. doi: 10.1016/s0012-8252(01)00071-x

Ibáñez Asensio, S., Gisbert Blanquer, J. M., & Moreno Ramón, H. (2011). *Entisoles.* Uni-

versitat Politècnica de València. Retrieved from `http://hdl.handle.net/10251/12883`

Ibañez Asensio, S., Gisbert Blanquer, J. M., & Moreno Ramón, H. (2011). *Mollisoles.* Universitat Politècnica de València. Retrieved from `http://hdl.handle.net/10251/13609`

Ilbay-Yupa, M., Lavado-Casimiro, W., Rau, P., Zubieta, R., & Castillón, F. (2021). Updating regionalization of precipitation in ecuador. *Theoretical and Applied Climatology*, *143*(3), 1513–1528. doi: 10.1007/s00704-020-03476-x

Jaillard, E., Caron, M., Dhondt, A., Ordoñez, M., Andrade, R., Bengtson, P., . . . Díaz, R. (1997). Síntesis estratigráfica y sedimentológica del cretáceo y paleógeno de la cuenca oriental del ecuador. *Orstom-Petroproduccion eds*, 164. Retrieved from `https://www.researchgate.net/publication/46298379_Sintesis_estratigrafica_y_sedimentologica_del_Cretacico_y_Paleogeno_de_la_Cuenca_Oriental_del_Ecuador`

Jaillard, E., Ordonez, M., Suárez, J., Toro, J., Iza, D., & Lugo, W. (2004). Stratigraphy of the late cretaceous–paleogene deposits of the cordillera occidental of central ecuador: geodynamic implications. *Journal of South American Earth Sciences*, *17*(1), 49–58. doi: 10.1016/j.jsames.2004.05.003

Janeau, J.-L., Grellier, S., & Podwojewski, P. (2015). Influence of rainfall interception by endemic plants versus short cycle crops on water infiltration in high altitude ecosystems of ecuador. *Hydrology Research*, *46*(6), 1008–1018. doi: 10.2166/nh.2015.203

Jorgensen, P. M., & Leon-Yanez, S. (1999). *Catálogo de las plantas vasculares del ecuador* (Vol. 75). Missouri Botanical Garden St Louis, USA.

Leonarduzzi, E., Molnar, P., & McArdell, B. W. (2017). Predictive performance of rainfall thresholds for shallow landslides in switzerland from gridded daily data. *Water Resources Research*, *53*(8), 6612–6625. doi: 10.1002/2017wr021044

Li, Z., Tang, G., Hong, Z., Chen, M., Gao, S., Kirstetter, P., ... others (2021). Two-decades of gpm imerg early and final run products intercomparison: Similarity and difference in climatology, rates, and extremes. *Journal of Hydrology*, *594*, 125975.

Litherland, M., Aspden, J., & Jemielita, R. (1994). The metamorphic belts of ecuador, british geological survey, overseas memoir 11. *British Geological Survey, Keyworth*, *147*. Retrieved from `https://archive.org/details/1994 -litherland-los-cinturones-metamorficos-del-ecuador-sp/ page/n7/mode/2up`

MAG. (n.d.). *Catálogo de objetos geográficos y productos cartográficos agropecuarios del ministerio de agricultura y ganadería.* Retrieved from `http://geoportal .agricultura.gob.ec/pdf/catalogo_mag_volumenV.pdf`

Markgraf, V., & Hastenrath, S. (1982). *The glaciation of the ecuadorian andes* (Vol. 14) (No. 4). Informa UK Limited. doi: 10.2307/1550805

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., & Guzzetti, F. (2015). Rainfall thresholds for the possible landslide occurrence in sicily (southern italy) based on the automatic reconstruction of rainfall events. *Landslides*, *13*(1), 165–172. doi: 10.1007/s10346-015-0630-1

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., Roccati, A., & Guzzetti, F. (2017). Ctrl-t algorithm, calculation of thresholds for rainfall-induced landslides -tool. *Geomorphology Research Group @ CNR IRPI Perugia*. Retrieved from `http://geomorphology.irpi.cnr.it/tools/rainfall-events -and-landslides-thresholds/ctrl-t-algorithm`

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., Roccati, A., & Guzzetti, F. (2018). A tool for the automatic calculation of rainfall thresholds for landslide occurrence. *Environmental Modelling & Software*, *105*, 230-243. doi: 10.1016/j.envsoft.2018.03.024

Morán-Tejeda, E., Bazo, J., López-Moreno, J. I., Aguilar, E., Azorín-Molina, C., Sanchez-Lorenzo, A., ... Martín-Hernández, N. (2016). Climate trends and vari-

ability in ecuador (1966–2011). *International Journal of Climatology*, *36*(11), 3839–3855. doi: 10.1002/joc.4597

Moreno, H., Ibáñez, S., & Gisbert, J. (2011). *Andisoles.* Retrieved from `https://riunet.upv.es/bitstream/handle/10251/13676/Andisoles.pdf`

Nerini, D., Zulkafli, Z., Wang, L.-P., Onof, C., Buytaert, W., Lavado-Casimiro, W., & Guyot, J.-L. (2015). A comparative analysis of trmm–rain gauge data merging techniques at the daily time scale for distributed rainfall–runoff modeling applications. *Journal of Hydrometeorology*, *16*(5), 2153–2168. doi: 10.1175/jhm-d-14-0197.1

Noble, S., Aspden, J., & Jemielita, R. (1997). Northern andean crustal evolution: New u-pb geochronological constraints from ecuador. *Geological Society of America Bulletin*, *109*(7), 789–798. doi: 10.1130/0016-7606(1997)109<0789:nacenu>2.3.co;2

Peruccacci, S., Brunetti, M. T., Gariano, S. L., Melillo, M., Rossi, M., & Guzzetti, F. (2017). Rainfall thresholds for possible landslide occurrence in italy. *Geomorphology*, *290*, 39–57. doi: 10.1016/j.geomorph.2017.03.031

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

Ramadhan, R., Yusnaini, H., Marzuki, M., Muharsyah, R., Suryanto, W., Sholihun, S., . . . others (2022). Evaluation of gpm imerg performance using gauge data over indonesian maritime continent at different time scales. *Remote Sensing*, *14*(5), 1172.

Ray, R. L., & Jacobs, J. M. (2007). Relationships among remotely sensed soil moisture, precipitation and landslide events. *Natural Hazards*, *43*(2), 211–222. doi: 10.1007/s11069-006-9095-9

Rivera, A. (2021). *Analysis of landslide triggering mechanisms in the imbabura province* (B.S. thesis, Universidad de Investigación de Tecnología Experimental Yachay). Retrieved from `http://repositorio.yachaytech.edu.ec/handle/123456789/387`

Samaniego, P., Robin, C., Chazot, G., Bourdon, E., & Cotten, J. (2010). Evolving metasomatic agent in the northern andean subduction zone, deduced from magma composition of the long-lived pichincha volcanic complex (ecuador). *Contributions to Mineralogy and Petrology*, *160*(2), 239–260. doi: 10.1007/s00410-009-0475-5

Sauer, W. (1965). *Geología del ecuador*. Talleres Gráficos del Ministerio de Educación.

Savtchenko, A. (2017). Readme document for the gpm data. In *Nasa goddard space flight center*. Retrieved from `https://disc.gsfc.nasa.gov/datasets/ GPM_3IMERGHHE_06/summary`

SIGTIERRAS. (2017). *Memoria explicativa del mapa de Órdenes de suelos del ecuador.* Retrieved from `http://metadatos.sigtierras.gob.ec/pdf/ MEMORIA_MAPA_DE_ORDENES_DE_SUELOS_MAG_SIGTIERRAS.pdf`

Soto, J., Galve, J. P., Palenzuela, J. A., Azañón, J. M., Tamay, J., & Irigaray, C. (2017). A multi-method approach for the characterization of landslides in an intramontane basin in the andes (loja, ecuador). *Landslides*, *14*(6), 1929–1947. doi: 10.1007/s10346-017-0830-y

Spikings, R. A., Seward, D., Winkler, W., & Ruiz, G. M. (2000). Low-temperature thermochronology of the northern cordillera real, ecuador: Tectonic insights from zircon and apatite fission track analysis. *Tectonics*, *19*(4), 649–668. doi: 10.1029/2000tc900010

Tamay, J., Galindo-Zaldívar, J., Martos, Y., & Soto, J. (2018). Gravity and magnetic anomalies of ecuadorian margin: Implications in the deep structure of the subduction of nazca plate and andes cordillera. *Journal of South American Earth Sciences*, *85*, 68–80. doi: 10.1016/j.jsames.2018.04.020

Tobar, V., & Wyseure, G. (2018). Seasonal rainfall patterns classification, relationship to enso and rainfall trends in ecuador. *International Journal of Climatology*, *38*(4), 1808–1819. doi: 10.1002/joc.5297

Urgilez, A., Robles, J., Bakker, M., Guzman, P., & Bogaard, T. (2020). Characterization and hydrological analysis of the guarumales deep-seated landslide in the tropical

ecuadorian andes. *Geosciences*, *10*(7), 267. doi: 10.3390/geosciences10070267

Varela, A., & Ron, S. (2018). Geografía y clima del ecuador. *BIOWEB. Pontificia Universidad Católica del Ecuador*. Retrieved from `https://bioweb.bio/geografiaClima.html/`

Villagómez, D., Egüez, A., Winkler, W., & Spikings, R. (2002). Pifo-quaternary sedimentary and tectonic evolution of the central inter-andean valley in ecuador. *City*, *10*, 5.

Vuille, M., Bradley, R. S., & Keimig, F. (2000). Climate variability in the andes of ecuador and its relation to tropical pacific and atlantic sea surface temperature anomalies. *Journal of climate*, *13*(14), 2520–2535. doi: 10.1175/1520-0442(2000)013<2520:cvitao>2.0.co;2

Wieczorek, G. F. (1996). Landslides: investigation and mitigation. chapter 4-landslide triggering mechanisms. *Transportation Research Board Special Report*(247).

Winkler, W., Villagómez, D., Spikings, R., Abegglen, P., & Egüez, A. (2005). The chota basin and its significance for the inception and tectonic setting of the inter-andean depression in ecuador. *Journal of South American Earth Sciences*, *19*(1), 5–19. doi: 10.1016/j.jsames.2004.06.006

Zaruba, Q., & Mencl, V. (2014). *Landslides and their control*. Elsevier.

Appendix

Appendix A │ Guide to download landslide database

1. Enter by the browser to the address

   `https://www.desinventar.net/DesInventar/index.jsp`.

2. Click on the analysis tab and select the Ecuador region.

3. Select the search by Query tab.

4. Select "landslide" the type of disaster.

5. Click on View data.
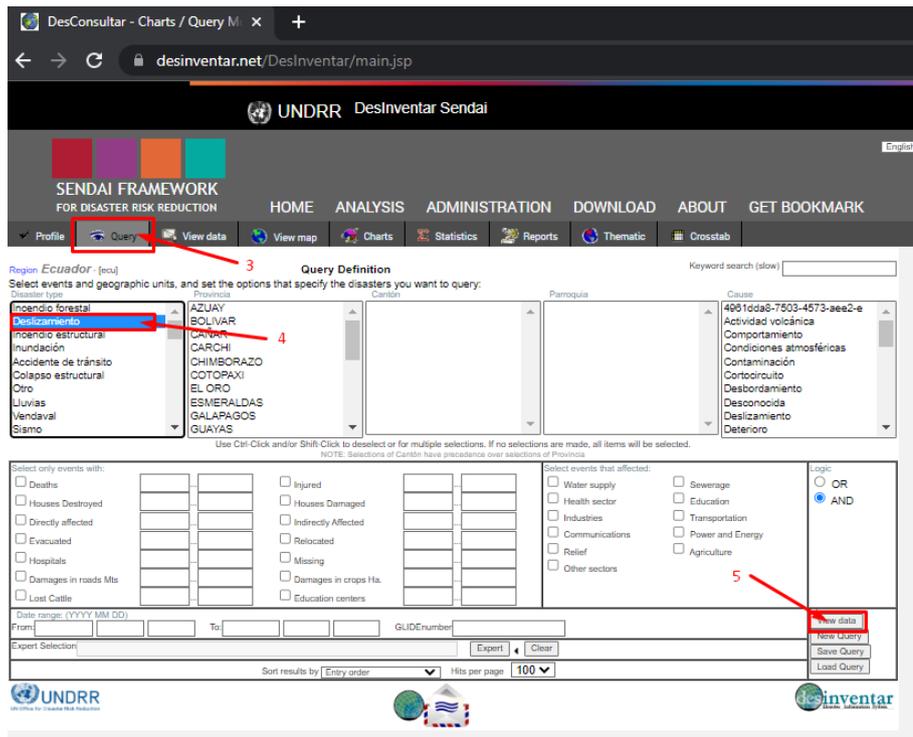
6. Go to the reports tab.

7. Select the available information you want and click continue.
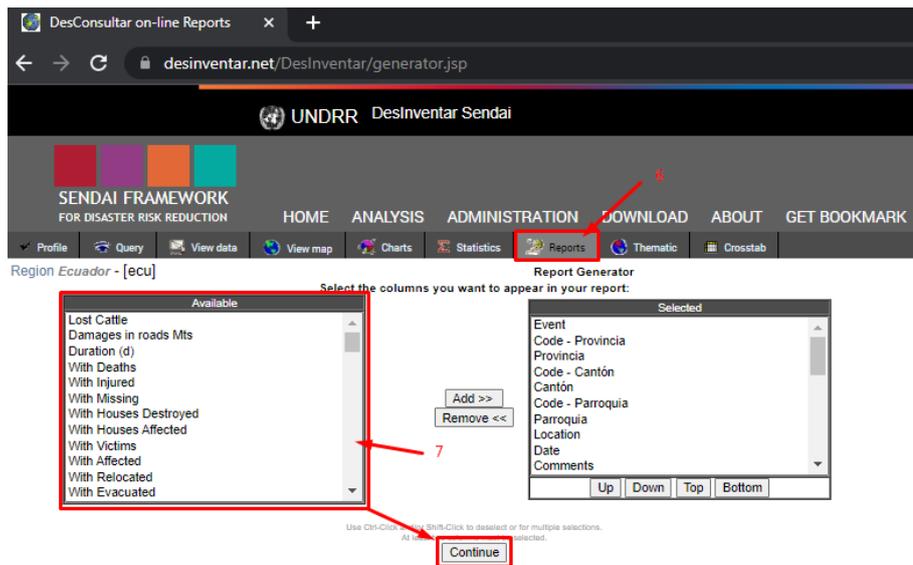
8. At the top of the table click on get it as Excel CSV.

Appendix B ｜ Guide to download GPM data

1. Enter by the browser to the address.

   `https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHH_06/summary.`

2. Click on Subset/Get Data.
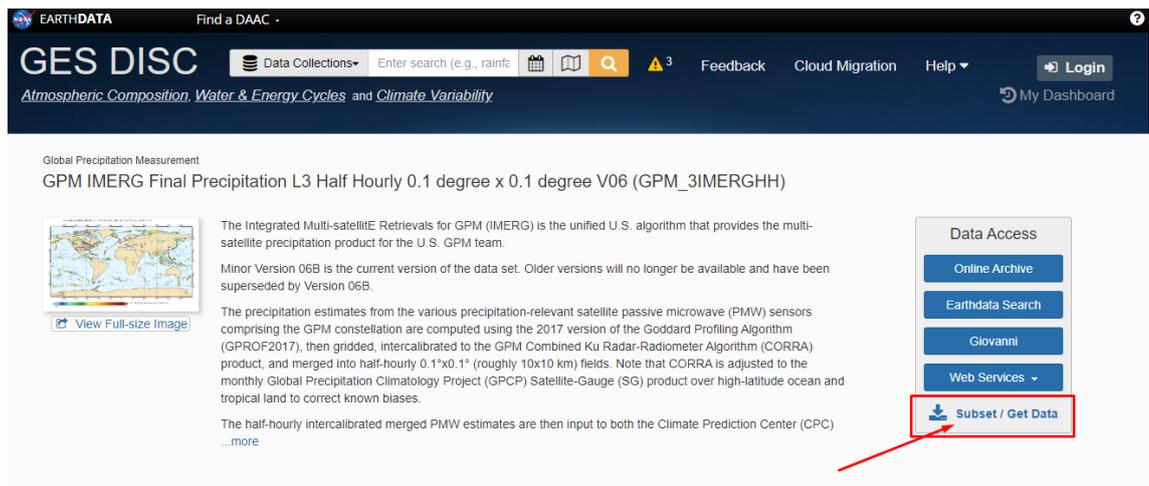
3. Select the download method.

4. Set the method options: Refine the range date.

5. Refine the region with the tool draw a rectangle.

Method Options ⑦

| ▸ Refine Date Range: | ✔ 2013-01-01 to 2018-12-31 | Reset |

| ▾ Refine Region: | ✔ -79.76, -0.66, -76.24, 1.36 | Reset |

-79.76,-0.66,-76.24,1.36

Available Range: -180, -90, 180, 90   Cursor Coordinates: 0.518, -84.155

6. Choose the variable to download.

Method Options ⑦

| ▸ Refine Date Range: | ✔ 2013-01-01 to 2018-12-31 | Reset |

| ▸ Refine Region: | ✔ -79.76, -0.66, -76.24, 1.36 | Reset |

☑ ✂ Use 'Refine Region' for geo-spatial subsetting ⑦

| ▾ Variables: | ✔ 1 variable(s) selected | Reset |

☐ HQobservationTime = Microwave satellite observation time
☐ HQprecipitation = High Quality precipitation from all available passive microwave sources
☐ HQprecipSource = Microwave satellite source identifier
☐ IRkalmanFilterWeight = IR-data weights in MW Kalman smoothing
☐ IRprecipitation = Precipitation (IR-only) from all available geo-satellites, passive-microwave-calibrated
☑ precipitationCal = Precipitation (combined microwave-IR) estimate with post-processing gauge calibration
☐ precipitationQualityIndex = Quality Index of precipitationCal
☐ precipitationUncal = Precipitation (combined microwave-IR) estimate before the post-processing gauge calibration
☐ probabilityLiquidPrecipitation = Probability of liquid precipitation phase
☐ randomError = Random error for precipitationCal

7. Set the interpolation, this value is not mandatory.

8. Download the list of files and follow the intrusions to the automatic download of each file.

Appendix C | Guide to download

FONAG data

1. Enter by the browser to the address.

   `https://sedc.fonag.org.ec/reportes/consultas/.`

2. Select automatic transmission.

3. Select one station for downloading the data.

   For each station the following steps are the same:

4. Select the "precipitation" variable.

5. Choose the start date (01-01-2013) and the end date (31-12-2018).

6. Select the hourly frequency.
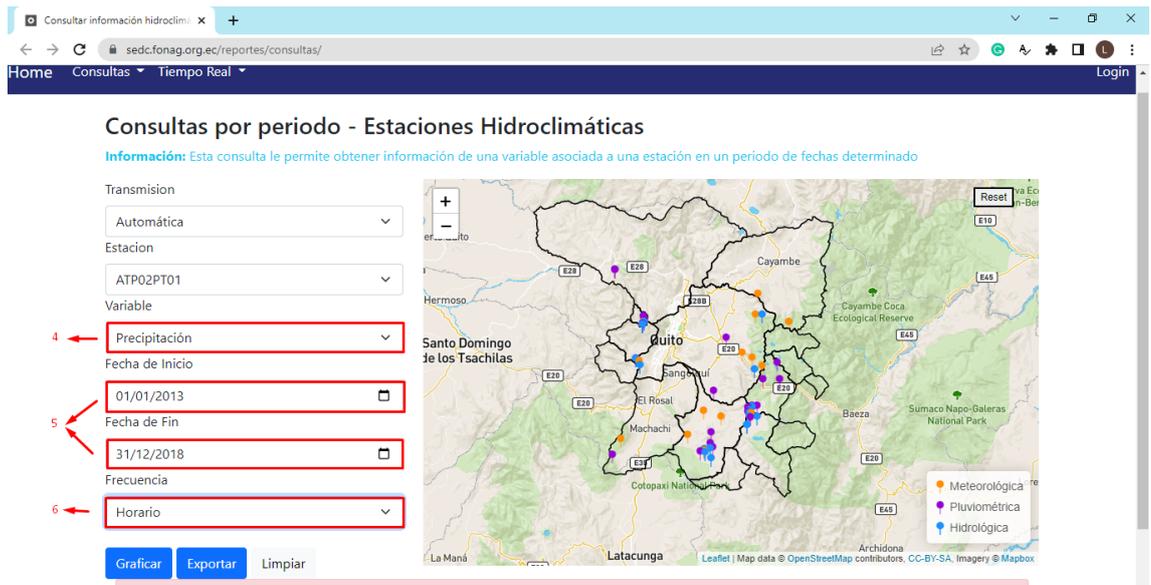
7. Finally, click on export.

Appendix D | R scrips

To see R scrips go to Drive folder \ scrips

Table D.1: R scrips

Id	Name	Description
C001	Ctrl-T.R	This code reconstruct rainfall events, select the representative rain gauges, get multiple failure rainfall conditions in terms of D and E, models the antecedent rainfall, give a probability to each rainfall condition, and calculates probabilistic rainfall thresholds and their associated uncertainties.
C002	get_data.R	This code extract and join the variable "precipitationCal" to form the de rainfall series for each pixel in the ndf4 files downloaded from GPM

Appendix E │ Output Phase 5

For further output files see Drive folder \ outputs
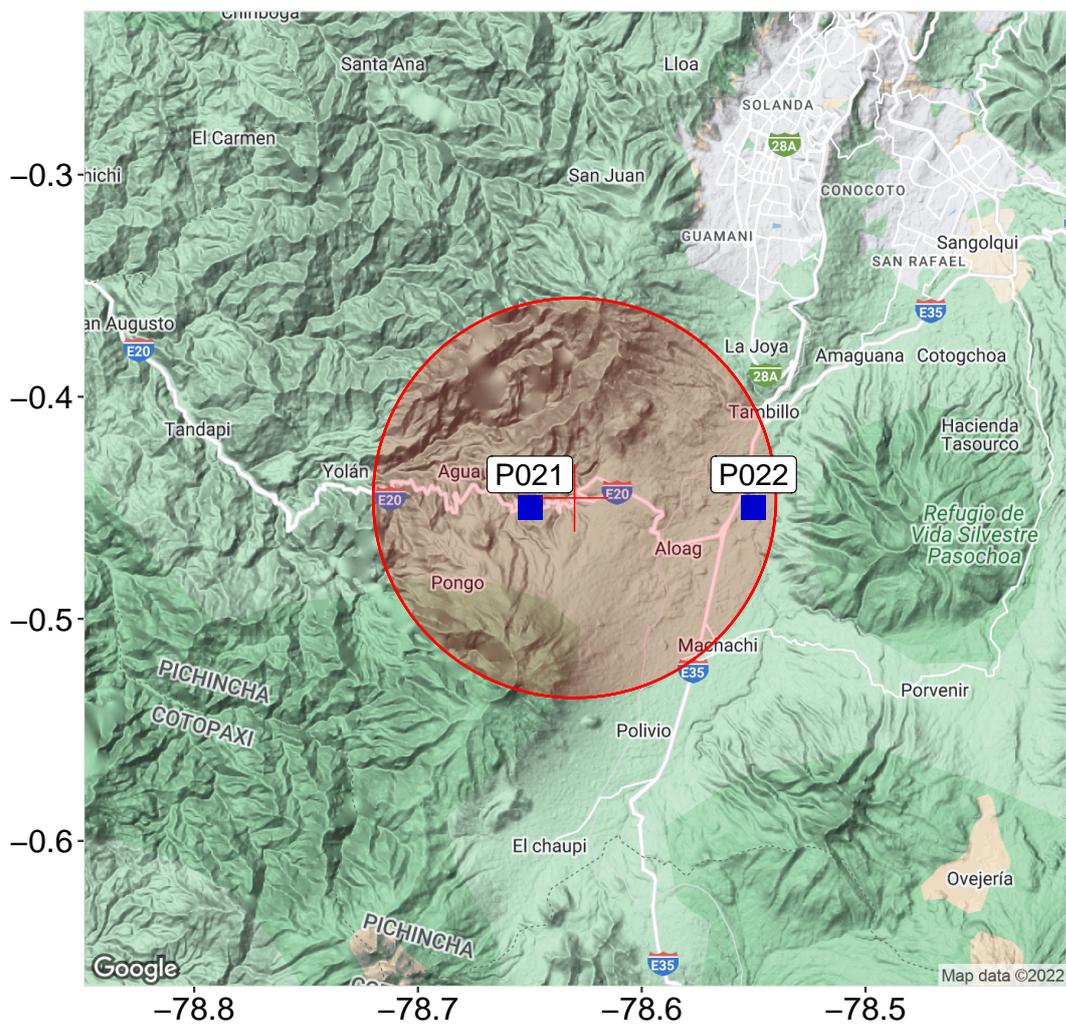
RAIN GAUGE : `P021`
Distance from the landslide: 2.28 km
Temporal coverage
Start Date: 2013–01–01 07:00+0000
Stop Date: 2018–12–31 23:00+0000
Data Resolution
Temporal Resolution: hourly
Rainfall events
#: 248
Statistics:

	Min	Max	Mean	Median	SD
D_E (h)	1	1052	96.67	31.5	158.65
E_E (mm)	1.02	358.96	32.57	10.41	52.8

ID LANDSLIDE : `L009_a`   **DELAY:0 h**

Landslide date: 14–01–08 12:00
Rainfall event associated with the landslide
#: 42
Start Date: 2014–01–05 22:00+0000
Stop Date: 2014–01–15 12:00+0000
D_E **:** 231 h
E_E **:** 106.87 mm
Rainfall conditions responsible for the landslide
#: 3
Discarded rainfall conditions (threshold value:10%)
#: 1

Rainfall event D_E = 231 h E_E = 106.87 mm

LRC $1_{of\,2}$ D_L = 13 h E_L = 11.6 mm I_L = 0.889 mm/h

LRC $2_{of\,2}$ D_L = 39 h E_L = 42.8 mm I_L = 1.096 mm/h

DRC $1_{of\,1}$ D_L = 63 h E_L = 46.2 mm I_L = 0.733 mm/h
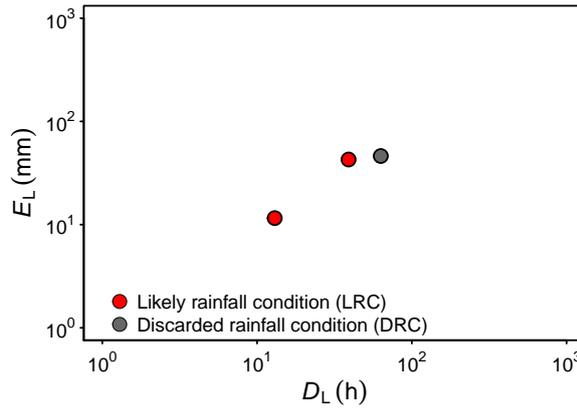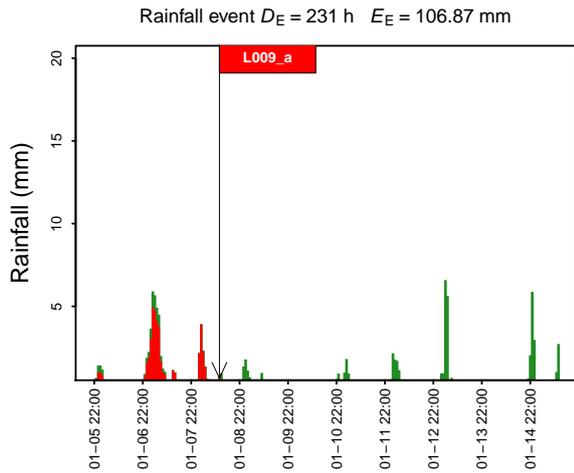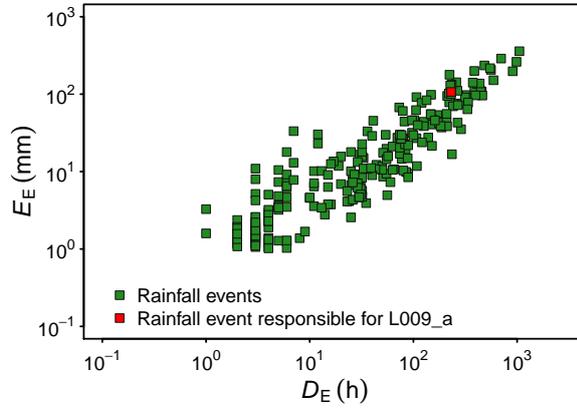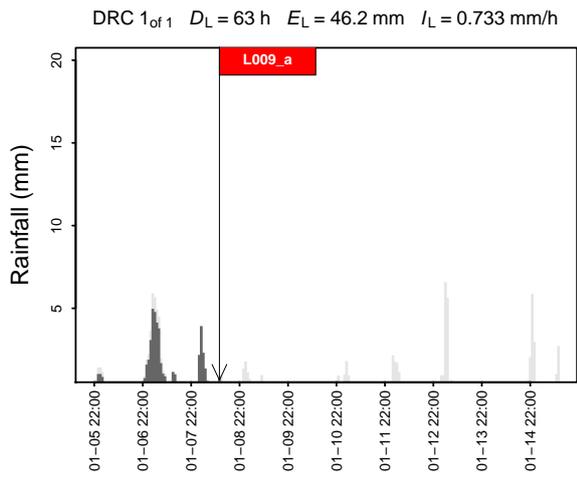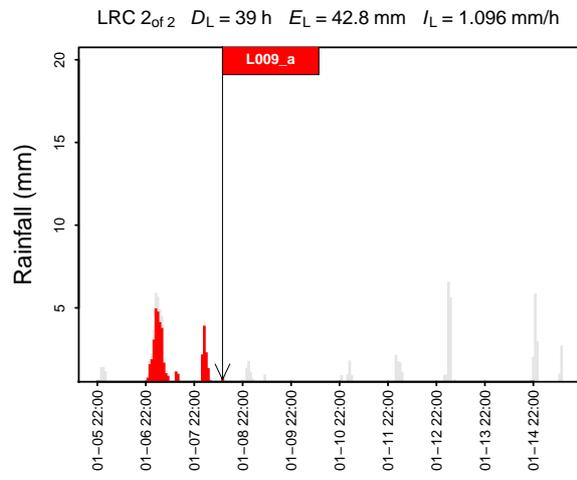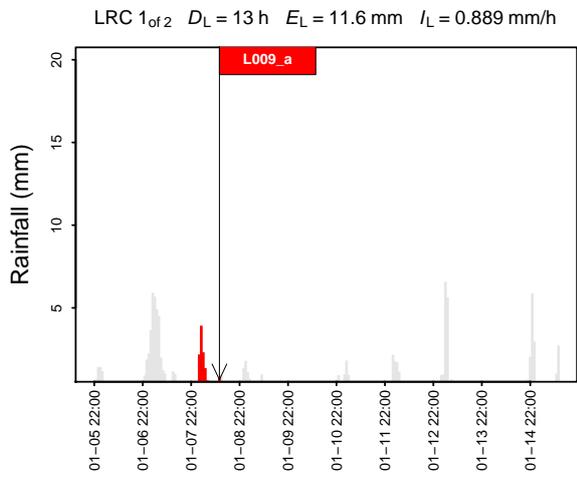
RAIN GAUGE : P022
Distance from the landslide: 8.91 km
Temporal coverage
Start Date: 2013–01–01 07:00+0000
Stop Date: 2018–12–31 23:00+0000
Data Resolution
Temporal Resolution: hourly
Rainfall events
#: 259
Statistics:

	Min	Max	Mean	Median	SD
D_E (h)	1	1002	93.87	31	153.16
E_E (mm)	1.01	265.3	29.6	9.81	46.72

ID LANDSLIDE : L009_a   **DELAY: 0 h**

Landslide date: 14–01–08 12:00
Rainfall event associated with the landslide
#: 48
Start Date: 2014–01–06 01:00+0000
Stop Date: 2014–01–23 17:00+0000
D_E : 425 h
E_E : 115.07 mm
Rainfall conditions responsible for the landslide
#: 4
Discarded rainfall conditions (threshold value: 10%)
#: 0
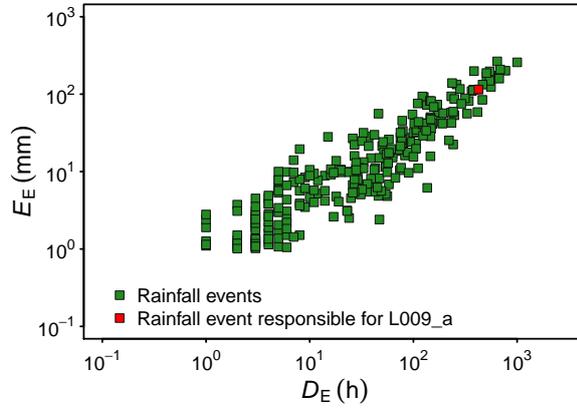
Rainfall event D_E = 425 h E_E = 115.07 mm

LRC $1_{of\ 4}$ $D_L = 1$ h $E_L = 0.2$ mm $I_L = 0.24$ mm/h

LRC $2_{of\ 4}$ $D_L = 14$ h $E_L = 7.9$ mm $I_L = 0.561$ mm/h

LRC $3_{of\ 4}$ $D_L = 38$ h $E_L = 21.5$ mm $I_L = 0.565$ mm/h

LRC $4_{of\ 4}$ $D_L = 60$ h $E_L = 25.9$ mm $I_L = 0.432$ mm/h
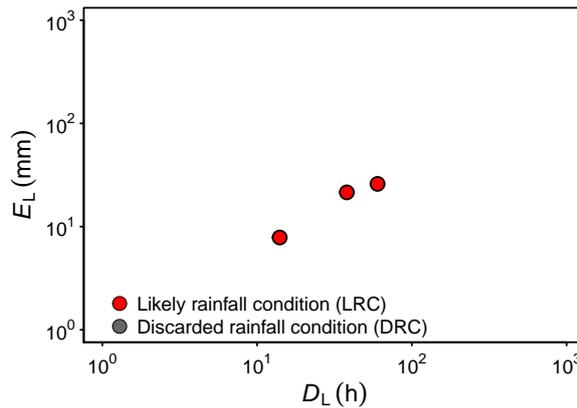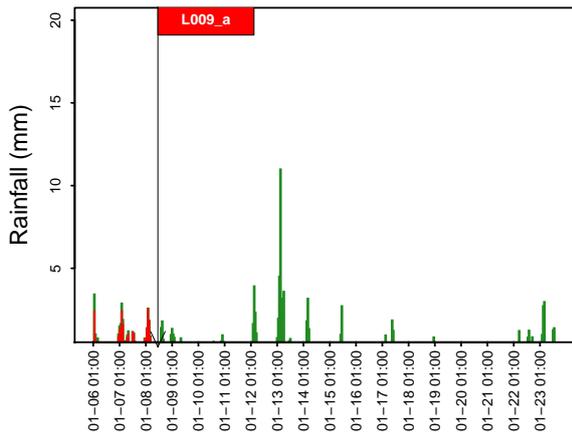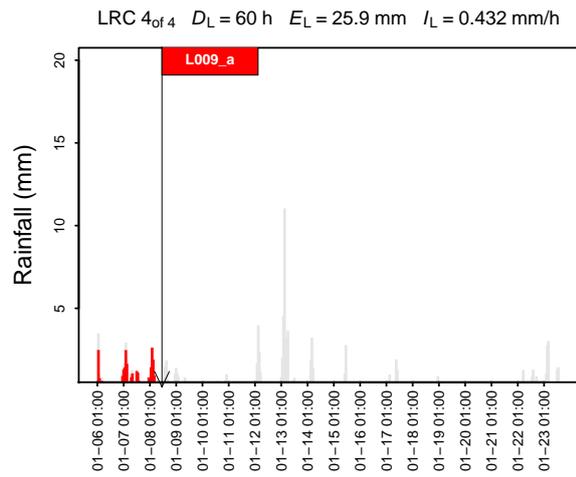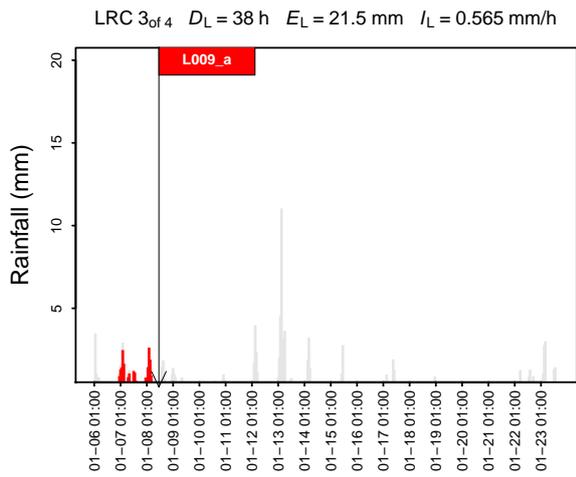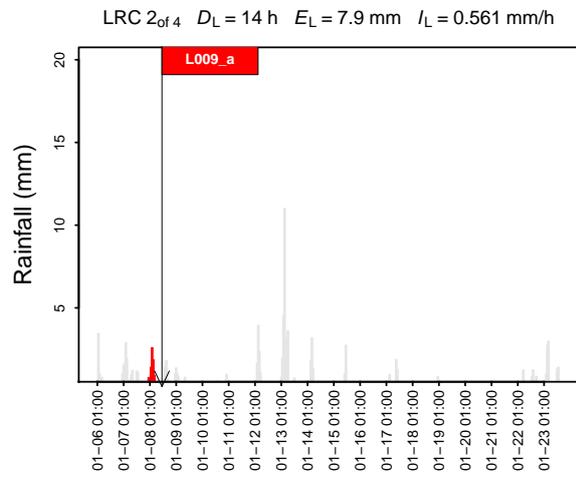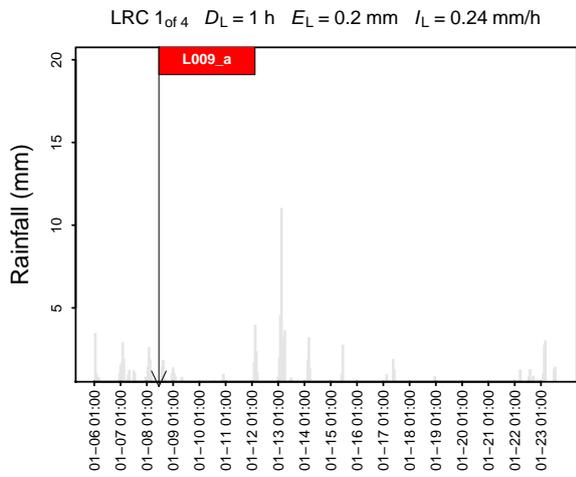
Appendix F | Output Phase 6

For further output files see Drive folder \ outputs

α (7 %)

T 7 %: $E = (2.8 +- 0.3) D^{(0.37\ ?\ 0.03)}$

alpha (7%)

alpha (7%)

$\Delta \alpha = 17.2\, n^{-0.72}$
$R^2 = 0.958$
$\Delta \alpha/\alpha \le 0.1\ n \ge 316$

α (6 %)

T 6 %: $E = (2.7 +- 0.3) D^{(0.37\ ?\ 0.03)}$

alpha (6%)

alpha (6%)

$\Delta \alpha = 16.8\, n^{-0.73}$
$R^2 = 0.958$
$\Delta \alpha/\alpha \le 0.1\ n \ge 315$

α (5 %)

T 5 %: $E = (2.5 +- 0.3) D^{(0.37\ ?\ 0.03)}$

alpha (5%)

alpha (5%)

$\Delta \alpha = 16.2\ n^{-0.73}$
$R^2 = 0.957$
$\Delta \alpha / \alpha \leq 0.1\ n \geq\ 318$

α (4 %)

T 4 %: $E = (2.4 +- 0.3) D^{(0.37\ ?\ 0.03)}$

alpha (4%)

alpha (4%)

$\Delta \alpha = 15.5\ n^{-0.73}$
$R^2 = 0.956$
$\Delta \alpha / \alpha \leq 0.1\ n \geq\ 322$

$\alpha\,(\,3\,\%\,)$

$T\,3\,\%:\,E = (\,2.2 +- 0.3\,)\,D^{(0.37\,?\,0.03)}$

alpha (3%)

alpha (3%)

$\Delta\,\alpha = 14.8\,n^{-0.74}$
$R^2 = 0.955$
$\Delta\,\alpha/\alpha \leq 0.1\,n \geq\,324$

$\alpha\,(\,2\,\%\,)$

$T\,2\,\%:\,E = (\,2.0 +- 0.2\,)\,D^{(0.37\,?\,0.03)}$

alpha (2%)

alpha (2%)

$\Delta\,\alpha = 13.9\,n^{-0.74}$
$R^2 = 0.954$
$\Delta\,\alpha/\alpha \leq 0.1\,n \geq\,328$

α (0.5 %)

E_l (mm)

D_l (h)

T 0.5 %: $E = (1.5 +- 0.2) D^{(0.37 ? 0.03)}$

E_l (mm)

D_l (h)

alpha (0.5%)

α

Number of events,n

alpha (0.5%)

$\Delta \alpha$

$\Delta \alpha = 11.5\ n^{-0.75}$
$R^2 = 0.951$
$\Delta \alpha/\alpha \leq 0.1\ n \geq 348$

Number of events,n

α (0.005 %)

E_l (mm)

D_l (h)

T 0.005 %: $E = (0.7 +- 0.1) D^{(0.37 ? 0.03)}$

E_l (mm)

D_l (h)

alpha (0.005%)

α

Number of events,n

alpha (0.005%)

$\Delta \alpha$

$\Delta \alpha = 7.43\ n^{-0.79}$
$R^2 = 0.944$
$\Delta \alpha/\alpha \leq 0.1\ n \geq 429$

Number of events,n

127