The Potential of Satellite and Model Derived Variables for Rainfall-Induced Landslide Initiation Thresholds in Rwanda

Alessia Riveros Pavez

Technische Universiteit Delft

TU Delft

Deltares

The Potential of Satellite and Model Derived Variables for Rainfall-Induced Landslide Initiation Thresholds in Rwanda

by

Alessia Riveros Pavez

to obtain the degree of Master of Science in
Civil Engineering
Track: Environmental Engineering
Specialization: Environmental Science
at the Delft University of Technology,
to be defended publicly on **January 14 2022**.

Student number:	4818547	
Project duration:	May 2021 – December 2021	
Thesis committee:	Dr. T. Bogaard,	TU Delft
	Dr. ir. F. Sperna Weiland,	Deltares
	MSc. J. Uwihirwe,	TU Delft
	Prof. Dr. ir. S. Steele-Dunne	TU Delft

TU Delft

Abstract

Empirical-statistical rainfall landslide initiation thresholds are popularly used for early warning systems to discriminate between the occurrence and non-occurrence of rainfall-induced landslides. However, the few studies that have derived landslide initiation thresholds for landslide-prone and data-scarce Rwanda rely solely on the limited in situ data. Therefore, our objective is to explore the feasibility of using satellite data and hydrological model derived data to derive both trigger and trigger-cause thresholds for landslides in Rwanda. We firstly evaluated seven precipitation products (TRMM 3B42v7, CHIRPS, PERSIANN-CDR, GLDAS 2.1, CFSv2, IMERG, and ERA5) using the rain gauge data as a reference and found that IMERG was the most suitable product for obtaining rainfall triggering conditions. We then studied the added value of incorporating the antecedent soil moisture from both a high spatial satellite data and from a distributed hydrological model following the trigger-cause framework. The results showed that the event precipitation volume E, the event duration D and the bilinear threshold E-D are the landslide initiation thresholds that accurately predict the highest number of landslide events while keeping the false and the failed alarms low. Including the antecedent soil moisture products as the causal variables -expected to account for the hillslope hydrologic processes predisposing the slopes to near failure- did not lead to any improvement with respect to the trigger only thresholds for predicting landslides in Rwanda.

Contents

List of Figures

List of Tables

1
Introduction

Landslides are natural hazards that occur worldwide. They do not only cause deaths every year but also economic damages. Landslides are often unnoticed and compared to other hazards, their consequences are understated (Guzzetti, 2021). Furthermore, the continuous development in mountainous areas increments the exposure of people and properties to landslide hazards (Bogaard and Greco, 2018). To address the devastating effects of landslides on human lives and the economy, landslide early warning systems (LEWSs) have been developed. An early warning system (EWS) is a device, system or set of capabilities that generates and disseminates timely and meaningful information to enable individuals, communities, and organizations threatened by a hazard to act timely and appropriately to avoid or to reduce the impact of the threat (Guzzetti et al., 2020). As such, operational LEWSs contribute to mitigating landslide risks by minimizing the fatalities and the economic losses (Guzzetti et al., 2020) yet they are in place only in limited areas.

Although landslides can be triggered by a number of factors such as earthquakes and mining, the most frequent trigger remains rainfall (Bogaard and Greco, 2016; Zhao et al., 2019). More precisely, the most widespread are shallow rainfall-induced landslides. Predicting these landslides is generally based on physically-based deterministic models or empirical-statistical rainfall thresholds. The former rely mainly on distributed models for slope stability, hydrology and infiltration. At a slope or catchment scale, the physically-based models have been successful though they can be sensitive to small errors in the subsurface data. However, at larger scales, the soil heterogeneity and variability result in a poor knowledge of the hydrological and geotechnical parameters' spatial distribution. Thus, the high data requirements, (and the need for a well-calibrated model) limit its application in LEWSs (Bogaard and Greco, 2018). On the other hand, empirical rainfall thresholds have been commonly used in both local and regional landslide hazard assessment, in particular, the intensity-duration (ID) threshold. These require precipitation time series alongside a landslide inventory with a high spatial and temporal resolution (Bogaard and Greco, 2018). Meteorological thresholds are usually based on the recent, current, or imminent storm conditions, and sometimes jointly used with previous rainfall accounting for antecedent wetness (Mirus et al., 2018a).

Nevertheless, landslides are not directly triggered by rainfall (Mirus et al., 2018a) rather by subsurface hydrology (Sidle et al., 2019). Indeed, the accumulation of water in the subsurface leads to an increased buoyancy force (pore water pressure buildup) exerted on the soil: this results in the reduction of shear strength in the saturated soil, and in the reduction of soil suction in the unsaturated soil until failure occurs (Bogaard and Greco, 2016). Accordingly, processes such as precipitation, infiltration or bedrock exfiltration contribute to the failure of slopes whereas drainage, evaporation, and transpiration contribute to the stabilization of hillslopes.

Hence, to include the relevant hydrologic processes that play a key role in landslide initiation, some empirical rainfall threshold studies have also included antecedent precipitation, e.g. Chleborad et al. (2008). The antecedent precipitation (or the accumulated rainfall over a certain period of time) indirectly considers the soil moisture conditions of the ground before the materialization of landslides (Bogaard and Greco, 2018). However it still does not add much information as it is solely based on precipitation. In fact, representing only with rain the hydrological processes that play a role in predisposing the slopes to failure such as infiltration, evaporation, transpiration and soil drainage may be troublesome (Mirus et al., 2018a). Failing to account for such hydrological processes could partially explain why meteorological thresholds that are designed to predict the occurrence of rainfall induced landslides also forecast landslides when there are none (false alarms) and miss predicting landslides that do take place (missed alarms) (Mirus et al., 2018a; Marino et al., 2020). In this research, we test this hypothesis, i.e. whether incorporating antecedent soil moisture increases the explanatory power of the empirical-statistical thresholds. For landslide thresholds to be effective, there should be a high true positive rate (correctly predicted landslide occurrence), and a low false positive rate (incorrectly predicted landslide occurrence) and missed alarms.

Consequently, explicitly accounting for hydrological processes in these empirical-statistical thresholds such as a direct measure or proxy for antecedent soil water content adds physically relevant information (Bogaard and Greco, 2018). Specifically, Bogaard and Greco (2018) propose the trigger-cause concept for regional landslide hazard assessment. The cause is the antecedent hydrological conditions predisposing the slopes to near-failure while the trigger is the rainfall intensity responsible for the last push initiating the landslide. An example of cause is the antecedent soil moisture which is a determining factor to include in shallow landslide forecasting (Marino et al., 2020).

Rain gauge data is the most common source to derive rainfall thresholds (Nikolopoulos et al., 2015) though the spatial variability of rainfall is an issue. The same is true for in situ soil moisture observations despite their accuracy, they remain only point measures (Owe et al., 2008). However, employing real-time gauge networks is more troublesome in poorer, landslide-prone regions because maintaining these networks is not only expensive but also complicated in complex terrains (Brunetti et al., 2018). Moreover, Nikolopoulos et al. (2015) find that the resulting meteorological landslide initiation threshold is conditioned to the rain gauge density: the lower the density, the larger the underestimation and estimation variance. Therefore, the applications are limited to denser rain gauge networks (Wang et al., 2021).

Alternatively, remotely sensed data can provide a comprehensive view of landslide hazard and promote landslide monitoring and prediction (Stanley et al., 2021). The advantages of using satellite data include human-independent information, wide and consistent coverage and operational delivery of information (Skakun et al., 2016). Nonetheless, the number of studies developing forecasts for landslide events based on satellite data is limited even though remote sensing precipitation products provide estimates at regional and global scales. This is especially relevant for monitoring areas with scarce rain gauge networks (Marra et al., 2017). Besides, employing satellite precipitation products allows a direct and consistent comparison between thresholds developed for distinct regions (Marra et al., 2017). The same applies to soil moisture remote sensing estimates. Additionally, another option for soil moisture estimation is to use conceptual hydrologic models (Zhao et al., 2019). As opposed to the previously mentioned spatially distributed physical models, these do not require large data input nor calibration. Thus, hydrological simulations providing relevant subsurface data have already been employed to study landslide initiation thresholds (e.g. Zhao et al., 2019 and Bezak et al., 2019).

This study focuses on Rwanda, a tropical landslide-prone country located in Central-East Africa. This region is characterized by a lack of detailed data on landslide occurrences in time and space (Monsieurs et al., 2018) and a lack of adequate rainfall records from ground monitoring networks. Consequently, there has been limited research on landslides over the tropics in Africa and specifically Rwanda. As a first step towards robust LEWS in Rwanda, Uwihirwe et al. (2020) used an empirical-statistical approach based on rain gauges to derive landslide initiation thresholds. Along the same line, Uwihirwe et al. (2021) incorporated regional groundwater level measurements extended with a model to the empirical-statistical landslide initiation thresholds for the Kivu, the upper Nyabarongo, and the Mukungwa catchments in Rwanda.

Notwithstanding, these studies rely exclusively on in situ data. Therefore, the objective of this research is to instead use data from satellites and from a distributed hydrological model to derive landslide initiation thresholds for Rwanda. Specifically, the following research questions are answered:

(i) From the freely available satellite precipitation products on the Google Earth Engine (GEE) platform, which one is most suitable to derive the rainfall triggering conditions for landslide initiation in Rwanda?

(ii) Which satellite derived variables or combination thereof increase the predictability of landslides in Rwanda?

(iii) Which distributed hydrological model derived variables or combination thereof increase the predictability of landslide in Rwanda?

The structure of this report is as follows: chapter two describes the study area along with the satellite data, and the distributed hydrological model data. In chapter three, the seven satellite precipitation products are analyzed to find the product that best matches the rain gauges' observations. Chapter four provides a quality assessment of both the satellite and the distributed hydrological model soil moisture products. In chapter five, the landslide initiation thresholds are derived based on (i) the satellite data, and (ii) the distributed hydrological model to find the variable(s) or combinations that have the highest explanatory power to predict landslides in Rwanda following the trigger and trigger-cause framework. Lastly, chapter 6 presents the conclusions of this research.

2

Study Area and Data

2.1. Study Area

The Republic of Rwanda, south of the Equator, is a landlocked-country located in the western branch of the East African Rift and is part of the African Great Lakes region. It is bordered by Uganda to the north, Tanzania to the east, Burundi to the south and the Democratic Republic of the Congo to the west. Its total area is equal to 26338 km^2. The north and western regions are dominated by the Virunga volcano chain reaching a maximum elevation of 4519 m.a.s.l.. The lowest point (950 m.a.s.l.) is situated in the south-west, in the Rusizi river. The lake Kivu is located on the west edge of the country. Towards the east and south of the country, the mountains decrease in height.

2.1.1. Climate

Rwanda's climate is classified as tropical Savannah (Peel et al., 2007). The average annual temperature ranges between 15 °C to 17 °C in the mountains, and up to 30 °C in the lowlands (Figure 2.2) (The World Bank Group, 2021). In the high altitude areas, the mean annual rainfall varies from 1200 to 1500 mm while in the plateaus it ranges from 900 to 1200 mm (Demarée and Van deVyver, 2013) and, in the Savannah it reduces to less than 1000 mm (Uwihirwe et al., 2020). The complex topography (mountains, low lands, inland water bodies) greatly influences the climate of the region on a local scale playing a role in the low-level circulation and moisture transport (Cattani et al., 2016). Rwanda experiences two rainy seasons per year: the "long rains" from March to May (MAM) which are heaviest and the "short rains" from October to November (ON) (Nicholson, 2017) or traditionally from October to December (OND). During the MAM, both convective and orographic rainfall regimes coexist whereas during the OND, (warm) orographic rainfall is dominant (Kimani et al., 2017). The short rains are influenced by El Niño Southern Oscillation (ENSO), the Indian Ocean Zonal Mode (IOZM) and the zonal winds. Conversely, the variability of the long rains are due to the Madden-Julian Oscillation (MJO) and to the Pacific and Indian Ocean anomalies. Recent trends indicate that the interannual variability has increased (Nicholson, 2017). Rwanda is characterized by frequent extreme rainfall events and flooding which are both precursors to landslide hazards (Uwihirwe et al., 2020).

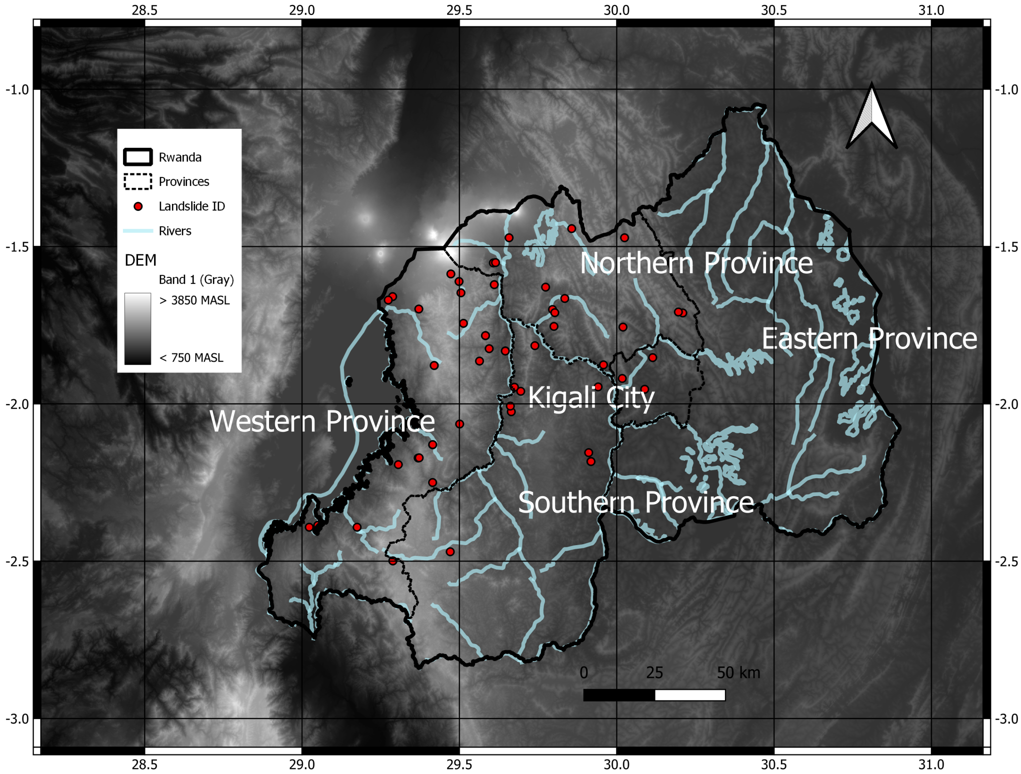

Figure 2.1: Digital Elevation Model (DEM) of Rwanda with Shuttle Radar Topography Mission (SRTM) 90 m digital elevation model (Jarvis et al., 2008) available at: `https://srtm.csi.cgiar.org/srtmdata/` along with the recorded landslides from 2007-2019 (red dots).

2.1.2. Lithology and Land Cover

In the areas that have recorded landslide events, the prevailing lithological units are mica schists and pegmatite rocks which are weak due to rapid weathering, easy splitting and loss of strength (Uwihirwe et al. 2020). Figure 2.3 shows that most of the land use in Rwanda corresponds to cultivated and managed vegetation or agriculture.

2.2. Data

2.2.1. Landslide Inventory

The landslide inventory is based on the NASA global landslide catalogue (`https://data.nasa.gov/Earth-Science/Global-Landslide-Catalog/h9d8-neg4`) primarily uploaded by the landslide inventory for the central section of the western branch of the East African Rift (LIWEAR) project which has been further extended by Uwihirwe et al., (2020). We used the landslide events ranging from 2007 to 2019, excluding the events that did not have a specific date or were not caused by rainfall leading to a total of 55 recorded landslides. The inventory provides the exact location of each recorded landslide event but with a varying accuracy from 5 km to 25 km. It should be noted that this catalog is likely to miss the non-hazardous landslides which are less reported than hazardous landslides that lead to fatalities/injuries and damages.

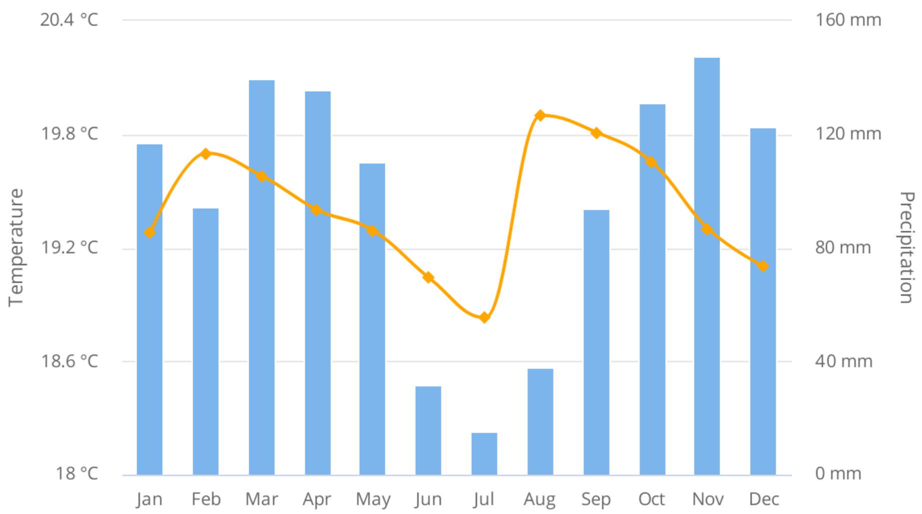

Figure 2.2: Monthly mean temperature (in orange) and precipitation (in blue) 1991-2020 (The World Bank Group, 2021).

2.2.2. Precipitation

We used precipitation data from the rain gauges installed in Rwanda provided by the Rwanda Meteorology Agency (Figure 2.4) to select the most adequate satellite precipitation product. The rain gauge data are provided at an (almost) daily frequency covering a varying period between 2006-2018. The table in the appendix (Table A.1) provides more details about the selected rain gauges.

Satellite products' estimates primarily come from infrared (IR) sensors on board geostationary satellites and/or passive and active microwave (MW) sensors on board low-Earth orbiting satellites (Kimani et al., 2017). The combination of both allows to take advantage of the high temporal resolution of IR and the better accuracy of rainfall estimation from MW sensors. Furthermore, many satellite rainfall products use rain gauge data for bias correction (le Coz, 2021).

As many satellite rainfall products exist, we carried out an assessment of several distinct ones. We made the pre-selection on the basis that the dataset (i) provides coverage spanning the entire landslide inventory (2007-2019), (ii) is at least daily, and (iii) can be obtained via Google Earth Engine (GEE). It should be noted that the global datasets available on GEE are only a subset of all the available ones. Table 2.1 shows the datasets that are compliant with these conditions. A brief overview of each of the data set downloaded can be found in Appendix A.

After determining the most suitable satellite precipitation product, we used a 10-km diameter buffer at every landslide location to download the precipitation time series in GEE (Figure 2.4).

2.2.3. Soil Moisture

We employed the automatic weather stations (AWSs) (Figure 2.5) provided by the Rwanda Meteorology Agency to assess the quality of the soil moisture products. The selected AWSs measure the soil moisture at a depth of 20 cm with a varying temporal resolution, the highest being every 5 minutes though the measurements contain large data gaps.

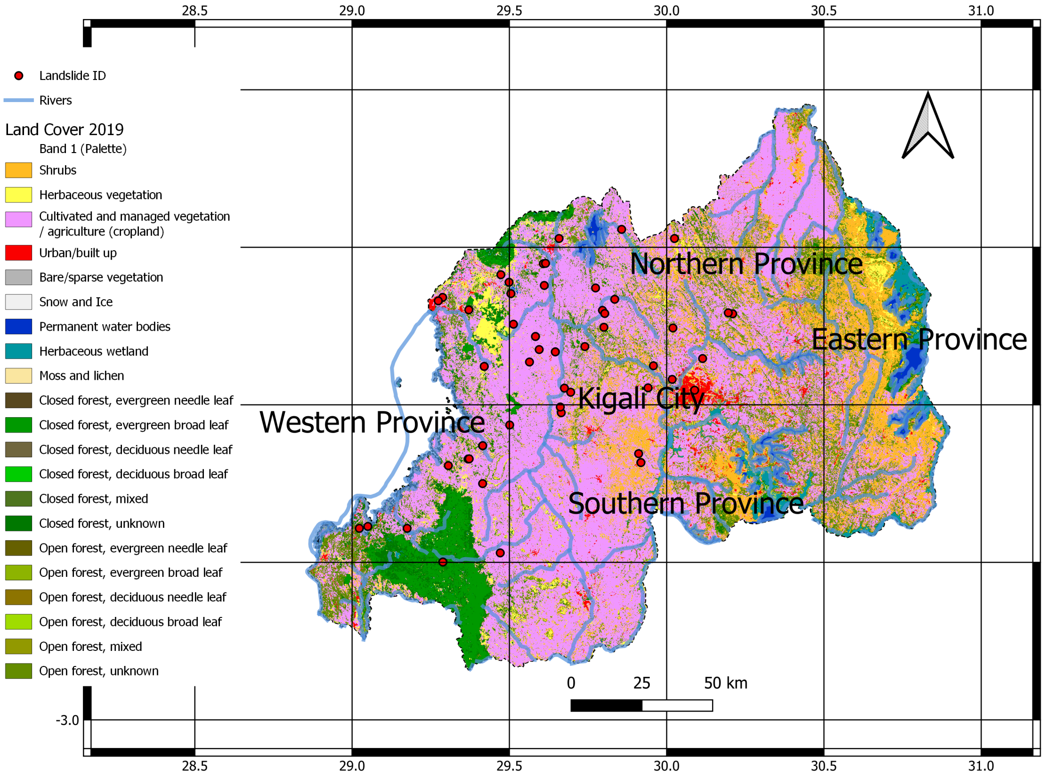

Figure 2.3: Land Cover 2019 (Buchhorn et al., 2020) along with the recorded landslide events from 2007-2019 (red dots).

VanderSat provided the L-band volumetric soil moisture product which estimates the top 5 cm soil moisture θ_{top}. It is retrieved from satellite microwave measurements and down-scaled from typical spatial resolution of 25 km x 25 km to 100 m x 100 m using a patented algorithm. The volumetric soil moisture is derived using the Land Parameter Retrieval Model (LPRM) (Owe et al., 2001; Owe et al., 2008; de Jeu et al., 2014) that has been further developed (van der Schalie et al., 2017). The L-band product is retrieved from the Soil Moisture Active and Passive (SMAP) satellite for the period of April 2015 until present. In this period, an observation is provided every varying number of days, hence the values are extrapolated using a 20-day backward moving average window. The SMAP dataset is extended back until 2002 using the modeled ERA-5 data providing an almost daily soil moisture value.

Since a 5-km location accuracy is the most commonly reported in the landslide catalog, we used a 5-km radius buffer as a region of interest (ROI) around each landslide location. As the ROIs cannot overlap each other, we modified some of the buffers (Table A.2) leading to the final ROIs shown in Figure 2.5. The ROIs that had unrealistic soil moisture time series were not included in the remainder of the analysis and instead we used the closest available buffer to the landslide event.

Wflow_sbm is a distributed hydrological model using the conceptual bucket approach (Schellekens et al., 2021; Figure 2.6). This model is based on model parameters that are estimated a priori using pedotransfer functions (reducing the number of parameters) and global or local datasets. For this particular case, we included the presence of natural lakes and reservoirs. The spatial resolution of the output data is about 1 km.

Table 2.1: Pre-selected satellite datasets.

Dataset	Resolution	Frequency	Period	Data source	Reference
TRMM 3B42 v7	0.25°	Daily	1998 end of 2019	Gauge, MW, IR	(Huffman et al., 2007; Huffman and Bolvin, 2018)
CHIRPS	0.05°	Daily	1981 near present	Gauge, MW, IR	(Funk et al., 2015)
PERSIANN -CDR	0.25°	Daily	1983 near present	Gauge, IR	(Ashouri et al., 2015)
GLDAS 2.1	0.25°	3 hourly	2000 near present	MW, IR	(Rodell et al., 2004)
CFSv2	0.2°	6 hourly	1979 near present	Gauge, MW, IR	(Saha et al., 2010; Saha et al., 2014)
IMERG	0.1°	30 min	2000 near present	Gauge, MW, IR	(Hou et al., 2014; Huffman et al., 2018)
ERA5	0.25°	Daily	1979 near present	NWP	(Hersbach et al., 2020)

Wflow_sbm uses the kinematic wave routing approach for channel, overland, and lateral subsurface flow. The soil is divided into a saturated and an unsaturated storage where the depth of the saturated zone constantly varies. For Rwanda, the total soil profile is subdivided into four layers of 100, 300, 800, and 800 mm. Water experiences evapotranspiration and interception following the Gash model (Gash, 1979). Moreover, it has a spatially distributed gridded cell network based on the D8-network flow routing. Additionally, the vertical water transfer is controlled by the saturated hydraulic conductivity, the effective saturation degree of the layer, and the Brooks-Corey power coefficient based on the pore size distribution.

We forced the hydrological model from 2001-01-02 to 2020-12-31 with ERA 5 (Hersbach et al., 2020) precipitation, temperature, radiation and pressure and computed the potential evaporation following de Bruin (1983) on the same three catchments (Kivu, upper Nyabarongo and Mukungwa) of Rwanda as in Uwihirwe et al., (2021) (Figure 2.5). The first two years are considered as spin-up period, hence, were excluded from the analysis. To increase the comparability with the satellite-based soil moisture, we employed the same buffers (Figure 2.5). The variables of interest derived from the model are the volumetric water content in the root zone, θ_{root} [-], representative of the top 50 cm approximately, and the part of the soil water capacity occupied, θ_{uz} [-] representative of the top two meters approximately. θ_{uz} is computed following Equation 2.1 where θ_{sat} is the saturated store [mm], U is the total amount of available water in the unsaturated zone [mm], and θ_{cap} is the soil water capacity in [mm].

$$\theta_{uz} = \frac{\theta_{sat} + U}{\theta_{cap}} \tag{2.1}$$

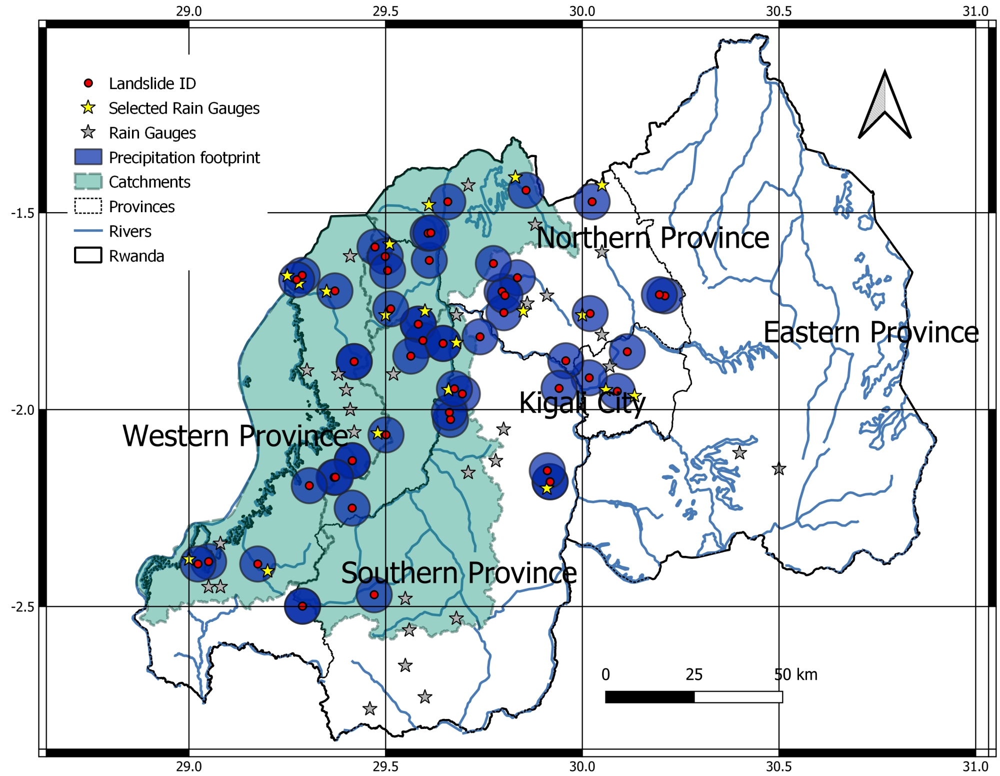

Figure 2.4: Buffers (in blue) used to download the selected satellite precipitation product via GEE. The rain gauges are plotted as stars. In particular, the ones used to select the best satellite product are colored in yellow.

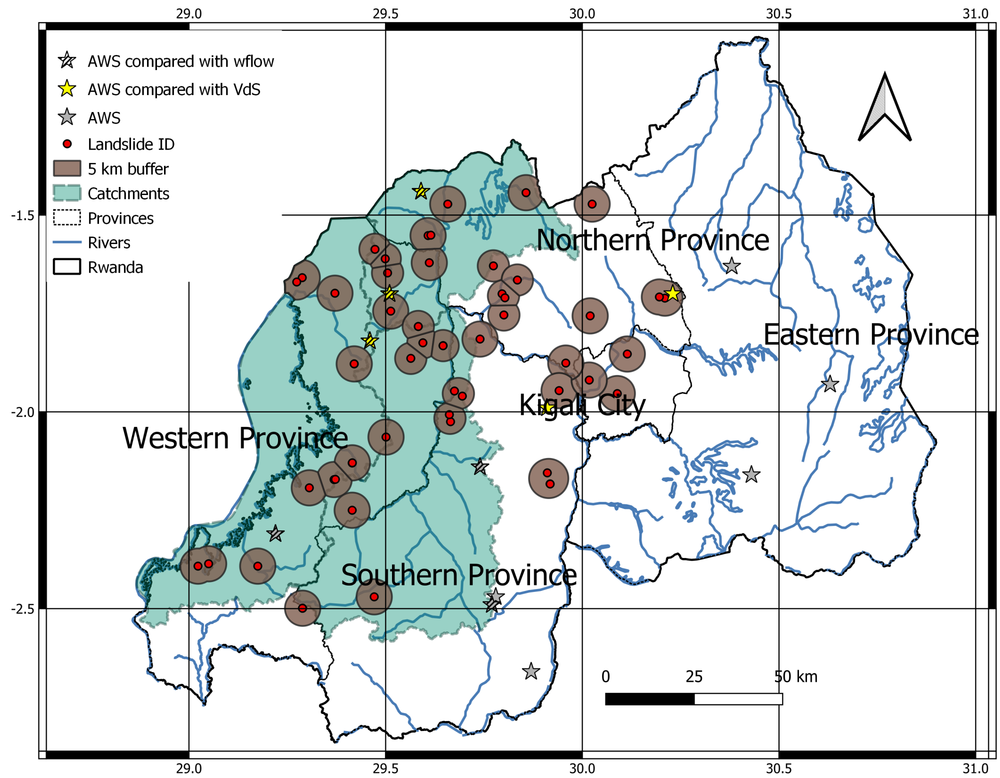

Figure 2.5: ROIs uploaded to the VanderSat API portal. The ROIs inside the Kivu, upper Nyabarongo and Mukungwa catchments are also used to extract data from the wflow_sbm model simulations.

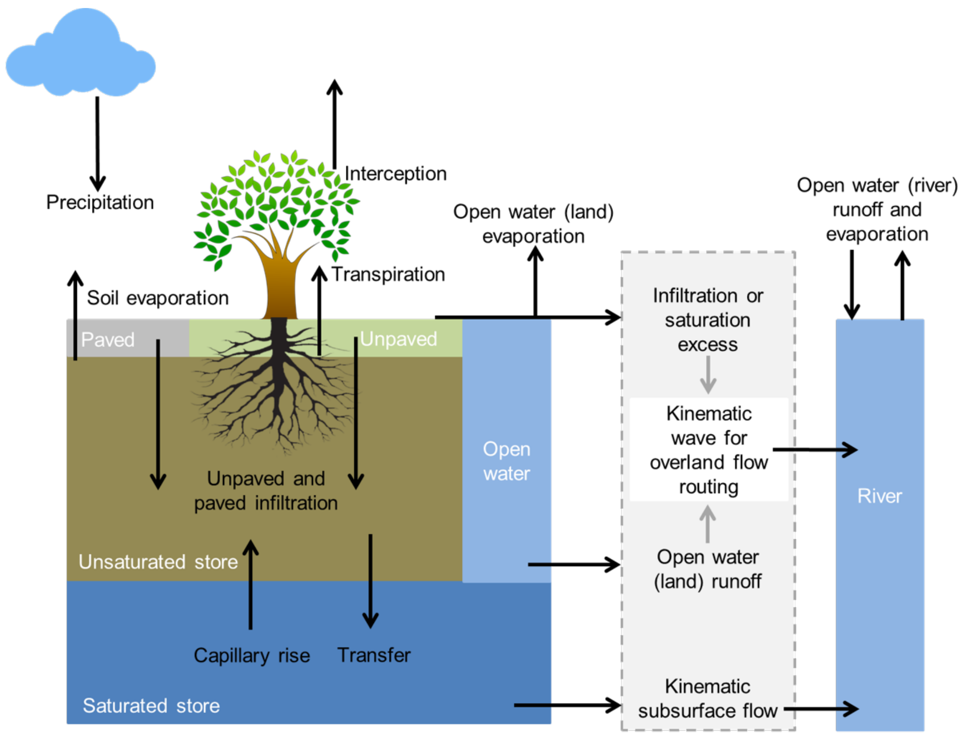

Figure 2.6: Overview of the different processes and fluxes in the wflow_sbm model (Schellekens et al., 2021).

3

Satellite Precipitation Products

3.1. Introduction

Temporally and spatially accurate rainfall data are fundamental for robust LEWS development. Guzzetti et al., (2020) recommended employing multiple sources of rainfall information in LEWSs. In situ rain gauges have already been explored to derive empirical-statistical landslide initiation thresholds for Rwanda (Uwihirwe et al., 2020; Uwihirwe et al., 2021). Rain gauges directly measure the rain that reaches the ground surface though they are land-based, sparse, and remain point measurements (Ashouri et al., 2015). Furthermore, Wang et al. (2021) emphasized that any assessment relying on rain gauges is constrained to areas that are well equipped. This is not generally observed in landslide-prone regions due to the mountainous terrains. Alternatively, remote sensing rainfall products provide estimates particularly important for areas with a sparse gauge network (Marra et al., 2017) although these are limited by (i) the coarse resolution, (ii) the uncertainty intrinsically related to the retrieval algorithms, and (iii) the underestimation of rainfall (Brunetti et al. 2018; Chikalamo et al., 2020). Brunetti et al. (2018) showed that even though the satellite rainfall products underestimate the observed rain gauge rainfall, it is not an issue to the development of landslide warning system, as long as the product is not biased in terms of rainfall regimes and locations. Recently, Wang et al. (2021) employed satellite rainfall data sets to detect the rainfall conditions for the initiation of hydro-morphological processes. Yet, the use of satellite rainfall data for landslide forecasting is still limited as opposed to rain gauges or weather radars (Abancó et al., 2021; Brunetti et al., 2018). Satellite precipitation products can play a key role, in particular, in the East African region where the rain gauge density is low (Cattani et al., 2016). Nonetheless, it is important to keep in mind that satellite precipitation estimates over this region are challenging because of the complex terrain, the clear seasonal and geographic-dependence of rainfall (chapter 2) (Cattani et al., 2016).

The purpose of this chapter is to answer the following research question: from the freely available satellite precipitation products on the GEE platform, which one is most suitable to derive the rainfall triggering conditions for landslide initiation in Rwanda?

3.2. Methods

We used rain gauge data as the reference data to assess each satellite product. From the 48 total available rain gauge locations, we chose the closest 19 to past landslide events (Figure 2.4). Thus, we evaluated each dataset using a point-to-pixel approach for the selected rain gauge locations. It should be noted that we did not make the assessment at a scale finer than the daily one because the time of occurrence of landslide events is available on a daily basis.

We assessed the consistency or correlation between the rain gauges observations and the satellite precipitation products via commonly used indicators as shown in Table 3.1. Y_i is the rain gauge observation at date i, X_i is the satellite estimate at the same date i, and n is the total number of data pairs for each satellite precipitation product considered.

Table 3.1: Correlation indicators.

Indicator	Acronym	Equation	Unit
Pearson's correlation	CC	$CC = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}}$	[-]
Root mean square error	RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - X_i)^2}{n}}$	[mm]
Long term relative bias	RB	$RB = \frac{\overline{Y_i} - \overline{X_i}}{\overline{Y_i} + \overline{X_i}}$	[-]

To measure aspects of frequency, we calculated the number of rainy days above various thresholds (Table 3.2) for all products including the rain gauge observations following the guidelines established by WMO (Tank et al., 2009) on analysing extremes. This type of standardized indices allow for consistency and comparison between different sources of rainfall data.

Table 3.2: Rainfall frequency indicators.

Indicator	Acronym	Definition	Unit
Number of rainy days	RD	Count of days when rainfall 0 mm	[days]
Number of heavy rainfall days	R10mm	Count of days when rainfall 10 mm	[days]
Number of very heavy rainfall days	R20mm	Count of days when rainfall 20 mm	[days]
Number of even heavier rainfall days	R30mm	Count of days when rainfall 30 mm	[days]
Number of extremely heavy rainfall days	R50mm	Count of days when rainfall 50 mm	[days]

We plotted the accumulated moving average of 5-days, 15-days, and 30-days rainfall for all satellite precipitation products versus the same accumulated period of the rain gauges' observations along with an x=y line to visualize how closely the satellite products match the in situ measurements.

3.3. Results

Table 3.3 shows that the microwave satellite products (IMERG, TRMM42 and CHIRPS) have the highest correlation coefficient, specifically IMERG with a coefficient of 0.35. This value is quite low and can be attributed to the overestimation of rainy days (Table 3.4). The infrared-based product PERSIANN-CDR has the lowest RMSE. Conversely, CFSv2 and ERA5, both reanalysis, have the highest RMSE. For the rest of the products

(TRMM, CHIRPS, GLDAS and IMERG), the RMSE have a value between 7 and 8. CHIRPS exhibits the lowest RB followed by IMERG and GLDAS. It is important to mention that the only satellite product with missing values during the studied period is PERSIANN-CDR which is a considerable drawback in selecting it to study landslide initiation.

Table 3.3: Mean of correlation between satellite precipitation products and rain gauge observations at selected locations.

	TRMM42	**CHIRPS**	**PERSIANN-CDR**	**GLDAS-2.1**	**CFSv2**	**IMERG**	**ERA5**
CC	0.31	0.27	0.25	0.24	0.17	0.35	0.22
RMSE [mm]	8.17	8.53	7.42	8.55	10.58	8.18	12.60
RB	-0.08	-0.01	-0.15	0.03	0.11	0.02	0.29

CHIRPS has on average 1256 RD, the closest to the gauges' average 1259 RD as opposed to all other products which overestimate the number of rainy days (Table 3.4). This is in line with the results from Beck et al., (2017) in which the reanalysis products consistently underestimated the number of dry days across the globe. These values can be visualized spatially in Figure 3.1.

Regarding higher values of rainfall (Table 3.4), IMERG only slightly underestimate them with R10D = 383, R20D = 126, R30D = 42, R50D = 6 days compared with the gauges' 397, 132, 49 and 9 days respectively. CFSv2 and ERA5 consistently overestimate the number of days above the various thresholds whereas TRMM42, PERSIANN-CDR and GLDAS persistently underestimate R10D, R20D, R30D and R50D. In particular, PERSIANN-CDR experiences the heaviest underestimation out of all the products for all indicators except rainy days. Indeed, PERSIANN-CDR misses the localized high intensity rain in high ground areas in East Africa (Kimani et al., 2017). Lastly, CHIRPS overestimates R10D but underestimates R20D, R30D and R50D.

Table 3.4: Mean of frequency indicators at selected locations.

	Gauge	**TRMM42**	**CHIRPS**	**PERSIANN -CDR**	**GLDAS2.1**	**CFSv2**	**IMERG**	**ERA5**
RD	**1259**	1691	*1256*	2732	3086	2835	2842	3520
R10D	**397**	307	424	138	377	617	*383*	879
R20D	**132**	87	101	9	79	199	*126*	250
R30D	**49**	29	25	0	22	84	*42*	78
R50D	**9**	4	3	0	2	22	*6*	21

Additionally, the scatter plots of accumulated rainfall for 5-days, 15-days and 30-days for the satellite precipitation products versus the rain gauge observations shed light on how they relate to each other (Figure 3.2, Figure 3.3, Figure 3.4 respectively). It is evident that the reanalysis products (GLDAS, CFSv2 and ERA5) tend to systematically overestimate rainfall, especially for lower rainfall depths compared to the in situ measurements, becoming clearer for longer periods of accumulated precipitation. Conversely, PERSIANN-CDR, and to a lesser extent, TRMM42 underestimate rainfall compared to the rain gauge recorded values. Both CHIRPS and IMERG scatter plots follow more closely the $y = x$ line which means that their values are closer to the ones recorded by the in situ observations. The correlations are higher for these multiple day events than for the daily rainfall as calculated by the Pearson correlation in Table 3.3.

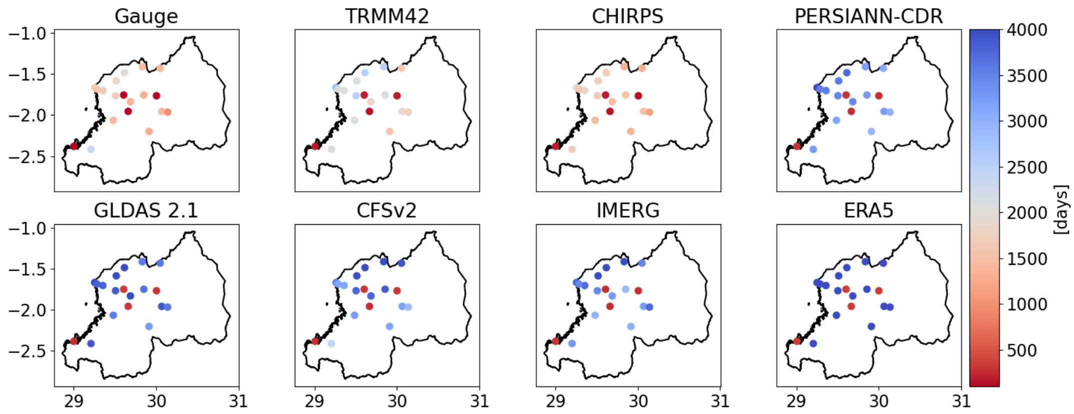

Figure 3.1: Number of rainy days (RD) over the period covered by the rain gauges per selected location.

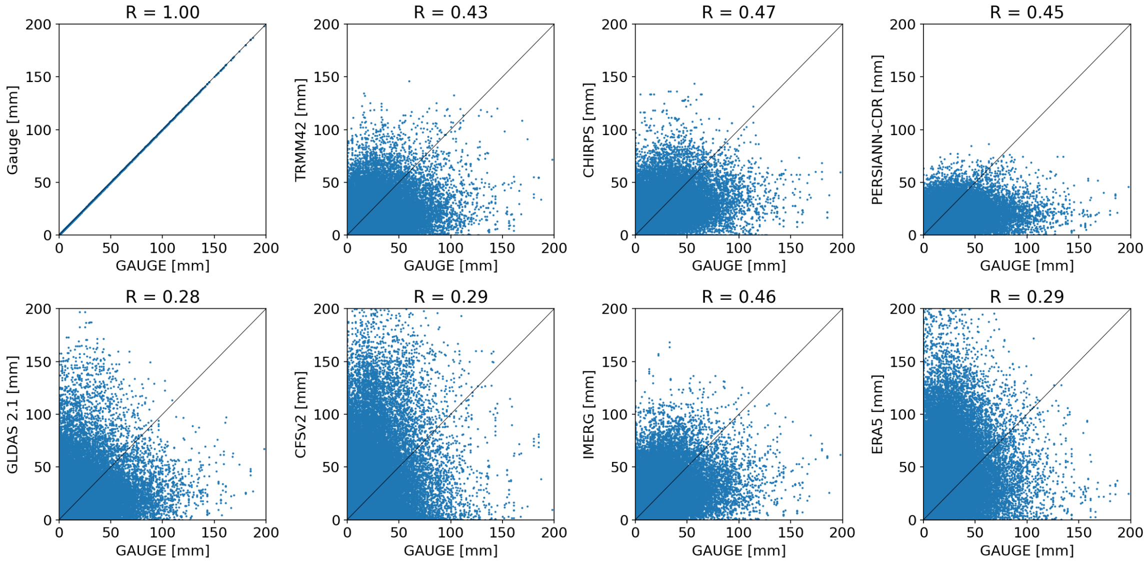

Figure 3.2: Scatter plots of 5-day accumulated rainfall for each precipitation product versus the rain gauge observations.

3.4. Discussion and Selection of Precipitation Product

Regarding the correlation between the satellite products and the rain gauges, IMERG appears to be the product performing the best as it has the highest CC (0.35), the second lowest RB (0.02) after CHIRPS, and the third lowest RMSE after PERSIANN-CDR and TRMM42. On the other hand, the ones that look the least promising are PERSIANN-CDR, CFSv2 and ERA5. Despite having the best RMSE (7.42), the infrared-based PERSIANN is the single one with missing values which is not recommended for landslide initiation studies. The reanalysis products CFSv2 and ERA5 show the lowest CC, the highest RMSE, and along with PERSIANN-CDR the highest RB. In their global-scale evaluation of precipitation products, Beck et al., (2017) determined that the reanalyses exhibit lower skill levels than the microwave- and infrared-based satellite datasets in the tropics.

As to the frequency indicators, although CHIRPS has the closest number of rainy days to the rain gauges', IMERG, which has the most similar R10D, R20D, R30D and R50D to the in situ measurements is preferred. It is most important to accurately obtain the heavier rainfalls events because they are most likely to be respon-

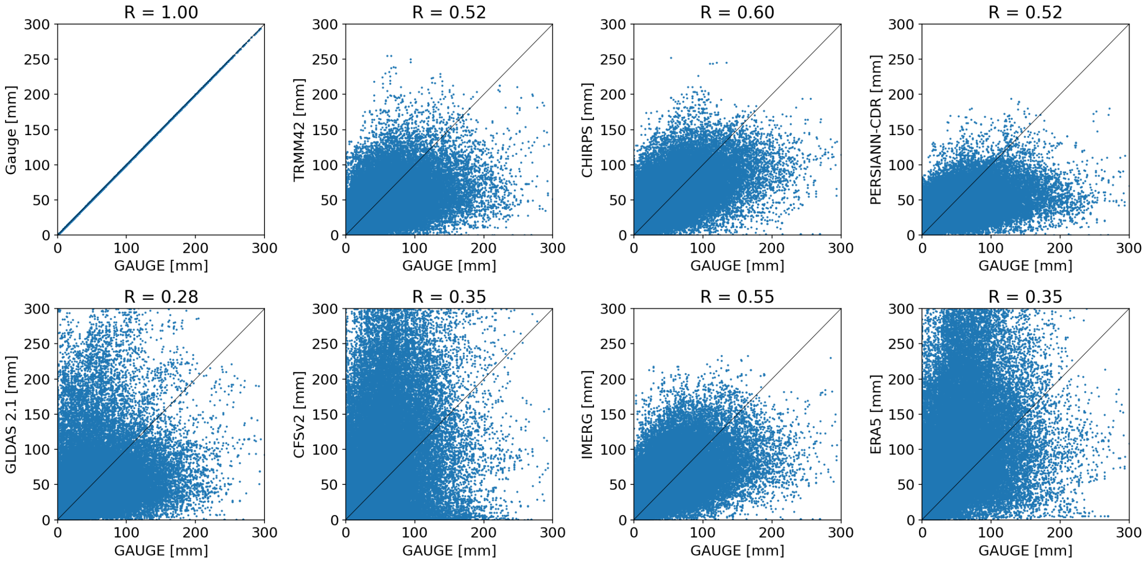

Figure 3.3: Scatter plots of 15-day accumulated rainfall for each precipitation product versus the rain gauge observations.

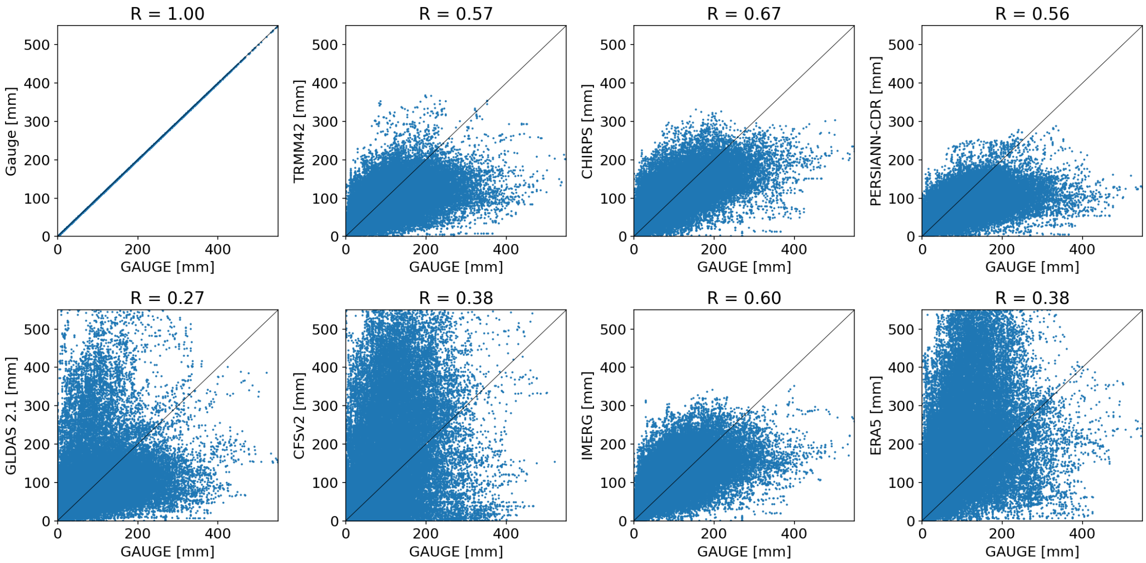

Figure 3.4: Scatter plots of 30-day accumulated rainfall for each precipitation product versus the rain gauge observations.

sible for triggering landslides. TRMM, CHIRPS, GLDAS and PERSIANN-CDR all heavily underestimate the number of heavier rainfall events while CFSv2 and ERA5 considerably overestimate it. Furthermore, from the scatter plots of accumulated rainfall of the satellite precipitation products versus the rain gauge observations, both CHIRPS' and IMERG's values are the most concordant with the in situ measurements.

More generally, the strong underestimation of infrared-only products as is the case of PERSIANN-CDR occurs during the warm orographic rainfall regime. Indeed, it associates these warm clouds mistakenly as non-precipitating (Kimani et al., 2017) because the algorithm relies solely on the cloud-top temperatures (Cattani et al., 2016). Conversely, microwave-derived products can retrieve both convective and orographic rainfall regimes (Kimani et al., 2017). Notwithstanding, they also suffer from the moderate to low signal from ice scattering used in the algorithm, because of its scarcity at the top of the warm orographic clouds (Cattani et al., 2016). Moreover, Kimani et al. (2017) found that the better performance by CHIRPS and TRMM-3B43 over East Africa can be attributed to the direct inclusion of rain gauge data and microwave images during calibration. By extension, the same applies to IMERG which is the improved successor of TRMM and was built on its success (Hou et al., 2014).

In addition, Brunetti et al., (2018) concluded that a high temporal resolution is crucial for forecasting landslides as using the daily aggregated data may result in considerable overestimation depending if the landslide occurred at the beginning of the day. Thus, for future applications, it would be beneficial to have a satellite product with a sub-daily rainfall data such as the IMERG with a frequency of 30 min rather than CHIRPS with its daily frequency. Due to the above statements, we have selected the IMERG precipitation product to study landslide initiation in Rwanda.

When choosing the IMERG product or any other satellite precipitation product for landslide regional assessment, certain considerations should be made:

- The uncertainty of the satellite product over East Africa is controlled by both the precipitation intensity and the topographic complexity (Cattani et al., 2016). The higher the rainfall intensity, the higher the uncertainty. Moreover, precipitation intensity increases with elevation.
- Not only does the rainfall intensity affect the performance of the satellite product but also the rainfall regime. Indeed, Kimani et al., (2017) deduced that the increase in underestimation of satellite products with respect to gauge data in high elevation areas over East Africa can be due to the enhanced stratifications during the deep convections taking place in the MAM season. Furthermore, they also found that all satellite precipitation products have difficulties retrieving the orographic rainfall.
- In their landslide nowcasting at the global scale, Stanley et al., (2021) found that some important false negatives were not predicted due to the absence of heavy rainfall shown by IMERG. This can be due to the lack of a recent overpass by the orbiting passive microwave sensors failing to capture short intense peaks in rainfall. In the case of Rwanda, the rain gauges have a significantly lower temporal frequency (daily) compared to IMERG (30-min).

It should be noted that there are limited studies dealing with product validation and intercomparison of satellite precipitation products over Central East Africa, crucial because of the complex topography, the rainfall geographic variability and the low number of rain gauges. Besides, the studies carried out by Kimani et al., (2017) and Cattani et al., (2016) although important, are done at a monthly scale (and yearly) but not at the higher daily temporal resolution which is especially relevant for the use case of landslides initiation thresholds. Cattani et al., (2016) determined East Africa to be a region where satellite precipitation estimates are still difficult. Addressing the dependence between the satellite products and the elevation to enhance the ability of the algorithms to better represent orographic rainfall regimes will improve the performance of these products which is much needed in these data-scarce regions.

4

Soil Moisture Products

4.1. Introduction

Unlike the satellite precipitation product for which we followed a selection procedure, we obtained the soil moisture time series directly from VanderSat and from the wflow_sbm hydrological model. Hence, the purpose of this chapter is to evaluate the quality of these soil moisture products.

4.2. Methods

We compared the trends rather than the absolute values of the satellite and of the hydrological model derived soil moisture with the AWSs soil moisture time series. In the case of the VanderSat soil moisture data θ_{top}, we plotted the closest ROI to the AWSs in a single graph for five locations. As for the wflow_sbm, we extracted the shallower soil moisture product θ_{root} at the same coordinate as the AWS stations and plotted them together for six locations using a pixel-to-pixel comparison.

Although they are not representative of the same depth, we compared the L-band volumetric soil moisture θ_{top} time series with the wflow_sbm θ_{root} time series over the catchments' ROIs. For each of these locations, we computed the Pearson correlation and resumed the results in a box plot to quantitatively assess the similarity.

We plotted the IMERG precipitation time series along with each soil moisture product (θ_{top}, θ_{root}, and θ_{uz}) averaged over all ROIs time series to investigate the response of the latter to the former. Additionally, we marked the landslide events on the soil moisture time series to analyze whether they occurred for values above or below the mean soil moisture.

4.3. Results

Figure 4.1 and Figure 4.2 show an example of how closely the VanderSat and the wflow_sbm soil moisture match the AWS measured soil moisture time series. Due to its sub-daily frequency, the in situ soil moisture θ_{20cm} exhibits the most fluctuations. As previously mentioned, the AWS data have large gaps. Nevertheless, both satellite and model derived soil moisture time series reproduce the drier periods (e.g. 2018-08 and 2018-

07 respectively) and wetter periods as measured by the AWSs.

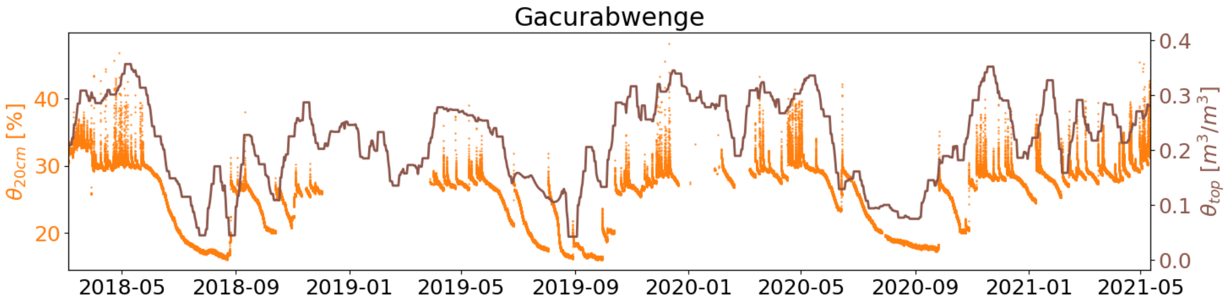

Figure 4.1: Comparison of L-band volumetric soil moisture θ_{top} (brown) with the measured soil moisture θ_{20cm} (orange) time series at the Gacurabwenge AWS.

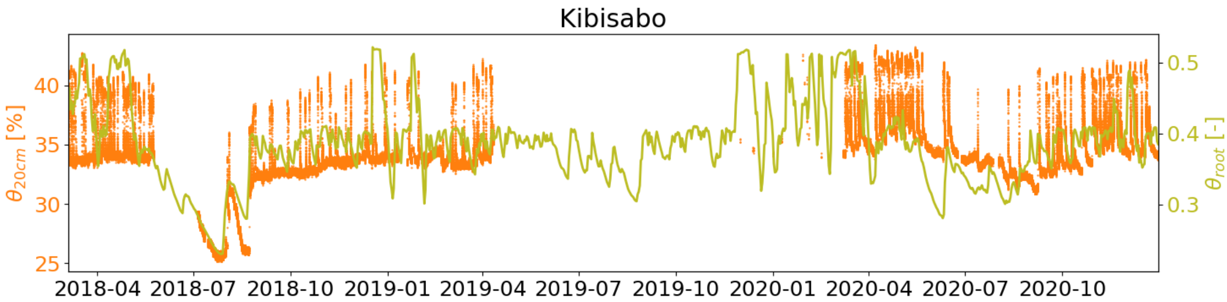

Figure 4.2: Comparison of the wflow modelled volumetric water content in the root zone θ_{root} (lime green) with the measured θ_{20cm} (orange) time series at the Kibisabo AWS.

The satellite soil moisture θ_{top} and the wflow_sbm modeled θ_{root} time series exhibit very similar trends over all ROIs as can be confirmed by the high Pearson correlation values (Figure 4.4) where the mean correlation is above 0.79. An example of their similarity at one particular location is shown in Figure 4.3.

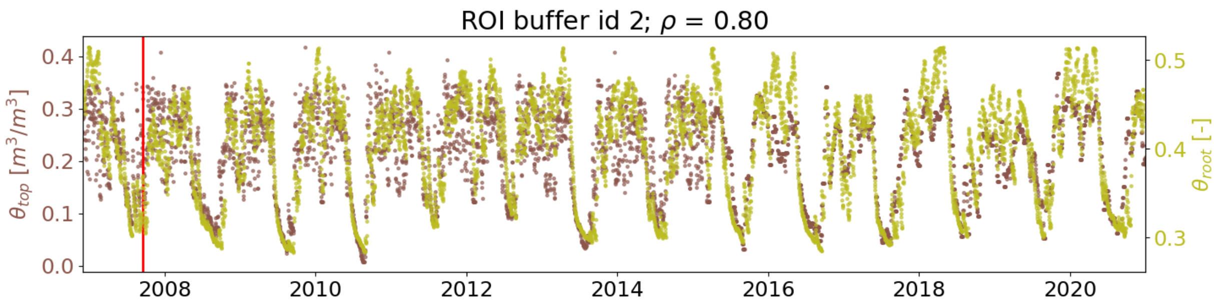

Figure 4.3: Comparison of the θ_{top} (brown) with the θ_{root} (lime green) time series shows a Pearson correlation of 0.80. The red line indicates the timing of the landslide for this location.

Periods of no rain correspond to drops in both satellite (Figure 4.5) and wflow_sbm (Figure 4.6) soil moisture, while periods of consecutive heavy rain coincide with an increase in the soil moisture. The majority of

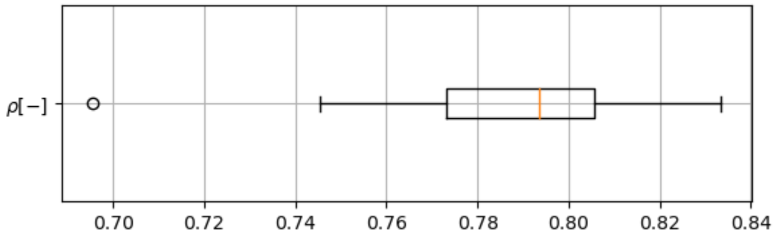

Figure 4.4: Box plot of the Pearson correlation between the VanderSat soil moisture θ_{top} and the wflow_sbm θ_{root} for the ROIs over the catchments.

landslide events occur for above average soil moisture values (Figure 4.5; Figure 4.6). Since the wflow_sbm modeled θ_{root} corresponds to an upper, shallower (approx. 50 cm) soil compartment, it exhibits more fluctuations and responds a bit earlier in time compared to the deeper (approx. 2 m), slower storage wflow_sbm modeled θ_{uz}.

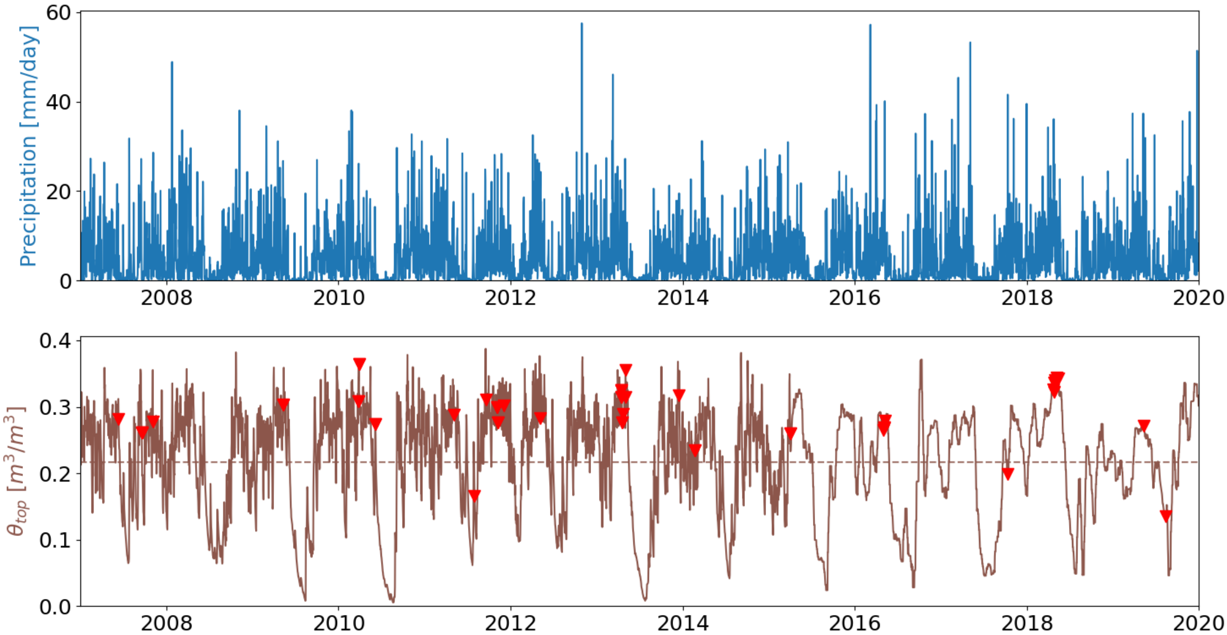

Figure 4.5: IMERG precipitation (top) and VanderSat soil moisture θ_{top} (bottom) averaged over all ROIs. In the bottom graph, the dashed line indicates the all-time soil moisture mean and the red triangles are the landslide events.

4.4. Discussion

Although only a few AWS measurements of the in situ soil moisture for a short period of time are available, both θ_{top} and the θ_{root} are able to reproduce the most important trends regardless of the different depths they represent (approx. 5 cm and 50 cm respectively). Furthermore, comparing the two soil moisture products results in a high Pearson correlation despite originating from different sources meaning that they capture the same hydrological processes. As we found that the IMERG product was the most suitable precipitation product for landslide initiation studies (chapter 3), we analyzed the soil moisture products' response to the IMERG variations. Specifically, whether the soil moisture values become drier in periods of no rain or wet-

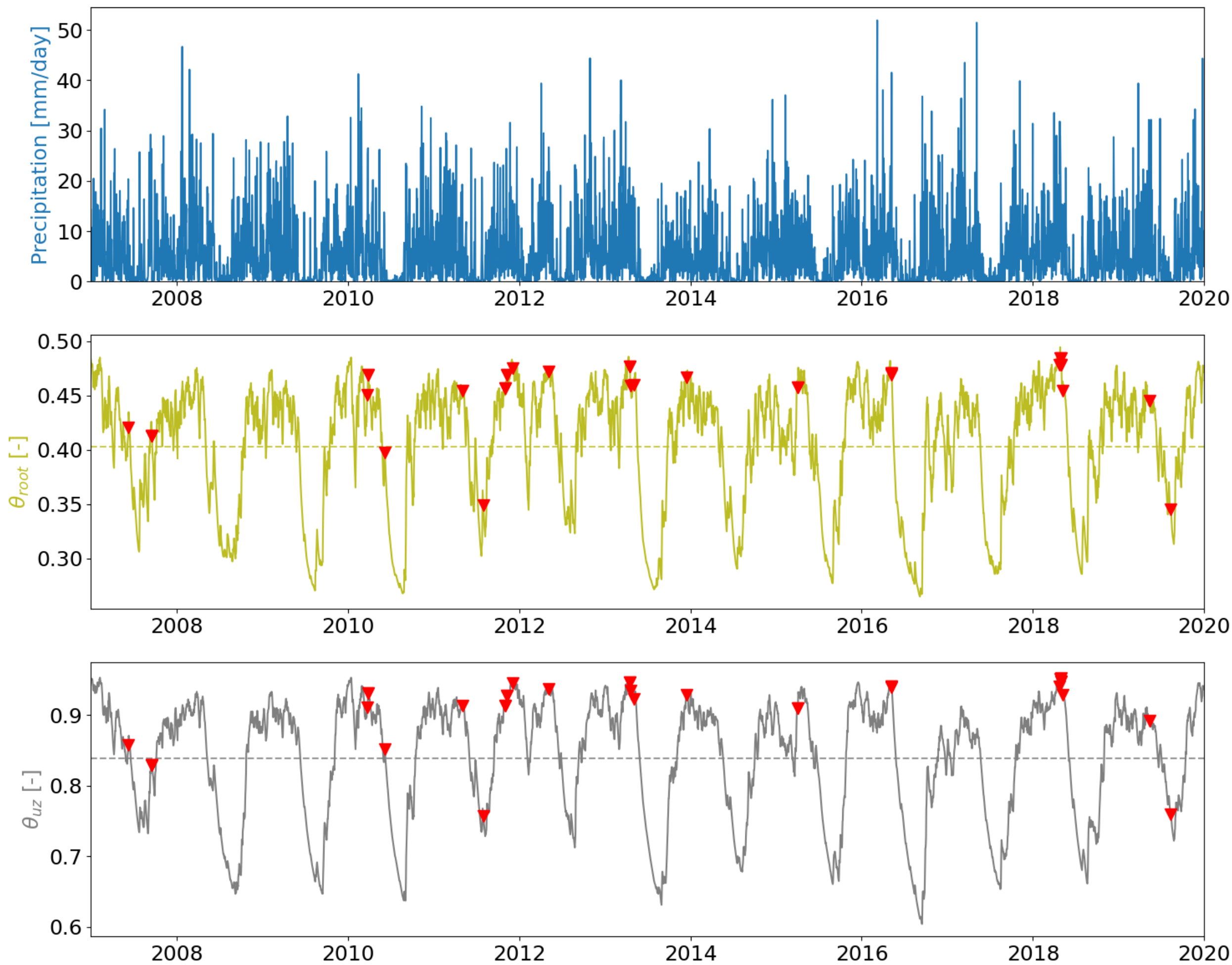

Figure 4.6: IMERG precipitation (top), wflow_sbm modeled θ_{root} (middle) and wflow_sbm modeled θ_{uz} (bottom) averaged over all ROIs inside the catchments. The dashed horizontal lines indicate the all-time mean θ_{root} and θ_{uz} while the red triangles indicate the landslide events.

ter in periods of consecutive heavy rain, which all soil moisture products seem to satisfy. In addition, most landslide events occur with above average soil moisture conditions in agreement with the effects of water accumulation in the subsurface as mentioned in chapter 1. Hence, the satellite-based and the model derived soil moisture are compliant with the plausibility checks and reproduce the most important trends. Therefore the added value of incorporating these soil moisture products to the empirical-statistical rainfall thresholds is studied in chapter 5.

It should be noted that soil moisture retrievals over complex topography or densely vegetated areas exhibit a low quality and should be used cautiously (Brunetti et al., 2018). On the other hand, hydrological simulations are affected by model parameter uncertainty and the selection of the model structure and its parameterization (Bouaziz et al. 2021). In particular, Koch et al., (2016) carry out an inter-comparison of spatio-temporal soil moisture variability of three distributed hydrological models and find that their ability to predict soil moisture is diverging and conclude that the studied catchment poses major challenges to the models in terms of soil moisture heterogeneity and seasonality.

5

Landslide Initiation Thresholds

5.1. Introduction

Most landslide early warning criteria rely exclusively on empirical rainfall thresholds and other indirect proxies for subsurface wetness. The most common rainfall threshold type is the mean rainfall intensity-duration (ID) (Guzzetti et al., 2020; Segoni et al. 2018) in which the separation between landslide occurrence and non-occurrence is based on a power law ($I = \alpha D^{-\beta}$). However, Bogaard and Greco (2018) highlighted several problems associated with this model such as the limited physical meaning and the fact that the mean intensity rather than the peak intensity (which is likely the actual trigger of landslide events) is employed. Hence, in this research, we derived other common rainfall (trigger-based) landslide initiation thresholds. On the other hand, Bogaard and Greco, (2018) proposed the trigger-cause concept for regional landslide initiation thresholds which combines two predominant drivers with distinct timescales: the antecedent hydrological cause and the precipitation trigger. We explored both the traditional precipitation based thresholds, and the proposed trigger-cause thresholds in this chapter. Thus, the aim of this chapter is to answer the research question: which variables or combination thereof have the highest explanatory power for predicting landslides in Rwanda for both the satellite-based variables and for the distributed hydrological model derived variables. As a first step, we derived single-variable landslide initiation thresholds. Secondly, we combined these variables following a bilinear trigger and trigger-cause framework. It should be noted that the choice for the trigger and cause implicitly accounts for the timescale definition and should relate to the characteristics of the landslide (Bogaard and Greco, 2018). The results constitute a preliminary step towards robust landslide early warning systems in Rwanda.

5.2. Methods

5.2.1. Definition of Meteorological Variables

For each landslide location, we downloaded the daily IMERG rainfall time series from 2007 to 2019 via GEE as this product proved to perform the best (chapter 3). For the landslide events that are spatially close to each other such that more than one IMERG buffer cover these events, we selected only one IMERG time series to account for these landslide events. We chose the precipitation time series having the highest weight W (Equation 5.1) according to Melillo et al., (2018) where E is the cumulated rainfall event corresponding to the landslide event, D is the cumulated event duration, and d is the distance between two neighboring IMERG (and thus landslide location) buffers.

$$W = \frac{E^2}{d^2 D} \tag{5.1}$$

To facilitate the comparison between studies, we used the same rainfall definition that yielded the best results in Uwihirwe et al. (2020) which is the maximum probable rainfall event (MPRE). It is defined as the individual periods of days with recorded rain ≥ 1 mm day^{-1} interrupted by dry periods of at least two dry days. The rainfall event E (mm/E) is the accumulated rainfall during the MPRE, the duration D is the number of days each MPRE lasts, and the event/duration (mm/day/E) is the ratio of E over D. We compiled a single dataset of rainfall characteristics from all the rainfall time series analyzed to distinguish between landslide triggering and non-triggering conditions. An advantage of using MPRE instead of daily rainfall is that it copes better with the possible inexact dates of landslides (i.e. landslide could be spotted a few days after the actual triggering day).

5.2.2. Definition of Hydrological Variables

We normalized all soil moisture products (θ_{top}, θ_{root}, and θ_{uz}) to obtain the effective soil moisture Se [-] (Equation 5.2) where θ is the volume of water per unit volume of soil (soil moisture), θ_{res} [-] and θ_{sat} [-] are the residual and saturated moisture contents respectively, corresponding to the minimum and maximum value ever recorded per ROI for the entire time series.

$$Se = \frac{\theta - \theta_{res}}{\theta_{sat} - \theta_{res}} \tag{5.2}$$

We used the effective soil moisture one day before the start of the MPRE, Se_{t-1} as the antecedent hydrological conditions for deriving the landslide initiation thresholds resulting in $Se_{top,t-1}$, $Se_{root,t-1}$ and $Se_{uz,t-1}$. We also tested the $Se_{top,t-1}$ limited to the ERA 5 period (results not shown) to see whether using the moving average decreased the quality of the results. However, by excluding the SMAP period, the already limited landslide inventory further decreases resulting in a reduced AUC of 0.606 compared to the one based on 2007-2019.

5.2.3. Quantification of Landslide Explanatory Variables

The Receiver Operating Characteristic (ROC) curve is a way to visualize the capability of a test variable to distinguish between landslide occurrence and non-occurrence at various threshold settings. Specifically, it is a plot of the true positive rate (TPR) (Equation 5.3) versus the false positive rate (FPR) (Equation 5.4).

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{5.3}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} \tag{5.4}$$

The true positives (TP) are the number of landslides that took place when the threshold was breached. The false negatives (FN) correspond to the number of landslides that occurred although the threshold was not exceeded. The false positives (FP) amount to the number of incorrect predictions of landslides (threshold was breached) by the model while no landslide had been reported. The true negatives (TN) are the number of times the model correctly predicted the non-occurrence of landslides events (threshold was not exceeded). Moreover, Brunetti et al. (2018) stressed that the FP can be overrated by the lack of information on landslide occurrence, i.e. landslides may have occurred but not reported.

The area under the ROC curve (AUC) is a measure of the discriminatory power of the test variable, in other words, of its capability to discern between a hit and a false alarm. An excellent (poor) model has an AUC close to 1 (0) whereas an AUC of 0.5 indicates that the model has no discriminatory power and is referred to as random guessing.

5.2.4. Threshold Definition

Selecting the optimal cutoff value of each test variable (E, D, E/D, $Se_{top,t-1}$, $Se_{root,t-1}$ and $Se_{uz,t-1}$) is a trade-off between maximizing the TPR and minimizing the FPR. To increase the intercomparability of the results from this research (satellite- and hydrological model-based data) with the ones from Uwihirwe et al. (2020) (in situ based data), we used the same threshold formulations. Hence, we computed the optimum cutoff for each test variable employing (i) the maximum true skill statistics (TSS) (Equation 5.5) and (ii) the minimum radial distance (RD) (Equation 5.6). Ideally, the TSS threshold is equal to 1 whereas the RD threshold is equal to 0, this means that for both, their location on the ROC curve is on the top left corner.

$$\mathrm{TSS} = \mathrm{TPR} - \mathrm{FPR} \tag{5.5}$$

$$\mathrm{RD} = \sqrt{\mathrm{FPR}^2 + (\mathrm{TPR} - 1)^2} \tag{5.6}$$

We obtained the resulting optimal cutoff values for the single variable thresholds. In the case of the bilinear thresholds, represented by a 2D plane with one variable on the x- and the other one on the y-axis, we drew the optimal thresholds of the two variables on both axis respectively. Thus, the correctly predicted landslide events are located in the upper right quadrant.

Finally, we quantified the performance of each threshold by the TPR, the FPR and the false negative rate (FNR) (Equation 5.7).

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{TP} + \mathrm{FN}} \tag{5.7}$$

For the IMERG and VanderSat data (satellite-based thresholds), we derived single variable thresholds for the trigger variables E, D, E/D and for the cause variable $Se_{top,t-1}$. Additionally, we constructed two bilinear thresholds: (i) the trigger E-D and (ii) the trigger-cause E/D-$Se_{top,t-1}$. It should be noted that we explored the 3-D threshold E-D-$Se_{top,t-1}$ as well (not shown) but as it does not include new information, the visualization is difficult, and the performance is poor, this threshold was excluded from the research.

Similarly, for the IMERG and wflow_sbm data (hydrological model-based thresholds) we computed single variable thresholds for the trigger variables E, D, E/D and for the cause variables $Se_{root,t-1}$ and $Se_{uz,t-1}$. In addition, we derived three bilinear thresholds: (i) the trigger E-D, (ii) the trigger-cause E/D-$Se_{root,t-1}$, and (iii) the trigger-cause E/D-$Se_{uz,t-1}$.

5.3. Results

5.3.1. Satellite-Based Landslide Initiation Thresholds

Classifying the precipitation into MPRE leads to a total of 13377 rainfall events from 2007 to 2019. From those MPRE, 52 are responsible for the occurrence of one or more landslides.

The Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) results are shown in (Figure 5.1). The AUC is highest for the meteorological variables event E (AUC=0.870), followed by the duration D (AUC=0.860). This means that the ROC curve is closest to the ideal point located on the top left corner with the highest hit alarm rate and lowest false positive rate. Notwithstanding, when normalizing the event by the duration as E/D, the explanatory power reduces significantly (AUC=0.684). Lastly, the hydrological variable $Se_{top,t-1}$ has the least explanatory power (AUC=0.631).

Table 5.1: Optimal cut-off values for each tested variable according to the maximum TSS and the minimum RD (refer to subsection 5.2.4) along with their corresponding TPR, FPR, and FNR. When the optimal threshold established by the TSS and RD differ, the selected threshold for the construction of the bilinear thresholds is written in italic.

	TSS					**RD**				
Variables	**Value**	**Threshold**	**TPR**	**FPR**	**FNR**	**Value**	**Threshold**	**TPR**	**FPR**	**FNR**
E [mm/E]	0.65	102.25	0.83	0.18	0.17	0.25	102.25	0.83	0.18	0.17
D [d]	0.61	11.5	0.79	0.18	0.21	0.28	11.5	0.79	0.18	0.21
E/D [mm/E/d]	*0.38*	*7.00*	*0.87*	*0.47*	*0.13*	0.45	7.94	0.77	0.39	0.23
$Se_{top,t-1}$ [-]	0.25	0.55	0.58	0.32	0.42	0.53	0.55	0.58	0.32	0.42

As quantified by the AUC, the single variable event E has the highest explanatory power (Table 5.1): a high number of correctly predicted landslides (TPR=83%) and a relatively low false alarm rate (FPR=18%) and missed alarms (FNR=17%). Then, the duration D has the same FPR but with a diminished TPR (-4%) and consequently with an increased FNR (+4%). The event/duration threshold (TSS=7 mm/E/day; RD=7.94 mm/E/day) correctly predicts many landslides (TSS=87%; RD=77%) in spite of issuing many false alarms (TSS=47%; RD=39%) making E/D not suitable as a landslide predictor variable. Finally, the optimal $Se_{top,t-1}$ cutoff, 0.55 translates to a low TPR (58%) with a quite high FPR (32%) and FNR (42%) implying a low discrimination capability between the occurrence and non-occurrence of landslides.

The meteorological bilinear threshold E-D has a slightly better performance (Table 5.2) (FPR=-4%) than the

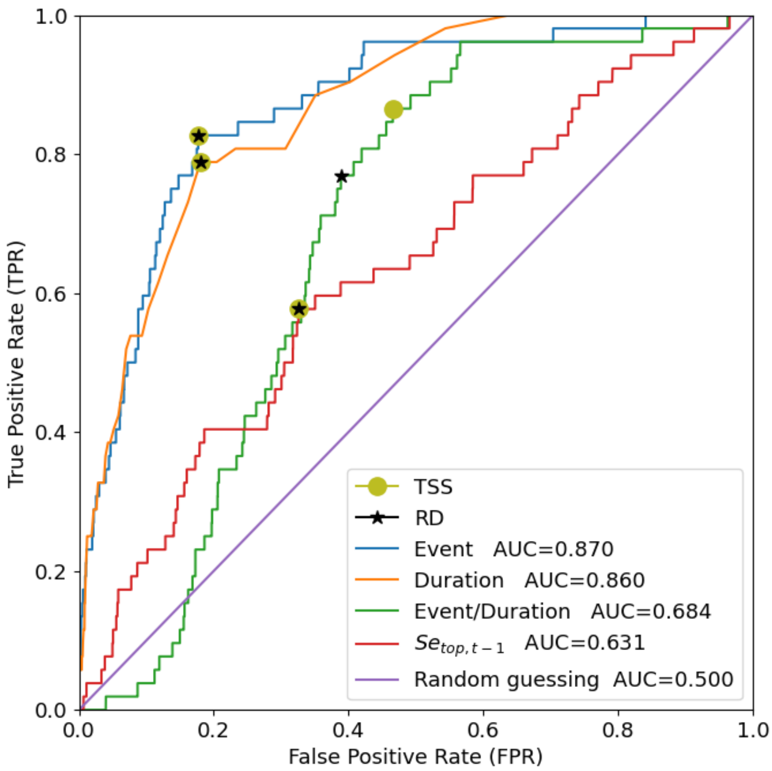

Figure 5.1: ROC curve for the event, duration, event/duration and effective soil saturation one day before the start of the event $Se_{top,t-1}$ variables. The optimal thresholds based on the TSS and the RD are marked on each curve.

Table 5.2: Bilinear thresholds along with their corresponding TPR, FPR, and FNR.

Bilinear Thresholds	**TSS value**	**RD value**	**TPR**	**FPR**	**FNR**
E-D	0.65	0.25	0.79	0.14	0.21
E/D-$Se_{top,t-1}$	0.36	0.51	0.52	0.16	0.48

single variable D threshold. This improvement can be visualized in the bi-logarithmic threshold (Figure 5.2) where the same number of landslides are correctly predicted but with a reduced number of false positives: the lower right quadrant corresponds now to true negatives instead of false positives due to the inclusion of the event threshold. On the other hand, comparing E-D to the single variable E threshold, the FPR is lower (-4%) at the expense of a lower TPR (-4%) and higher FNR (+4%). Indeed, the addition of the duration threshold modifies two correctly predicted landslide events to two missed alarms though it does reduce the false positives that would be otherwise located in the upper left quadrant.

The trigger-cause threshold E/D-$Se_{top,t-1}$ has a very poor overall performance (Table 5.2): a low landslide prediction rate (TPR=52%) and a high missed alarm rate (FNR=48%) despite a low false alarm rate (FPR=16%). From Figure 5.3, it is clear that most landslide events are recorded for an E/D value higher than 5 mm/E/days, slightly lower than the 7 mm/E/days threshold used (horizontal line). However, the landslide events are spread out for $Se_{top,t-1}$ (values range from approx. 0.07 to 0.9), thus this variable is not a good discriminator between the occurrence and non-occurrence of landslide events.

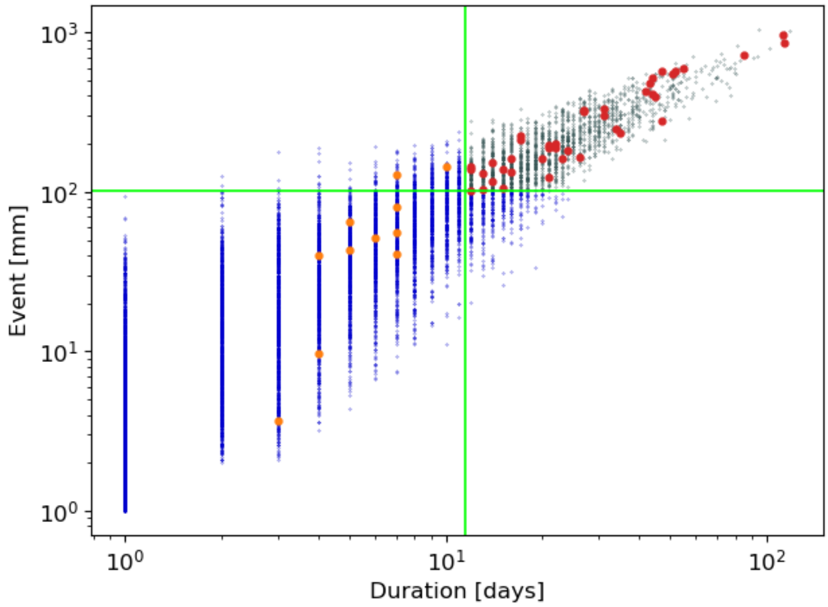

Figure 5.2: Event-duration bilinear threshold. The lime green lines are the optimal cutoff values corresponding to the event and the duration. The bigger dots denote landslide occurrences; the red ones are TP whereas the orange ones are FN. No landslides were reported for the smaller dots: the black ones are FP while the blue ones are TN.

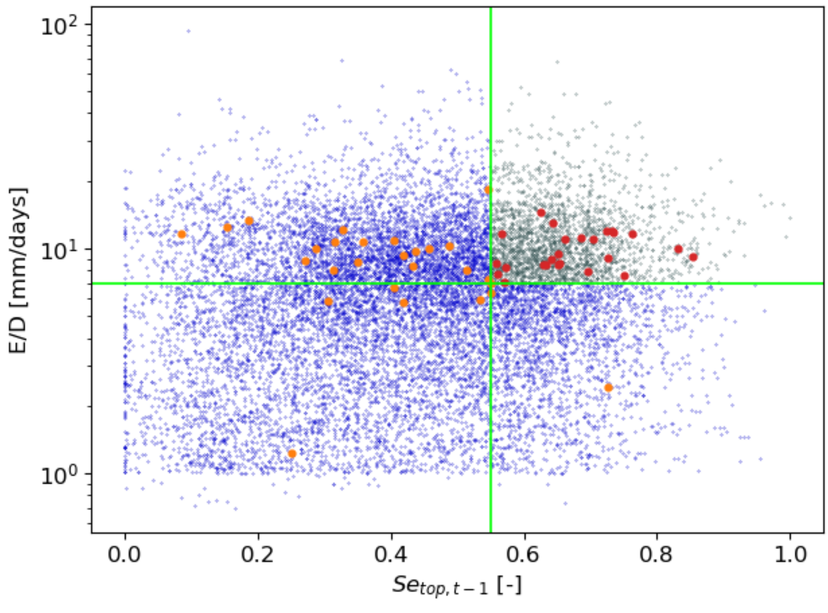

Figure 5.3: Event/duration-$Se_{top,t-1}$ bilinear threshold.

5.3.2. Hydrological Model-Based Landslide Initiation Thresholds

Figure 5.4 shows the Receiver Operating Characteristic (ROC) curve along with the Area Under the Curve (AUC) of each test variable. The AUC is greatest for the duration D (AUC=0.890) and the event E (AUC=0.886) meaning that they are good at discriminating between the occurrence and non-occurrence of landslide events. Similarly to subsection 5.3.1, when normalizing the latter by the former, the predictive power of the E/D variable decreases considerably (AUC=0.662). As before, both hydrological variables have a lower explanatory power (AUC=0.631, AUC=0.642) than the meteorological variables.

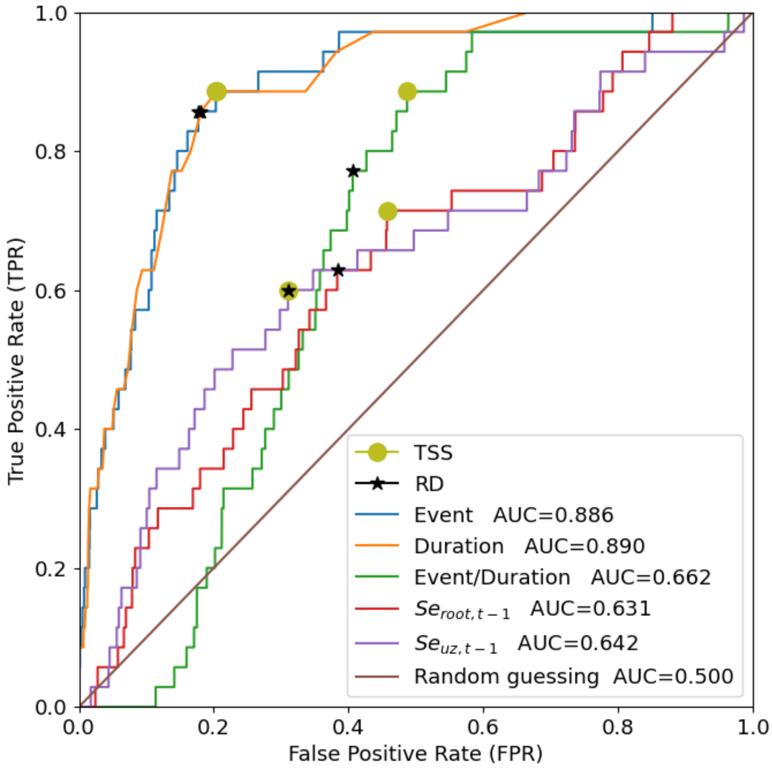

Figure 5.4: ROC curve for the event, duration, event/duration, $Se_{root,t-1}$ and $Se_{uz,t-1}$. The optimal thresholds based on the TSS and the RD are marked on each curve.

Table 5.3: Cut-off values for each tested variable according to the maximum TSS and the minimum RD along with their corresponding TPR, FPR, and FNR. When the optimal threshold established by the TSS and RD differ, the selected threshold for the construction of the bilinear thresholds is written in italic.

	TSS					RD				
Variables	**Value**	**Threshold**	**TPR**	**FPR**	**FNR**	**Value**	**Threshold**	**TPR**	**FPR**	**FNR**
E [mm/E]	*0.68*	*103.17*	*0.89*	*0.20*	0.11	0.23	114.89	0.86	0.18	0.14
D [d]	*0.68*	*11.50*	*0.89*	*0.20*	0.11	0.23	12.50	0.86	0.18	0.14
E/D [mm/E/d]	*0.40*	*6.99*	*0.89*	*0.49*	0.11	0.47	7.94	0.77	0.41	0.23
$Se_{root,t-1}$ [-]	*0.26*	*0.64*	*0.71*	*0.46*	0.29	0.53	0.68	0.63	0.38	0.37
$Se_{uz,t-1}$ [-]	0.29	0.81	0.60	0.31	0.40	0.51	0.81	0.60	0.31	0.40

The test variables E, D and E/D optimal thresholds and predictive capabilities (Table 5.3) are very similar to the previously found ones in subsection 5.3.1, as expected because they are based on the same IMERG product and MPRE definition but over a reduced area. Both E and D show a slight increase (higher TSS value and lower RD value) in explanatory power comparing with the previous results possibly attributed to the smaller and more homogeneous area even if less landslides were used to construct these thresholds. E/D accurately predicts many landslides but at the expense of an elevated false alarm rate and missed alarm rate which is not suitable for landslide forecasting. Both $Se_{root,t-1}$ and $Se_{uz,t-1}$ have a somewhat low TPR (0.6-0.71) with a rather high FPR (0.31-0.46) and FNR (0.29-0.4) suggesting a poor distinction capability between landslide occurrence and non-occurrence.

Table 5.4: Bilinear thresholds along with their corresponding TPR, FPR, and FNR.

Bilinear Thresholds	**TSS value**	**RD value**	**TPR**	**FPR**	**FNR**
E-D	0.73	0.19	0.89	0.16	0.11
E/D - $Se_{root,t-1}$	0.35	0.47	0.60	0.25	0.40
E/D - $Se_{uz,t-1}$	0.34	0.52	0.51	0.17	0.49

The bilinear threshold E-D compared with E or D alone yields the same elevated correctly predicted landslide rate (TPR=0.89) but with a slight reduction of the false alarm rate (FPR -4%) (Table 5.4). This improvement can be visualized in Figure 5.5 where almost all landslide events are clustered in the top right corner except for four missed alarms in the bottom left corner. Hence, the addition of the event E threshold to the duration D or vice versa does not modify any correctly predicted landslides into missed alarms but only decreases the number of false alarms while increasing the true negatives (top left and bottom right quadrant respectively).

On the contrary, the performance of the trigger-cause E/D-$Se_{root,t-1}$ threshold is rather poor (Table 5.4). In Figure 5.6, most landslides are reported for an E/D higher than approx. 6 mm/E/d, slightly below the optimal threshold (6.99 mm/E/d). Hence, including the $Se_{root,t-1}$ threshold modifies the upper left quadrant from many false positives and 10 true positives to many true negatives and 10 missed alarms respectively. This results in a reduction of the FPR by 0.24 at the expense of a decrease in TPR by 0.29 compared to E/D alone. In the case of $Se_{root,t-1}$, the landslide events are spread out (0.1 - 0.9) meaning that the corresponding ideal cutoff value generates many missed alarms implying that this variable is not a good landslide predictor. Then, adding the E/D threshold to the $Se_{root,t-1}$ only transforms two correctly predicted landslides into two missed alarms (bottom right corner) yet it drastically reduces the number of false positives that would otherwise be located on the left side. Consequently, this bilinear threshold has a better performance than the single variable $Se_{root,t-1}$ threshold because it reduces a bit the TPR and significantly the FPR.

We can make analogous observations regarding the trigger-cause E/D-$Se_{uz,t-1}$ threshold (Figure 5.7): (i) compared to E/D alone, the performance decreases, (ii) since the landslide events reported are spread out ranging from approximately 0.2 to 1 for $Se_{uz,t-1}$ values, the ideal cutoff value of this variable misses many landslide occurrences; therefore, it is not a good landslide predictor, and (iii) compared to the single variable $Se_{uz,t-1}$ threshold, the predictive capability remains close.

5.4. Discussion

5.4.1. Satellite-Based Landslide Initiation Thresholds

As previously mentioned, Uwihirwe et al. (2020) defined both trigger and trigger-cause landslide thresholds for Rwanda in an empirical-statistical approach. They used almost the same landslide inventory whereas they employed in situ rain gauge data rather than IMERG data. However, they defined rainfall events identically. Furthermore, they applied the same methodology for defining the optimum thresholds (maximum TSS, minimum RAD) based on the ROC and AUC.

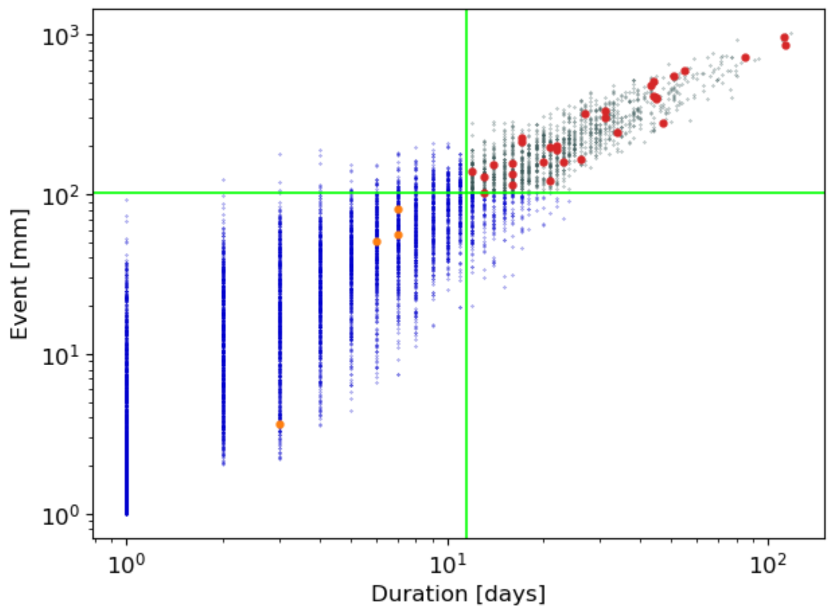

Figure 5.5: Event duration bilinear threshold.

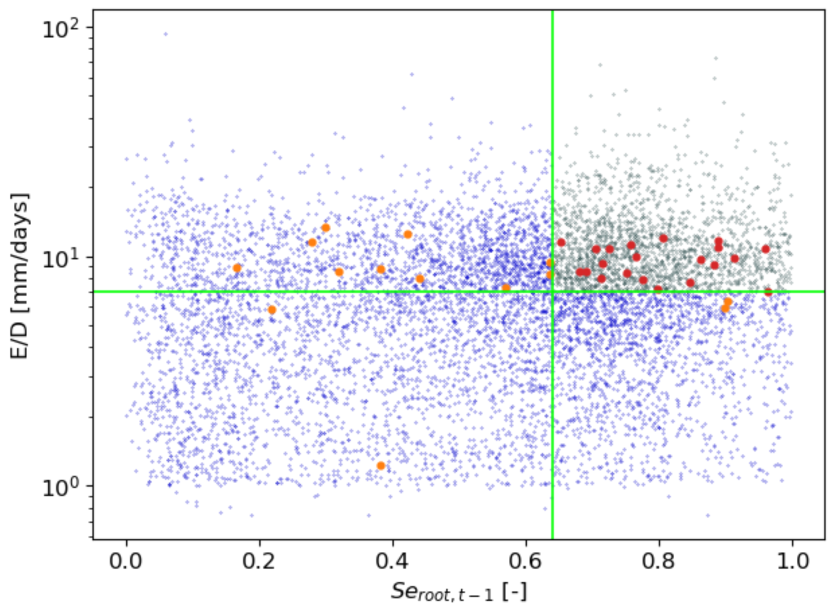

Figure 5.6: Event/duration $Se_{root,t-1}$ bilinear threshold.

The number of MPRE we computed is much higher than the 9353 found between 2006 to 2018 in Uwihirwe et al. (2020). As opposed to using rain gauges, we analyzed an increased number of locations with the IMERG product leading to more precipitation time series.

The optimal threshold established by the TSS and the RD lead to a duration cutoff value of 11.5 days which is more than double than the one calculated using rain gauges (4 days) in the study by Uwihirwe et al., (2020). This is probably due to the lower amount of dry days recorded by IMERG than by the rain gauges (chap-

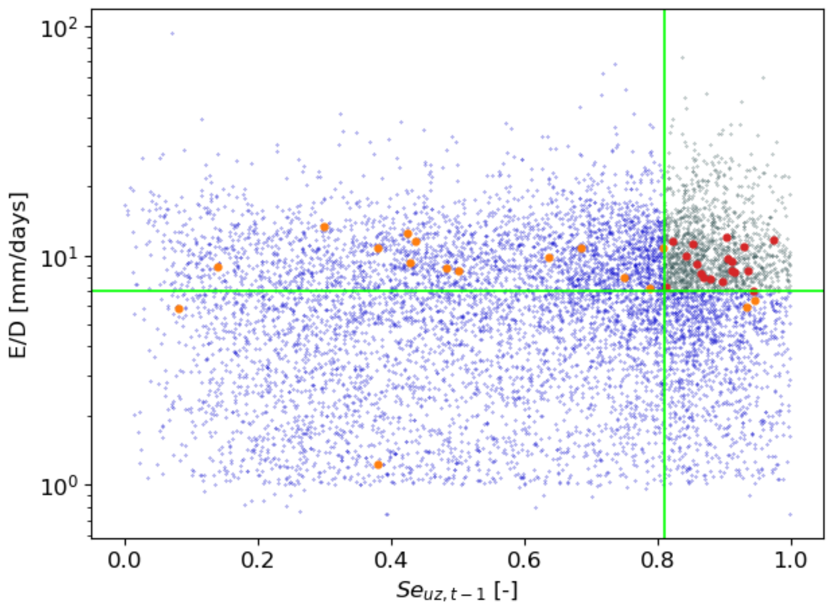

Figure 5.7: Event/duration $Se_{uz,t-1}$ bilinear threshold.

ter 3). We calculated the optimal event E threshold to be 102.25 mm/E, much higher than the ones obtained using rain gauges (TSS=29.9 mm/E, RD=45.90 mm/E). As the duration of the MPRE is longer for IMERG, the total event volume E is much larger because it is the sum of more rainy days as opposed to the in situ precipitation data. Conversely, Brunetti et al. (2018) and Chikalamo et al. (2020) found that the satellite products generally underestimate the cumulated rainfall responsible for the failures measured by the rain gauges. Despite having very different thresholds, the optimal IMERG derived E/D (TSS=7.00 mm/E/day, RD=7.94 mm/E/day) is quite similar to the rain gauge derived intensity I optimal cutoff value (TSS=7.87 mm/day, RD=10.05 mm/day).

The IMERG variables E and D have a high AUC (0.870 and 0.860 respectively), higher than the rain gauges' E and D (0.836 and 0.762 respectively). Hence, when deriving the single variable thresholds, the remotely-sensed E and D correctly predict a high number of landslide events (TPR=0.83; TPR=0.79) while the false alarms remain at low levels (FPR=0.18), significantly lower than the rain gauges'. The performance of E/D is similar, although slightly lower than the in situ derived intensity I. Uwihirwe et al. (2020) defined the landslide causal conditions by the antecedent precipitation index (API) of varying timescales (30, 10 and 5 days) as a proxy for soil moisture accumulation. The explanatory power of $Se_{top,t-1}$ is comparable, although poorer, to that of the APIs'.

5.4.2. Hydrological Model-Based Landslide Initiation Thresholds

Uwihirwe et al. (2021) explicitly accounted for hydrological processes to derive empirical-statistical landslide initiation thresholds for the same three catchments (Kivu, upper Nyabarongo and Mukungwa) in Rwanda by employing observed and modeled groundwater levels one day before h_{t-1} and on the day of the landslide event h_t. In particular, the single variable threshold h_t with an AUC of 0.76-0.80, accurately predicted the occurrence of many landslides ($0.82 \leq$ TPR ≤ 0.93) at the expense of yielding many false alarms ($0.25 \leq$ FPR ≤ 0.38). Contrarily, $Se_{root,t-1}$ and $Se_{uz,t-1}$ have a lower predictive capability (AUC=0.63; AUC=0.64) resulting in a decreased landslide prediction rate (0.63 ≤ TPR ≤ 0.71; TPR=0.60) with an elevated false alarm rate (0.38 ≤ FPR ≤ 0.46; FPR=0.31). When h_t is used in a bilinear trigger-cause threshold (I-h_t), the TPR was drastically reduced (0.64-0.85) along with the FPR (0.08-0.15). Similarly, when $Se_{root,t-1}$ and $Se_{uz,t-1}$ are combined with E/D in a bilinear trigger-cause framework, the correctly predicted landslide rate is reduced (TPR=0.60; TPR=0.51) and the false alarm rate (FPR=0.25; FPR=0.17) as well.

Uwihirwe et al. (2021) indicated that the most frequently recorded landslides in north western Rwanda are deep-seated which are presumably linked to the groundwater level and other hydrogeological factors. As such, the groundwater level on the day of the landslide event appears to have a higher discriminatory power than the antecedent soil moisture in distinguishing landslide from no-landslide conditions over these catchments.

5.4.3. General Considerations

In both subsection 5.3.1 and subsection 5.3.2, we saw that the IMERG derived single variable thresholds E and D are promising due to the high rate of correctly predicted landslides (83% ≤ TPR ≤ 89%), rather low false alarm rate (18% ≤ FPR ≤ 20%), and low missed alarm rate (11% ≤ FNR ≤ 17%). Furthermore, when combined into the bilinear E-D threshold, the correctly predicted landslide events remain high (79% ≤ TPR ≤ 89%) with a decrease of the false alarms (14% ≤ FPR ≤ 16%), and slight increase of the missed alarms (11% ≤ FNR ≤ 21%). Therefore, we advise using the IMERG estimated rainfall as an additional source of rainfall information apart from the rain gauges for future development of LEWSs in Rwanda. In fact, Guzzetti et al. (2020) recommended employing multiple sources of rainfall information and found that the majority of the LEWSs in operation rely on two sources of rainfall information.

The purpose of including the antecedent effective soil saturation variables ($Se_{top,t-1}$, $Se_{root,t-1}$, and $Se_{uz,t-1}$) was to account for the underlying hydrological processes that are responsible for predisposing the slopes to near-failure. Nevertheless, these variables have a poor landslide explanatory power as established by the low AUC (Figure 5.1, Figure 5.4) and by the quantification of the single variable thresholds (Table 5.1, Table 5.3). Notwithstanding, when combining these hydrological variables with the meteorological variable E/D following the trigger-cause concept, we expected the additional causal information to improve the landslide predictability with respect to the trigger only threshold. However, as can be seen in all the bilinear thresholds E/D-Se_{t-1} (Figure 5.3, Figure 5.6, and Figure 5.7), the recorded landslide events are spread out rather than clustered for all of the hydrological variables. Hence, the corresponding optimal cutoffs accurately predict few landslide events (0.51 ≤ TPR ≤ 0.60) and miss many (0.40 ≤ FNR ≤ 0.49). Mirus et al. (2018b) pointed out that if the landslide events do not plot in a cluster in the corner of the 2D selected threshold space, it is possible that other threshold formats may be more suitable.

The combination of the test variables into a trigger-cause framework implies the definition of the timescale separating the trigger from the cause (Bogaard and Greco, 2018). The antecedent soil moisture, taken as one day before the start of the MPRE, is likely to be too far ahead in time with respect to the occurrence of landslides besides having varying timescales. Since the meteorological variables defined here are good landslide predictors, it is possible that the trigger is more important than the cause for landslide initiation in Rwanda. Alternatively, because of their long timescale, the trigger variables may already be accounting for the antecedent soil moisture conditions and when combined into the trigger-cause framework, the Se_{t-1} variables become more noise than added value.

One recommendation that arises is to explore other rainfall definitions that have a shorter and fixed timescale such as the cumulated precipitation in the two or three days before the landslide activity. Hence, the associated soil moisture causal variable may increase the predictability of landslide events in Rwanda as it was observed in chapter 4 that most landslide events occurred for above average soil moisture conditions. Additionally, other hydrological variables such as the wflow_sbm modeled groundwater table could be tested.

As already stated, the landslide inventory is small and incomplete, possibly containing uncertainties. The landslide initiation model uncertainty has multiple sources, including the number, distribution, and accuracy of the empirical data points, and the definition of the rainfall conditions that initiate the landslides.

Lastly, our aim was to derive the most informative landslide initiation thresholds for Rwanda using satellite and model derived data. Nonetheless, we propose certain modifications that do not require expenditure from the government of Rwanda for the future implementation of LEWSs. Operational LEWSs use a variety of rainfall information including rainfall data from rain gauge networks, forecasts from numerical weather models, nowcasts from weather radars, and satellite-based rainfall estimates (Guzzetti et al., 2020). Since the IMERG product performed well for triggering conditions in Rwanda but has a latency of 2-3 days, the early IMERG product with a latency of 4 hours can be instead used for landslide nowcasting (Stanley et al., 2021). Additionally, weather forecasts can be employed for deriving the meteorological trigger variables and also as input data rather than the historical ERA5 data for running the wflow_sbm model to model hydrological variables.

6

Conclusion

Rwanda is both a landslide-prone and data-scarce region. In this research, we explored other sources of data not limited to the in situ data for predicting landslides. As rainfall is the main triggering mechanism behind landslides, we firstly made an assessment of freely available satellite precipitation products accessed through GEE over Rwanda using rain gauges as a reference. Other than finding which precipitation product is the most suitable, chapter 3 constitutes a contribution to the few studies dealing with product validation and intercomparison of satellite products over Rwanda. In the evaluation, IMERG proved to be a reliable source of precipitation data for the determination of rainfall thresholds in Rwanda. Secondly, we assessed two sources of soil moisture data: the satellite soil moisture data from VanderSat and the wflow_sbm hydrological model derived soil moisture. Unlike other remote sensing products that give good results in spite of their coarse spatial resolution (and thus, limiting their application), the L-band volumetric soil moisture from VanderSat has an almost daily resolution with a very high spatial resolution. On the other hand, the wflow_sbm model can be applied to data-scarce areas such as Rwanda, only requiring ERA5 precipitation, temperature, radiation and pressure and computing the potential evaporation as forcings to obtain daily and high spatial resolution data. As the VanderSat and wflow_sbm data complied with the plausibility checks and exhibited reasonable trends (chapter 4), we used them both for deriving landslide initiation thresholds for Rwanda.

Following an empirical-statistical approach, we studied the added value of incorporating antecedent soil moisture to the landslide initiation thresholds in Rwanda (chapter 5). We defined both trigger and trigger-cause based thresholds and we quantified objectively their performance employing the same metrics as in Uwihirwe et al. (2020, 2021). The results indicate that the meteorological variables event E, duration D and the corresponding E-D bilinear threshold hold the highest predictive power to discriminate between landslide occurrence and non-occurrence. Even if the IMERG derived meteorological thresholds have a higher predictive capability than their in situ counterpart (Uwihirwe et al., 2020), rain gauge precipitation should not be overlooked. Instead, following the recommendation of Guzzetti et al., (2020), both the IMERG satellite data and the rain gauge data should be used for predicting landslides in Rwanda despite yielding different thresholds. As also found in Uwihirwe et al. (2020, 2021), single variable thresholds tend to have a higher TPR in spite of an elevated FPR. Conversely, the bilinear thresholds lead to a decrease of the FPR at the expense of a reduction of TPR. We did not witness any improvement by including the antecedent soil moisture variables ($Se_{top,t-1}$, $Se_{root,t-1}$, $Se_{uz,t-1}$) following the trigger-cause concept with respect to the trigger only thresholds.

Defining the hydrological variables closer or on the day of the landslide event could lead to an increased prediction as we observed that most landslides occurred during above average soil moisture conditions. Besides, the trigger-cause threshold using the groundwater level on the day of the landslide event as a causal variable in Uwihirwe et al., (2021) showed encouraging results. In our research, the timescale of the triggering events is not constant and probably too long making the contribution of the antecedent soil moisture not significant. Hence, a plausible solution could be the reduction of the triggering variables' timescale to the cumulated two or three days combined with the corresponding antecedent soil moisture Se_{t-2}, Se_{t-3}.

A

Appendices

Table A.1: Selected rain gauges.

Station	Longitude	Latitude	First date of recorded rainfall	Last date of recorded rainfall	Count of missing values within first and last dates
Gitega	30.06000	-1.950	2006-01-01	2018-12-31	0
Kigali Aero	30.13278	-1.965	2006-01-01	2017-12-30	0
Kinazi	29.91000	-2.200	2006-01-01	2017-12-30	0
Kibangu	29.68000	-1.830	2006-01-01	2018-12-31	32
Rugabano	29.48000	-2.060	2006-01-01	2017-12-30	0
Gisenyi Aero	29.25000	-1.660	2006-01-01	2018-12-31	1
Kanama	29.35000	-1.700	2006-01-01	2017-12-30	0
Pfunda	29.28000	-1.680	2006-01-01	2017-12-30	0
Rwankeri-Nyabihu	29.51000	-1.580	2006-01-01	2018-12-31	1
Kabaya	29.50000	-1.760	2006-01-01	2017-12-30	0
Cyato	29.20000	-2.410	2006-01-01	2017-12-30	0
Rwankuba	29.85000	-1.750	2006-01-01	2017-12-30	0
Ruhengeri Aero	29.61000	-1.480	2006-01-01	2018-12-31	1
Butaro	29.83000	-1.410	2006-01-01	2017-12-30	0
Kabeza-Nyam	30.05000	-1.430	2006-01-01	2017-12-30	0
Rugendabari	29.66000	-1.950	2018-01-02	2018-12-31	31
Muramba Paroisse	29.60000	-1.750	2018-01-02	2018-12-31	0
Shangi	29.00000	-2.380	2018-01-02	2018-12-31	0
Cyinzuzi	30.00000	-1.760	2018-01-02	2018-12-31	0

Overview of pre-selected GEE precipitation products:

- The Tropical Rainfall Measuring Mission (TRMM) is a joint international program of the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA). The algorithm used is the TRMM Multisatellite Precipitation Analysis (TMPA) 3B42 version 7, **TRMM 3B427**. The product analyzed here is the 3-hourly combined microwave-IR estimates with gauge adjustment at a 0.25° x 0.25° spatial resolution covering 50°N-50°S. TMPA relies on data from the precipitation radar (PR), passive microwave (PMW) from a variety of low Earth orbit satellites, infrared (IR) data provided by the international constellation of geosynchronous-orbit meteorological satellites and the precipitation gauge supplied by the Global Precipitation Climatology Centre (GPCC).

- The **Climate Hazards group InfraRed Precipitation with Station data (CHIRPS)** was developed to assist the United States Agency for International Development Famine Early Warning Systems Network. It provides estimates at a daily resolution of resolution 0.05° x 0.05° extending from 50°N-50°S. The inputs required by CHIRPS are: (1) the monthly precipitation climatology, CHPClim; (2) the geostationary thermal infrared (IR) satellite observations; (3) the satellite estimates from TMPA 3B42; (4) the atmospheric model rainfall fields from the NOAA Climate Forecast System, version 2 (CFSv2); and, (5) the in-situ precipitation observations.

- The **Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Records (PERSIANN-CDR)** is created for hydrological and climate studies. This product covers 60°N-60°S, is daily and has a resolution of 0.25° x 0.25°. PERSIANN-CDR is generated from the PERSIANN algorithm. The algorithm is applied to GridSat-B1 infrared data and then, it is bias-adjusted using the Global Precipitation Climatology Project (GPCP) monthly product and accumulated to the daily scale resulting into the final product.

- The **Global Land Data Assimilation System (GLDAS)** has been jointly developed by the NASA and the National Oceanic and Atmospheric Administration (NOAA). GLDAS is an uncoupled land surface modeling system. The precipitation data are obtained from the U.S. Naval Research Laboratory (NRL) and from the NASA Goddard Space Flight Center (GSFC). The NRL is based on geostationary satellite infrared (IR) cloud-top temperature measurements and microwave observation techniques. The microwave product merges data from the Special Sensor Microwave/Imager (SSM/I), the TRMM, and the Advanced Microwave Sounding Unit (AMSU) instruments. Both NRL products have a spatial resolution of 0.25° x 0.25° and a temporal resolution of 6h and cover the area from 60°N-60°S. The optimal merging of the microwave and IR data is done at GSFC.

- The **second version of Climate Forecast System (CFSv2)** is developed by the NOAA National Centers for Environmental Information (NCEP) and entered into operations in 2011. It is a high-resolution coupled atmosphere-ocean-land surface-sea ice system. The model-generated precipitation which is generally too biased is replaced by observed precipitation resulting therefore in semi-coupled for the land section. Two global precipitation analyses are used: (1) the pentad dataset of CMAP, and (2) the CPC unified global daily gauge analysis. The former is defined as the 5-day mean precipitation on a 0.25° x 0.25° grid from gauge observations and satellite observations in the infrared and microwave channels. The latter is based on a 0.5° x 0.5° grid over the global land via interpolation of quality-controlled rain gauges. Specifically, the Optimal Interpolation (OI) algorithm is employed to partially account for the orographic enhancements in precipitation. These two precipitation inputs are blended with the CFSR background 6-hourly Global Data Assimilation System (GDAS) precipitation. The blending function is latitude dependent favoring thus the CMAP product in the tropics, the CPC analysis in the mid-latitudes and the model precipitation in the high latitudes.

- The Global Precipitation Measurement (GPM) launched in 2014 is the improved successor of the TRMM. The new product is the **Integrated Multi-satellitE Retrievals for GPM (IMERG)** and will replace the TMPA. IMERG provides half-hourly data at a 0.1° x 0.1° spatial resolution covering 60°N-60°S. Similarly to the TMPA, IMERG is based on PMW from various low Earth orbit satellites, IR from geosynchronous Earth orbit satellites, and precipitation gauge from two GPCC products. The major improvements are: (1) the increase in orbital inclination, covering therefore additional climate zones; (2) the upgrade of the radar increasing the sensitivity to light precipitation; and, (3) the inclusion of high frequency channels to the PMW imager aimed at sensing light and solid precipitation.

- Within the Copernicus Climate Change Service (C3S), the European Centre for Medium-Range Weather Forecasts (ECMWF) produced the **ERA5 reanalysis**, the fifth generation of atmospheric reanalysis replacing its predecessor ERA-Interim analysis. The precipitation is generated employing a convection scheme representing convection at spatial scales smaller than the grid box. This scheme along with the large-scale cloud scheme have been upgraded with an improved representation of mixed-phase clouds, and prognostic variables for precipitating rain and snow. Additionally, there were improvements to the parametrization of the microphysics, in particular, for warm-rain processes. The convection parametrizations has been changed resulting to an improvement in the distribution of the rain rate and the representation of tropical variability. Upgrades have been achieved in the diurnal cycle of convection.

Table A.2: 5-km radius buffer for the landslide events.

Landslide overlapping events	**Distance between events [m]**	**Solution**
39 and 11	3092	Dissolved the two buffers
10 and 35	0	Deleted buffer of 35
8 and 28	50	Deleted buffer of 28
12 and 38	370	Deleted buffer of 38 because 12 does not intersect with 21
14 and 18	69	Deleted buffer of 14 because 18 is a bit further away to 12 and 38
55 and 12	7391	Applied union to buffers, created line through intersection, split with lines tool, merged and dissolved the separated intersection to each of the vector layers
12 and 18	6973	Applied union to buffers, created line through intersection, split with lines tool, merged and dissolved the separated intersection to each of the vector layers
27 and 30	46	Deleted buffer of 27 because 30 is a bit further away to 32
30 and 32	3230	Created a single buffer of size 0.05ř halfway
47, 36, 34	7967, 8201 and 9048	Applied union to buffers, the intersections are merged and dissolved to the vector layers evenly
34 and 24	8777	Applied union to buffers, the intersection is merged and dissolved to 34
9 and 7	1592	Dissolved the two buffers
20 and 22	2035	Created single buffer halfway
1 and 50	2668	Created single buffer halfway
20/22 and 1/50	7328	Applied union to buffers, created line through intersection, split with lines tool, merged and dissolved the separated intersection to each of the vector layers
23 and 44	50	Deleted buffer of 23
10 and 35	0	Deleted buffer of 35
51, 10, 15 and 44	Distance to 15: 5564, 4783, 5741	Reduced buffers to 0.04ř, applied union to buffers, created line through intersection, split with lines tool, merged and dissolved the separated intersection to each of the vector layers
49 and 10	8907	Applied union to buffers, the intersection is merged and dissolved to 10
43 and 26	25	Deleted buffer of 43
53		Clipped buffer to Rwanda
42		Clipped buffer to Rwanda
41 and 4	1977	Dissolved the two buffers and clipped them to Rwanda
16, 29 and 33	3912 and 4056	Reduced buffers to 0.04ř, applied union to buffers, created line through intersection, split with lines tool, merged and dissolved the separated intersection to each of the vector layers
52 and 54	787	Deleted buffer of 54 to avoid the intersection with 46
52 and 31	7647	Applied union to buffers, created line through intersection, split with lines tool, merged and dissolved the separated intersection to each of the vector layers
3 and 25	1382	Created single buffer of size 0.04ř halfway
13, 3/25, 6	5397 and 5909	Reduce buffers to 0.04ř, applied union to buffers, created line through intersection, split with lines tool, merged and dissolved the separated intersection to each of the vector layers
6 and 45	7873	Applied union to buffers, the intersection is merged and dissolved to 6

Table A.3: Selected automatic weather stations.

Station	Longitude	Latitude	First date of recorded soil humidity	Last date of recorded soil humidity
Kibisabo	29.51	-1.7	2018-03-05	2021-05-18
Kinigi	29.59	-1.44	2018-11-14	2021-01-10
Muko	30.23	-1.7	2018-03-20	2021-05-02
Gacurabwenge	29.91	-1.99	2018-03-05	2021-05-12
Gisenyi	29.46	-1.82	2018-11-15	2020-10-19
Byimana	29.74	-2.14	2018-03-05	2019-11-01
Macuba	29.22	-2.31	2018-03-05	2021-05-18
Rubona	29.77	-2.49	2018-11-24	2021-05-18

Table A.4: Events that are represented by a single IMERG precipitation and VanderSat soil moisture time series. The IMERG time series are available for all the events, hence the criteria used to select only one event is explained (subsection 5.2.1). The VanderSat soil moisture corresponds to the only available one to represent these events.

Events	IMERG time series	VdS time series
18, 14	18 has a higher weight.	14
20, 22	20 has a higher weight.	20
44, 23	44 has a higher weight.	23
41, 4	41 has a higher weight.	4
52, 54	52 has a higher weight.	52
3, 25	3 so that it doesn't overlap with 13.	3
9, 7	9 has a higher weight.	7
50, 1	50 has a higher weight.	1
43, 26	43 has a higher weight.	26
27, 32, 30	27/30 because 32 is farther away. Between 27 and 30, 27 is kept because it has a higher weight.	27
38, 12, 55	55 has no available soil moisture values so pair it to 12/38 closest pair. Between 12 and 38, 38 is kept because it has a higher weight.	12
10, 35, 15	15 overlaps with 10/35 (IMERG), 10/35 farthest from other landslide points. Between10 and 35, they are both in the same location (=same time series), so 10 is chosen.	10
29, 16, 33	29 IMERG covers 16 and 33, otherwise 16 and 33 would both cover 29.	29
28, 8, 19, 39, 11	39, 11 and 19 have no available soil moisture values, 8 has and 28 is in the same location as 8. Between 8 or 28, 28 is kept because it has a higher weight.	8

Bibliography

Abancó, C., Bennett, G. L., Matthews, A. J., Matera, M. A. M., and Tan, F. J. (2021). The role of geomorphology, rainfall and soil moisture in the occurrence of landslides triggered by 2018 typhoon mangkhut in the philippines. *Natural Hazards and Earth System Sciences*, 21(5):1531–1550.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P. (2015). Persiann-cdr: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96(1):69 – 83.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, 21(12):6201–6217.

Bezak, N., Jemec Aufli, M., and Matja, M. (2019). Application of hydrological modelling for temporal prediction of rainfall-induced shallow landslides. *Landslides*, 16:1273–1283.

Bogaard, T. and Greco, R. (2018). Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds. *Natural Hazards and Earth System Sciences*, 18:31–39.

Bogaard, T. A. and Greco, R. (2016). Landslide hydrology: from hydrology to pore pressure. *WIREs Water*, 3(3):439–459.

Bouaziz, L. J. E., Fenicia, F., Thirel, G., de Boer-Euser, T., Buitink, J., Brauer, C. C., De Niel, J., Dewals, B. J., Drogue, G., Grelier, B., Melsen, L. A., Moustakas, S., Nossent, J., Pereira, F., Sprokkereef, E., Stam, J., Weerts, A. H., Willems, P., Savenije, H. H. G., and Hrachowitz, M. (2021). Behind the scenes of streamflow model performance. *Hydrology and Earth System Sciences*, 25(2):1069–1095.

Brunetti, M., Melillo, M., Peruccacci, S., Ciabatta, L., and Brocca, L. (2018). How far are we from the use of satellite rainfall products in landslide forecasting? *Remote Sensing of Environment*, 210:65–75.

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N., Linlin, L., and Tarko, A. (2020). Copernicus global land service: Land cover 100m: Version 3 globe 2015-2019.

Cattani, E., Merino, A., and Levizzani, V. (2016). Evaluation of monthly satellite-derived precipitation products over east africa. *Journal of Hydrometeorology*, 17(10):2555 – 2573.

Chikalamo, E. E., Mavrouli, O. C., Ettema, J., van Westen, C. J., Muntohar, A. S., and Mustofa, A. (2020). Satellite-derived rainfall thresholds for landslide early warning in bogowonto catchment, central java, indonesia. *International Journal of Applied Earth Observation and Geoinformation*, 89:102093.

Chleborad, A. F., Baum, R. L., Godt, J. W., and Powers, P. S. (2008). A prototype system for forecasting landslides in the Seattle, Washington, area. In *Landslides and Engineering Geology of the Seattle, Washington, Area*. Geological Society of America.

de Bruin, H. (1983). A model for the priestley-taylor parameter. *Journal of Applied Meteorology - J APPL METEOROL*, 22:572–578.

de Jeu, R. A., Holmes, T. R., Parinussa, R. M., and Owe, M. (2014). A spatially coherent global soil moisture product with improved temporal resolution. *Journal of Hydrology*, 516:284–296. Determination of soil moisture: Measurements and theoretical approaches.

Demarée, G. and Van deVyver, H. (2013). Construction of intensity-duration-frequency (idf) curves for precipitation with annual maxima data in rwanda, central africa. *Advances in Geosciences*, 35:1–5.

Funk, C., Peterson, P., Peterson, P., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J. (2015). The climate hazards infrared precipitation with stationsa new environmental record for monitoring extremes. *Scientific Data*, 2.

Gash, J. H. C. (1979). An analytical model of rainfall interception by forests. *Quarterly Journal of the Royal Meteorological Society*, 105(443):43–55.

Guzzetti, F. (2021). Invited perspectives: Landslide populations – can they be predicted? *Natural Hazards and Earth System Sciences*, 21(5):1467–1471.

Guzzetti, F., Gariano, S. L., Peruccacci, S., Brunetti, M. T., Marchesini, I., Rossi, M., and Melillo, M. (2020). Geographical landslide early warning systems. *Earth-Science Reviews*, 200:102973.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., and Thépaut, J.-N. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, 95(5):701 – 722.

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., and Xie, P. (2018). Nasa global precipitation measurement (gpm) integrated multi-satellite retrievals for gpm (imerg).

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F. (2007). The trmm multisatellite precipitation analysis (tmpa): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1):38 – 55.

Huffman, J. and Bolvin, D. (2018). Trmm and other data precipitation data set documentation.

Jarvis, A., Reuter, H., Nelson, A., and Guevara, E. (2008). Hole-filled seamless srtm data v4, international centre for tropical agriculture (ciat). `http://srtm.csi.cgiar.org`.

Kimani, M., Hoedjes, J., and Su, Z. (2017). An assessment of satellite-derived rainfall products relative to ground observations over east africa. *Remote Sensing*, 9.

Koch, J., Cornelissen, T., Fang, Z., Bogena, H., Diekkrüger, B., Kollet, S., and Stisen, S. (2016). Inter-comparison of three distributed hydrological models with respect to seasonal variability of soil moisture patterns at a small forested catchment. *Journal of Hydrology*, 533:234–249.

le Coz, C. (2021). *Setting Africa's rainfall straight: A warping approach to position and timing errors in rainfall estimates*. PhD thesis, Delft University of Technology.

Marino, P., Peres, D., Cancelliere, A., Greco, R., and Bogaard, T. (2020). Soil moisture information can improve shallow landslide forecasting using the hydrometeorological threshold approach. *Landslides*, 17(9):2041–2054.

Marra, F., Destro, E., Nikolopoulos, E. I., Zoccatelli, D., Creutin, J. D., Guzzetti, F., and Borga, M. (2017). Impact of rainfall spatial aggregation on the identification of debris flow occurrence thresholds. *Hydrology and Earth System Sciences*, 21(9):4525–4532.

Melillo, M., Brunetti, M. T., Peruccacci, S., Gariano, S. L., Roccati, A., and Guzzetti, F. (2018). A tool for the automatic calculation of rainfall thresholds for landslide occurrence. *Environmental Modelling Software*, 105:230–243.

Mirus, B., Becker, R., Baum, R., and Smith, J. (2018a). Integrating real-time subsurface hydrologic monitoring with empirical rainfall thresholds to improve landslide early warning. *Landslides*, 15.

Mirus, B., Morphew, M., and Smith, J. (2018b). Developing hydro-meteorological thresholds for shallow landslide initiation and early warning. *Water*, 10:1274.

Monsieurs, E., Kirschbaum, D. B., Tan, J., Mateso, J.-C. M., Jacobs, L., Plisnier, P.-D., Thiery, W., Umutoni, A., Musoni, D., Bibentyo, T. M., Ganza, G. B., Mawe, G. I., Bagalwa, L., Kankurize, C., Michellier, C., Stanley, T., Kervyn, F., Kervyn, M., Demoulin, A., and Dewitte, O. (2018). Evaluating tmpa rainfall over the sparsely gauged east african rift. *Journal of Hydrometeorology*, 19(9):1507 – 1528.

Nicholson, S. (2017). Climate and climatic variability of rainfall over eastern africa: Climate over eastern africa. *Reviews of Geophysics*, 55.

Nikolopoulos, E., Borga, M., Creutin, J., and Marra, F. (2015). Estimation of debris flow triggering rainfall: Influence of rain gauge density and interpolation methods. *Geomorphology*, 243:40–50.

Owe, M., de Jeu, R., and Holmes, T. (2008). Multisensor historical climatology of satellite-derived global land surface moisture. *Journal of Geophysical Research: Earth Surface*, 113(F1).

Owe, M., de Jeu, R., and Walker, J. (2001). A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. *IEEE Transactions on Geoscience and Remote Sensing*, 39(8):1643–1654.

Peel, M. C., Finlayson, B. L., and McMahon, T. A. (2007). Updated world map of the köppen-geiger climate classification. *Hydrology and Earth System Sciences*, 11(5):1633–1644.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, 85(3):381 – 394.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., ya Chuang, H., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P. V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M. (2010). The ncep climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8):1015 – 1058.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., ya Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E. (2014). The ncep climate forecast system version 2. *Journal of Climate*, 27(6):2185 – 2208.

Schellekens, J., van Verseveld, W., Visser, M., Winsemius, H., Euser, T., Bouaziz, L., Christophe Thiange, S. d. V., Boisgontier, H., Eilander, D., Tollenaar, D., Weerts, A., Baart, F., Hazenberg, P., Lutz, A., ten Velden, C., Jansen, M., and Benedict, I. (2021). openstreams/wflow: unstable-master. `https://github.com/openstreams/wflow`.

Segoni, S., Rosi, A., Lagomarsino, D., Fanti, R., and Casagli, N. (2018). Brief communication: Using averaged soil moisture estimates to improve the performances of a regional-scale landslide early warning system. *Natural Hazards and Earth System Sciences*, 18(3):807–812.

Sidle, R. C., Greco, R., and Bogaard, T. (2019). Overview of landslide hydrology. *Water*, 11(1).

Skakun, S., Kussul, N., Shelestov, A., and Kussul, O. (2016). The use of satellite data for agriculture drought risk quantification in ukraine. *Geomatics, Natural Hazards and Risk*, 7:901–917.

Stanley, T. A., Kirschbaum, D. B., Benz, G., Emberson, R. A., Amatya, P. M., Medwedeff, W., and Clark, M. K. (2021). Data-driven landslide nowcasting at the global scale. *Frontiers in Earth Science*, 9:378.

Tank, A. M. K., Zwiers, F. W., and Zhang, X. (2009). *Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation*, volume WCDMP-No. 72. World Meteorological Organization.

The World Bank Group (2021). World bank climate change knowledge portal. `https://climateknowledgeportal.worldbank.org/country/rwanda/climate-data-historical`.

Uwihirwe, J., Hrachowitz, M., and Bogaard, T. (2020). Landslide precipitation thresholds in rwanda. *Landslides*, 17.

Uwihirwe, J., Hrachowitz, M., and Bogaard, T. (2021). Integration of observed and model derived groundwater levels in landslide threshold models in rwanda.

van der Schalie, R., de Jeu, R., Kerr, Y., Wigneron, J., Rodríguez-Fernández, N., Al-Yaari, A., Parinussa, R., Mecklenburg, S., and Drusch, M. (2017). The merging of radiative transfer based surface soil moisture data from smos and amsr-e. *Remote Sensing of Environment*, 189:180–193.

Wang, N., Lombardo, L., Gariano, S. L., Cheng, W., Liu, C., Xiong, J., and Wang, R. (2021). Using satellite rainfall products to assess the triggering conditions for hydro-morphological processes in different geomorphological settings in china. *International Journal of Applied Earth Observation and Geoinformation*, 102:102350.

Zhao, B., Dai, Q., Han, D., Zhang, J., Zhuo, L., and Berti, M. (2019). Application of hydrological model simulations in landslide predictions. *Landslides*, 17:1–15.