Potential of satellite-derived hydro-meteorological information for landslide hazard assessment initiation thresholds in Rwanda

Judith Uwihirwe1,2, Alessia Riveros1,3, Hellen Wanjala4, Jaap Schellekens4, Frederiek Sperna Weiland1, Markus Hrachowitz1, Thom A. Bogaard1

1Section Water Resources, Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, PO Box 5048, 2600 GA, Delft, The Netherlands
2Department of Irrigation and Drainage, School of Agricultural Engineering, College of Agriculture Animal Sciences and Veterinary Medicine, University of Rwanda, PO Box 210, Musanze, Rwanda
3Deltares, PO Box 177, 2600 MH Delft
4Planet Labs, PBC, Wilhelminastraat 43A, 2011 VK Haarlem, The Netherlands

Correspondence to: Judith Uwihirwe (uwihirwejudith@yahoo.fr)

Abstract. Satellite and hydrological model-based technologies provide estimates of rainfall and soil moisture over larger spatial scales and now cover multiple decades, sufficient to explore their value for the development of landslide early warning system in data scarce regions. In this study, we used statistical metrics to compare gauge-based to satellite-based precipitation products and assess their performance in landslide hazard assessment and warning in Rwanda. Similarly, the value of high resolution satellite and hydrological model-derived soil moisture was compared to in situ soil moisture observations at Rwanda weather station sites. Based on statistical indicators, the NASA GPM-based IMERG rainfall product showed the highest skill to reproduce the main spatiotemporal precipitation patterns at the studies sites in Rwanda. Similarly, the satellite and model-derived soil moisture time series broadly reproduce the most important trends of in situ soil moisture observations. We evaluated two categories of landslide meteorological triggering conditions from IMERG satellite precipitation. First, the maximum rainfall amount during a multiple day rainfall event. Second, the cumulative rainfall over the past few day(s). For each category, the antecedent soil moisture recorded at three levels of soil depth, top 5 cm by satellite-based technologies as well as top 50 cm and 2 m through modelling approaches, was included in the statistical models to assess its potential for landslide hazard assessment and warning capabilities. The results reveal the cumulative 3 day rainfall $R_D3$ as the most effective predictor for landslide triggering. This was indicated not only by its highest discriminatory power to distinguish landslide from no landslide conditions (AUC ~0.72) but also the resulting true positive alarms TPR of ~80 %. The modelled antecedent soil moisture in the 50 cm root zone $S_{root(t-3)}$ was the most informative hydrological variable for landslide hazard assessment (AUC ~0.74 and TPR of 84 %). The hydro-meteorological threshold models that incorporate the $S_{root(t-3)}$ and $R_D3$ following the cause–trigger concept in a bilinear framework reveal promising results with improved landslide warning capabilities in terms of reduced rate of false alarms by ~20 % at the expense of a minor reduction of true alarms by ~8 %.
Introduction

Landslides are one of the most prevalent hazards in mountainous regions of the world associated with high rates of fatalities, injuries and economic loss globally (Froude and Petley, 2018; Haque et al., 2016; Kirschbaum et al., 2015; Petley, 2012). According to a recent estimate (Froude and Petley, 2018), precipitation-induced landslides were responsible for a global total of ~55000 deaths over the 13 year period from 2004 to 2016. In landslide prone regions, much effort is therefore put on the implementation of prevention and protection measures to control the most sensitive factors. Landslide early warning systems (LEWS) are used as non-structural and cost effective mitigation measures adopted to minimize landslides harms, loss of life and properties (Calvello et al., 2020; Glade and Nadim, 2014). However, the global landslide research indicated a bias in geographical distribution of LEWS and landslide research with a major gap in Africa (Gariano and Guzzetti, 2016; Guzzetti et al., 2020; Kirschbaum et al., 2015, 2010). According to Guzzetti et al. (2020), there are no LEWS in African countries despite the high number of fatal landslides recorded and the high landslide susceptibility (Broeckx et al., 2018; Kirschbaum et al., 2015). Previous landslide susceptibility analysis, revealed countries along the East African Rift region to be highly susceptible to landslides (Broeckx et al., 2018) despite that often only the most severe landslides are reported and thus a large number is missing. The East African Rift (EAR) was thus identified as a major hotspot of hazardous landslides in Africa with high rate of population exposure (Depicker et al., 2020, 2021a; Monsieurs et al., 2018a). On the long term, this is due to the active continental rifting caused by the persistent divergence of the Victoria and Nubia microplates (Glerum et al., 2020) while on the short term it is controlled by the interactions of prolonged and intense rainstorms in the region with hydrogeological and landscape processes. Rwanda is among the tropical countries located in the western branch of the East African Rift, threatened by landslide hazards (Bizimana and Sönmez, 2015; Nsengiyumva et al., 2018; Nsengiyumva and Valentino, 2020). About 43% of its surface area is classified as having moderate to very high susceptibility to landslides with 49% of local population exposed to landslide risks (Nsengiyumva et al., 2018). The long term landslide predisposing factors in Rwanda, include its pronounced topographic profile, the inherent geological and lithological units, weathering process, earthquakes, demographic pressure and related anthropogenic activities such as deforestation, agriculture expansion, and slope incision through roads construction activities (Bizimana and Sönmez, 2015; Depicker et al., 2015, 2021; Moeyersons, 1989; Monsieurs et al., 2018; Nsengiyumva et al., 2018; Valentino et al., 2021). The development of mining sites and the connected feeder roads also change the nature of natural hillslope through excavation and thus exacerbating landslide susceptibility and risks of slope failures. In addition, urban expansion pushes settlements, industry and infrastructure into hazardous areas that are naturally unstable, thereby further increasing the number of elements at landslide risks. The short-term landslide triggering factors include prolonged and intense rains in addition to the hydrological process that predispose slopes to near failure. In the past 15 years period from January 2006 to May 2021, the landslide inventory in Rwanda indicated about 425 landslide death (~0.6 % of the global landslide death) and about 2000 injuries induced by the above normal hydrological and meteorological factors (Uwihirwe, 2021). The lack of LEWS is one of the important factors for the increasing number of landslide victims in
The development of a robust LEWS hinges on the availability of hydro-meteorological data with sufficient spatio-temporal resolution and an accurate landslide inventory, both of which are scarce. Recently, numerous river catchments in Rwanda have been equipped with groundwater monitoring wells, river water level gauges, soil moisture sensors as well as the automated weather stations. However, the available data records are frequently of insufficient length to build historical time series that overlap with the time periods of landslide inventories and that could be incorporated into landslide hazard assessment thresholds. As a first step towards LEWS development in Africa, Monsieurs et al. (2018a, 2019) used the Landslide Inventory for the central section of the Western branch of the East African Rift (LIWEAR) to define the landslide susceptibility–rainfall and antecedent soil moisture thresholds in the East African Rift region. In Rwanda, Uwirwe et al. (2020) used a statistical approach to define gauge-based precipitation thresholds along with estimates of antecedent precipitation indices. Furthermore, Uwirwe et al., (2021) incorporated regional groundwater level measurements extended with a transfer function noise model to define the landslide hydro-meteorological thresholds for regional landslide hazard assessment. So far, these studies relied exclusively on in situ observed precipitation and hydrological data constrained by the sparsely distributed recording equipment with point scale resolution and gaps in the data record. There is a concern about the omission and/or overgeneralisation of information on the pre-wetting hydrological conditions at the locations of the landslide due to the sparsely distributed hydrological recording equipment (Uwirwe et al., 2021). These pre-wetting conditions regulate the disposition of a slope to near failure (Bogaard and Greco, 2018; Sidle et al., 2019). Including this information in a LEWS may thus be a promising opportunity to decrease the rate of both false and missed alarms (Bogaard and Greco, 2018; Peres et al., 2018). Similar to other hydrological variables, soil moisture exhibits high spatial variability particularly in tropical areas (Dewitte et al., 2021; Kirschbaum et al., 2012; Sekaranom et al., 2020). This spatial variability is hardly covered by on site monitoring equipment due to the sparse observation networks, themselves providing point scale observations only. Alternative ways of incorporating such hydrological state information into landslide hydro-meteorological thresholds have been attempted and include the use of soil moisture estimates from satellite products (Marino et al., 2020; Thomas et al., 2020; Zhuo et al., 2019) as well as from distributed hydrological models (Mostbauer et al., 2018; Preiner et al., 2018, 2019; Wang et al., 2019; Zhao et al., 2020). In this study, we aimed to explore the usefulness of combining soil moisture from satellite products and from a distributed hydrological model with satellite-based precipitation for the estimation of landslide hazard assessment thresholds in Rwanda. We specifically i) investigated the suitability of various satellite precipitation products as substitute for rainfall data from a sparsely distributed gauge network in Rwanda, ii) evaluated the added value of satellite and model-derived soil moisture information recorded at various soil depth and iii) assessed the potential of incorporating such information in empirical landslide hazard assessment threshold models and the warning capabilities in Rwanda.

2 Study area

Rwanda is an evergreen landlocked country geographically located between 1°–3° S and 28°–31° E in the great lakes region of central east Africa with a total area of 26,338 km². It is topographically and geomorphologically characterised by angular hills.
rounded hills and headlands, mountains and volcanoes (Fig. 3) with high elevation in the north western regions reaching up to about 4500m (Fig. 1) and steep slope reaching up to 55% (Fig. 1 and Fig. 2). The geomorphology of Rwanda is characterized by angular hills, rounded hills and headlands, mountains and volcanoes (Fig. 2). The subtropical highland climate, with a long-term mean annual rainfall >1200 mm year$^{-1}$ in the north western highland regions and < 1000 mm year$^{-1}$ in the eastern Savanah region (Fig. 1) and a mean annual temperature of about 19 $^\circ$C, prevails in the country. The north western region of the country is located in a tectonic region with a seismically active volcanic chain and earthquakes being among the possible landslide triggers in Rwanda. The hydrology is characterized by dense networks of lakes and rivers while the hydrogeology consists mainly of fractured aquifers of granite-gneisses, schist, mica schists and complex aquifers in volcanic rocks of the north and south western parts of the country (Fig. 4). The main hydro-geological units in the landslide area include low and semi-permeable fractured schist and mica schist and permeable fractured granites. The weathering products of granites are generally coarse-grained and tend to develop and preserve open joint systems that increase their permeability and thus prone to landslide hazards. The weathering product of schists include clay minerals that tend to fill up the fractures and thus slowing the permeability. However, mica schists are renown as unstable due to rapid weathering, easy splitting along the joints and bedding planes and loss of strength induced by the high content of mica.
3 Methods and data

3.1 Landslide inventory

The inventory for this study, contains landslides recorded from 2007 to 2019. It was accessed from a previous study in Rwanda (Uwihirwe et al., 2021) and was further extended and updated through compilation of additional rainfall-induced landslides as reported from local newspapers, blogs and government technical reports. This landslide inventory was compiled with respect to the methodology adopted by Bach et al. (2010), Kirschbaum et al. (2015); and Monsieurs et al. (2018c). Between 2007 and 2019, the inventory includes 55 accurately dated landslides, 32 of which are located in the catchments modelled for this study (Kivu, Nyabarongo upper and Mukungwa) (Fig. 4). However, it is important to note that this inventory is likely to miss the non-hazardous landslides which are less reported upon than hazardous landslides that led to fatalities/injuries and considerable damages. The inventory provides the location of each recorded landslide but with a varying spatial accuracy of 5 to 25 km depending on the smallest administrative unit recorded by the landslide event reporters. Therefore, a buffer zone of 5km, equivalent to the frequently recorded accuracy, was used around each landslide (Fig. 2) to support the choice of the landslide representative rain gauge. The same areal buffer was used as a footprint to avail the areal satellite precipitation and soil moisture as detailed in Sect. 3.2 and Sect. 3.3.
3.2 Precipitation products and performance evaluation

3.2.1 Gauge-based precipitation and selection of landslide representative data

We accessed daily precipitation data from 19 rain gauges operated by the Rwanda Meteorology Agency. These rain gauges were selected based on their location within the defined buffer of 5km around each landslide location (Fig. 2). Once two or more rain gauges fall within the same buffer zone, the gauges are weighted (Melillo et al., 2018) to select the most representative rain gauge following Eq. (1):

\[ W = \frac{E^2}{d^2D} \]

The weight \( W \) is estimated based on the cumulative event rainfall \( E \) (mm) until the landslide day, the distance between rain gauge and landslide \( d \) (mm), and duration \( D \) (days). A similar procedure was used to select the representative rain gauge for landslides located far (>5 km) from any rain gauge. The selected gauge-based precipitation was used as benchmarks to assess the suitability of satellite precipitation products.
3.1.1. Satellite precipitation products and suitability analysis for LEWS in Rwanda

With the gauge-based precipitation data as reference, we assessed the performance of seven satellite precipitation products summarized in Table 1. These satellite precipitation products were preliminary selected for analysis based on the criteria that their dataset i) at least partially overlap with the landslide inventory period (2007–2019), ii) has at least daily temporal resolution, and iii) is available on Google Earth Engine (GEE).
Table 1. Pre-selected precipitation products and short description.

<table>
<thead>
<tr>
<th>Satellite and reanalysis products</th>
<th>Resolution</th>
<th>Period</th>
<th>Data source description</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRMM 3B42 v7</td>
<td>0.25°</td>
<td>Daily</td>
<td>Passive microwave (PMW) from a variety of low Earth orbit satellites, infrared (IR) data and precipitation gauge supplied by the Global Precipitation Climatology Centre (GPCC) and the Global Precipitation Climate Project (GPCP) monthly product and accumulated to the daily scale.</td>
<td>Huffman et al., 2010</td>
</tr>
<tr>
<td>CHIRPS</td>
<td>0.05°</td>
<td>Daily</td>
<td>GridSat-B1 infrared data and bias-adjusted using the Global Precipitation Climatology Project (GPCP) monthly product.</td>
<td>Funk et al., 2015</td>
</tr>
<tr>
<td>PERSIANN CDR</td>
<td>0.25°</td>
<td>Daily</td>
<td>GridSat-B1 infrared data and bias-adjusted using the Global Precipitation Climatology Project (GPCP) monthly product and accumulated to the daily scale.</td>
<td>Ashouri et al., 2015</td>
</tr>
<tr>
<td>GLDAS 2.1</td>
<td>0.25°</td>
<td>3 hourly</td>
<td>Geostationary infrared (IR) cloud-top temperature measurements and microwave observation techniques.</td>
<td>Rodell et al., 2004</td>
</tr>
<tr>
<td>CFSv2</td>
<td>0.2°</td>
<td>6 hourly</td>
<td>Satellite observations in the infrared and microwave channels and gauge observations.</td>
<td>Saha et al., 2014</td>
</tr>
<tr>
<td>IMERG, GPM</td>
<td>0.1°</td>
<td>30 min</td>
<td>Passive Microwave from various low Earth orbit satellites, infrared from geosynchronous Earth orbit satellites and gauges precipitation (successor of TRMM).</td>
<td>Huffman et al., 2020</td>
</tr>
<tr>
<td>ERA5</td>
<td>0.25°</td>
<td>3 hourly</td>
<td>This is a non-satellite but re-analysis product. precipitation is generated employing a convection scheme along with the large-scale cloud scheme that have been upgraded with an improved representation of mixed-phase clouds, and prognostic variables for precipitating rain and snow.</td>
<td>Hersbach et al., 2020</td>
</tr>
</tbody>
</table>

Among the pre-selected satellite products, we have chosen the most suitable product for landslide hazard assessment in Rwanda based on the relative comparison with gauge-based precipitation. This was achieved using a number of statistical approaches that include: i) the use of statistical metrics of goodness of fit, ii) rainfall frequency indicators, and iii) intensity comparisons. The statistical metrics of goodness of fit include the root mean square error RMSE, Pearson correlation CC, and the long-term relative bias RB computed with Eq. (2) to Eq. (4):

\[
RMSE = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (Y_i - X_i)^2}, \tag{2}
\]

\[
CC = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y}) (X_i - \bar{X})}{\sqrt{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2}}, \tag{3}
\]

\[
RB = \frac{\sum_{i=1}^{n} (Y_i - \bar{Y}) - (X_i - \bar{X})}{\bar{Y}}, \tag{4}
\]

Where \(Y_i\) is the rain gauge observation at date \(i\), \(X_i\) is the satellite estimate at the same date \(i\), \(n\) is the total number of data pairs for each precipitation product considered, \(Y_{mean}\) and \(X_{mean}\) are the mean rainfall from rain gauge and satellite products respectively.
The rainfall frequency indicators specify the frequency of rainy days based on the predefined threshold indices (Joshi et al., 2014; Tank et al., 2009). We used five rainfall threshold indices that reflect the number of rainy days with >X mm of rain (RDx). The predefined indices are RD0, RD10, RD20, RD30, and RD50 indicating the number of rainy days with >0 mm as rainy days, >10 mm as heavy rainy days, >20 mm as very heavy rainy days, >30 mm as even heavier rainy days, and >50 mm as extremely heavy rainy days respectively. With intensity comparison, we compared the cumulative 30 day rainfall from the satellite precipitation products to the cumulative 30 day precipitation from rain gauges using scatter plots.

3.3 Soil moisture products and data acquisition

3.3.1 In situ soil moisture data from automatic weather stations

In situ soil moisture data, collected from the automatic weather stations (AWSs) equipped with soil moisture sensors, were accessed from the Rwanda Meteorological Agency for six AWSs as shown in Fig. 3. The AWSs recorded the soil moisture at 20 cm depth with a temporal resolution of 5-10 minutes from July 2018 to December 2019. Because the analysis focuses on a daily time-scale, we computed and used the daily average soil moisture time series recorded from July 2018 to December 2019. The in situ AWSs soil moisture data were used as a benchmark to comparatively get an insight on the quality of other sources of soil moisture products that include satellite and model-derived soil moisture estimates described in Sect. 3.3.2 and Sect. 3.3.3.

3.3.2 Satellite soil moisture and variable of interest

We used a satellite-derived near surface soil moisture product provided by Planet, formerly VanderSat (VdS) (https://vandersat.com/data/soil-moisture/, last access: 29 March 2022). The product relies on the Land Parameter Retrieval Model (LPRM) (De Jeu et al., 2014; Owe et al., 2001, 2008) to estimate the near surface soil moisture by combining raw data from the Advanced Microwave Scanning Radiometer 2 (AMSR-2), and Soil Moisture Active Passive (SMAP) (Bouaziz et al., 2020). The satellite product estimates volumetric soil water content or soil moisture (m³/m³) of the upper 5 cm of soil downscaled from a spatial resolution of 25 km x 25 km to 100 m x 100 m. From VdS, we accessed daily soil water content estimates from the top 5 cm of soil (θtop) for the 2007-2019 period for each of the defined regions of interest (ROIs) equivalent to the 5 km buffers shown in Fig. 3.
3.3.3 Hydrological model-derived soil moisture and variables of interest

We also used the soil moisture derived from the Wflow_sbm, a distributed hydrological model that uses the conceptual bucket model approach to estimate soil water content (Imhoff et al., 2020). With Wflow_sbm, the soil is considered as a bucket with a depth \( Z \) divided into 2 zones: the unsaturated store \( U \) and the saturated store \( S \). The interface between \( U \) and \( S \) is a pseudo water table located at depth \( Z_w \). The values of unsaturated storage \( U \) and saturated storage \( S \) are computed as in Eq. (5) and Eq. (6):

\[
U = \frac{Z - Z_w}{Z} \cdot \text{saturation}
\]

\[
S = \text{saturation} \cdot \text{depth}
\]

We also used the soil moisture derived from the Wflow, an open source software developed by the Deltares OpenStreams project (Schellekens, 2021; Schellekens et al., 2019). The Wflow distributed hydrological model platform currently contains 13 models (Schellekens, 2021) that include the wflow_sbm model. The models consist of a set of python programs that are
run on the PCRaster python framework to perform hydrological simulations (Karssenberg, 2014; Karssenberg et al., 2010). The Wflow sbm uses the conceptual bucket model approach to derive the hydrological variables of interest (Imhoff et al., 2020; Schellekens et al., 2019). With Wflow sbm, the soil is considered as a bucket with a depth ($Z$) divided into 2 zones: the unsaturated store $U$ and the saturated store $S$. The interface between $U$ and $S$ is a pseudo water table located at depth $Z_w$. The values of unsaturated storage $U$ and saturated storage $S$ are computed as in Eq. (5) and Eq. (6):

$$U = (\theta_s - \theta_r)Z_w - U_d,$$  \hspace{1cm} (5)

$$S = (\theta_s - \theta_r)(Z - Z_w).$$  \hspace{1cm} (6)

Where $\theta_r$, $\theta_s$, and $U_d$ are saturated and residual water content respectively and $U_d$ is the soil water deficit.

The unsaturated store $U$ was the variable of interest and was subdivided into 2 variables: the water content in the root zone $\theta_{\text{root}}$ representing the unsaturated soil water storage of the top 50 cm and the part of the soil water capacity occupied $\theta_{\text{uz}}$ representing the unsaturated soil water storage of the upper 2 m. For this study, the model area consisted of three catchments (Kivu, Upper Nyabarongo and Mukungwa) as highlighted in Fig. 4. We obtained time series of $\theta_{\text{root}}$ and $\theta_{\text{uz}}$. As with all other hydrological models, the Wflow sbm model calibration is needed for optimal performance. The model is calibrated based on the model parameters that are estimated a priori using pedo-transfer functions and global or local datasets. The model was run from January 2001 till December 2020 with the first years correspond to the model spin up time. We obtained time series of $\theta_{\text{root}}$ and $\theta_{\text{uz}}$ for 2007–2019, overlapping with the landslide inventory period, from a wflow sbm simulation based on ERA5 re-analysis meteorological data. To increase the comparability with the satellite-based soil moisture, the same ROIs represented by the buffers of 5 km around each landslide location were used to interpolate the unsaturated water storage time series for each ROIs located in the model catchment. Similarly, only the AWSs located in the model catchment (Fig. 4) were used for the comparative performance evaluation of the model-derived soil moisture products.
3.4 Landslide hazard assessment and thresholds definition

3.4.1 Landslide meteorological and hydrological conditions and test variables

The daily rainfall data from the best performing satellite product were used to define the landslide meteorological triggering conditions. We used two categories of landslide triggering conditions. The first category defined a landslide trigger as the maximum probable rainfall event (MPRE) during which or slightly after its end, one or more landslides occurred. The MPREs were defined as individual periods of rainy days interrupted by dry periods of at least two days. Given the constraint of overestimation of the number of rainy days with 0–10 mm by satellites (Pavez, 2021), a rainy day was objectively referred to as the day with ≥ 10 mm d\(^{-1}\).
while a dry day was referred to as the day with <10 mm d\(^{-1}\). This threshold was objectively selected using the rainfall frequency indicator metric explained in Sect. 3.2.2. The landslide predictor variables in this category were therefore the cumulated event rainfall, event volume (\(E (\text{mm})\)), event duration (\(D (\text{days})\)) and event-rainfall mean intensity (\(E/D (\text{mm d}^{-1})\)). The cumulated event rainfall, event volume \(E (\text{mm})\) was computed as the cumulative rainfall during each MPRE of duration \(D (\text{days})\). The event duration \(D\) equivalent to MPRE is the individual periods of days with recorded rainfall interrupted separated by inter event time (IET) of at least two dry days. The event-rainfall mean intensity \(E/D\) is the ratio of event rainfall volume \(E\) and event duration \(D\). The second category defined a landslide trigger as the recent cumulative rainfall \(R_p\) at the end of which one or more landslides occurred. This category considers the moving window total cumulative rainfall over the last three days \((R_{D3})\), two days \((R_{D2})\), and one day \((R_{D1})\) at the end of which, one or more landslides occurred. While MPREs time series are interrupted by the IETs, the \(R_{D3}, R_{D2}\) and \(R_{D1}\) for each day during the 2007–2019 study period were computed regardless of a rainy or dry day and thus resulting into longer time series and more data points compared to the MPREs time series. The time series of the defined meteorological triggering conditions from each category and for all precipitation foot prints were compiled in a single dataset for further statistical analysis. To provide a normalized comparison of the soil wetness, we transformed the satellite- and model-derived water content \(\theta\) to effective soil moisture \(S_e\) to define the landslide predisposing hydrological conditions using Eq. (7):

\[
S_e = \frac{\theta - \theta_{\text{min}}}{\theta_{\text{max}} - \theta_{\text{min}}} 
\]

Where \(S_e\) stands for the effective soil moisture [-], \(\theta\) is the actual soil moisture, \(\theta_{\text{max}}\) and \(\theta_{\text{min}}\) are the maximum and minimum values of the recorded or modelled soil water content.

The normalization of soil water content \(\theta\) was made for easy comparison of the observed, model-derived, and satellite-based soil moisture products. However, for all compared soil moisture products, \(\theta_{\text{max}}\) and \(\theta_{\text{min}}\) were 1 and 0 respectively which lead to almost similar values of \(S_e\) and \(\theta\).

The tested hydrological conditions include therefore, the near surface soil moisture \(S_{\text{tops}}\) representing the soil moisture of the upper 5 cm of soil, provided by the satellite techniques (VdS), the \(S_{\text{root}}\) representing the root zone soil moisture of the upper 50 cm, acquired through modelling approach (Wflow \(_{\text{sbm}}\)), and \(S_{\text{sat}}\) representing the soil moisture estimates from the upper 2 m of soil, obtained through modelling approach (Wflow \(_{\text{sbm}}\)). To assess the contribution of the pre-wetting state of the soil prior to the landslide triggering conditions, we have considered the antecedent soil moisture i.e. recorded or modelled prior to the start of the triggering meteorological conditions. The antecedent soil moisture referred to the time interval before the start of each of the defined categories of the meteorological triggering conditions. We have therefore used the \(S_{\text{tops}(t-x)}, S_{\text{root}(t-x)}, S_{\text{sat}(t-x)}\) with \(t\) (date) and \(x\) (days) expressing the end time and the duration of the triggering conditions respectively. However,
due to the transient changes in duration of the MPREs, x was hypothetically represented by a value of 1 standing for one entire MPRE while values of 1, 2, 3 represent the duration (days) of the triggering \( R_{Dx} \) conditions. A binary classification of the defined hydrological and meteorological conditions was undertaken to classify the landslide and no-landslide conditions. The meteorological or hydrological conditions are referred to as landslide conditions i.e. positive class, when at least one landslide occurs during its course or slightly after its end while they are referred to as no-landslide conditions i.e. negative class when no landslide occurred during its course or slightly after its end.

### 3.4.2 Discriminatory power of the landslide test variables and optimum thresholds for landslide initiation

The landslide test variables which include the predisposing hydrological conditions \( S_{top(t-1)} \), \( S_{top(t-2)} \), \( S_{top(t-3)} \), \( S_{root(t-1)} \), \( S_{root(t-2)} \), \( S_{root(t-3)} \), \( S_{ax(t-1)} \), \( S_{ax(t-2)} \), and \( S_{ax(t-3)} \) as well as the triggering meteorological conditions \( E \), \( D \), \( E/D \), \( R_D \), \( R_D \), and \( R_D \) were tested for their relevance on landslide occurrence. We used a receiver operating characteristic (ROC) and the area under the curve (AUC) metrics to evaluate the discriminatory power of each of the landslide test variables. The ROC curve is defined as a graphical plot indicating the performance of the test variable at all threshold levels by providing the trade-off between the true positive rate (TPR) and false positive rate (FPR) at each level. The AUC is a statistical metric that indicate the discriminatory power of the test variable i.e. the capacity of the test variable to correctly distinguish positive from negative classes i.e. landslide from no landslide conditions. It compares also the test variable to a random guess (AUC=0.5) and thereby indicates the statistical significance where the perfect test variable would have an AUC equal to unity. The rate of correctly (TPR, true positive rate) and incorrectly predicted landslides (FPR, false positive rate) corresponding to each cut off on the ROC curves are computed using Eq. (8) and Eq. (9):

\[
TPR = \frac{TP}{TP+FN} \quad (8)
\]

\[
FPR = \frac{FP}{FP+TN} \quad (9)
\]

The rate of unpredicted landslides (FNR, false negative rate) and the rate of correct prediction of no-landsides (TNR, true negative rates) are computed using Eq. (10) and Eq. (11):

\[
FNR = \frac{FN}{FN+TN} \quad (10)
\]

\[
TNR = \frac{TN}{TN+FP} \quad (11)
\]
Where TP are true positives or true alarms i.e. outcomes with correctly predicted landslides, FN are false negatives or missed alarms i.e. the number of landslides that occurred in reality but were not predicted, FP are false positives or false alarms i.e. predictions of landslide occurrence while in reality there was no landslide reported, and TN are true negatives i.e. correct predictions of no-landslide occurrence.

Since the ROC curve only indicates all possible thresholds and their relative balance between TPR and FPR, one is free to choose the optimum threshold depending on whether to maximize the TPR or minimize the FPR. However, according to Postance and Hillier (2017), the optimum threshold is the one that maximizes the TPR while minimizing the FPR. Therefore, that optimum threshold level above which landslides are highly likely to occur have been defined using two statistical metrics i.e. the maximum true skill statistic (TSS) and minimum radial distance (Rad). The TSS is expressed as a balance between the TPR and FPR as indicated in Eq. (12):

\[
\text{TSS} = \text{TPR} - \text{FPR}.
\]  

(12)

Where the maximum value of TSS indicates the optimum threshold for landslide initiation that maximizes the TPR while minimizing the FPR. For a perfect threshold, the TSS reaches a unity indicating a zero false positive rate (FPR). The radial distance (Rad) shows the relative distance from the defined threshold level on the curve to the perfect model or point whose TPR is a unit and zero FPR and is computed with Eq. (13):

\[
\text{Rad} = \sqrt{(\text{FPR})^2 + (\text{TPR} - 1)^2}.
\]  

(13)

3.4.3 Landslide hydro-meteorological thresholds and warning capabilities

The optimum thresholds defined based on the maximum TSS and or minimum Rad were plotted in 1D threshold space here referred to as single variable threshold line beyond which landslide are the probability of landslides is highly likely to occur. We also followed cause-trigger concept (Bogaard and Greco, 2018) that reflect the hydro-meteorological thresholds and hypothetically plotted the optimum thresholds of the landslide predisposing hydrological variables i.e the antecedent soil moisture on the x-axis and the meteorological triggering variables on the y-axis of a two dimensional 2D space here referred to as bilinear thresholds. The bilinear threshold models made of hydrological and meteorological variables are plotted in x, y pairs i.e antecedent soil moisture versus E/D or RX). Furthermore, the bilinear threshold from a traditional landslide prediction model event duration E-D, Furthermore, the bilinear threshold from a traditional landslide prediction model event duration E-D (Peruccacci et al., 2017), that exclusively rely on precipitation, has been also defined to serve as a benchmark for comparative performance evaluation.
4 Results and discussion

4.1 Performance of Satellite precipitation products

The suitability of satellite precipitation products in the study region was assessed using three statistical indicators as summarized in Table 2, Table 3 and illustrated in Fig. 4. From the statistical measures of fits (RMSE, CC, RB), it is generally observed that IMERG is consistently more suitable while ERA-5 was found to be the least suitable product as compared to other satellite precipitation products. The evaluation based on frequency indicators is summarized in Table 3. These indicators give an overview on whether a given satellite product would overestimate or underestimate the observed gauge precipitation based on the predefined threshold indices. IMERG_GPM displays the highest skill to estimate all ranges of rainfall from heavy to extremely heavy rainy days as recorded by the on-site gauges. CHIRPS and TRMM 3B42 v7 provide good estimates of precipitation with quite similar number of rainy days (R_{D0}=1250\text{days}) to gauge-based rainfall (R_{D0}=1256\text{days}). However, these satellites drastically underestimate the number of heavy to extremely heavy rainy days (R_{D20}, R_{D30} and R_{D50}) which are the root cause of landslide hazards. For example, TRMM and CHIRPS estimated R_{D20}=84\text{days} and 101\text{days} respectively out of 434\text{days} estimated by rain gauge. The products PERSIANN CDR, GLDAS 2.1, CFSv2 and ERA-5 overestimated low rainy days while underestimating the number of heavy to extremely heavy rainy days as shown in (Table 3).

The suitability of satellite products was also assessed using intensity comparison indicated by the density of the scatter points around 1:1 line as shown in Fig. 4. The scatter plots compare 30day cumulative rainfall from satellite precipitation products versus rain gauges. The scatter plots reveal that GLDAS, CFSv2 and ERA-5 tend to overestimate rainfall while underestimations are noticed from PERSIANN CDR as compared to the in situ gauge rainfall. Based on the closeness of scatter points to the 1:1 line, CHIRPS and IMERG exhibit a higher resemblance to gauge data (Pearson correlation R=0.67 and 0.60 respectively) than other satellite products and could thus be used as alternative to gauge-based precipitation. Overall, IMERG shows rainfall pattern that are most consistent with available gauge observations in Rwanda despite the over estimation of the number of rainy days with less than 10 mm (R_{D10}). According to Kimani et al. (2017), the overestimation of rainfall in areas with elevation >2500 m and underestimation in areas with elevation < 2500 m was observed before and is attributed to satellite inherent challenges to retrieve orographic rainfall. To overcome this constraint, 10 mm d^3 has been considered as a threshold to define a satellite-based rainy day and thus being relevant for landslide hazard assessment in Rwanda conditions. Other researchers in the regions also found CHIRPS and TRMM to be comparable to gauge-based precipitation in east Africa (Kimani et al., 2017; Monsieurs et al., 2018b). Monsieurs et al. (2018b) found the area-averaged TMPA rainfall estimates, the predecessor of IMERG, to be more suitable for assessing landslide hazard threshold than the sparsely distributed gauge data with limited representativeness in the context of high rainfall variability of the east African rift.
Table 2. Performance of Satellite precipitation products based on statistical metrics

<table>
<thead>
<tr>
<th>Metrics</th>
<th>TRMM 3B42 v7</th>
<th>CHIRPS</th>
<th>PERSIANN CDR</th>
<th>GLDAS 2.1</th>
<th>CFSv2</th>
<th>IMERG</th>
<th>ERA 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE (mm)</td>
<td>8.17</td>
<td>8.53</td>
<td>7.42</td>
<td>8.55</td>
<td>10.58</td>
<td>8.18</td>
<td>12.60</td>
</tr>
<tr>
<td>CC (•)</td>
<td>0.31</td>
<td>0.27</td>
<td>0.25</td>
<td>0.24</td>
<td>0.17</td>
<td>0.35</td>
<td>0.22</td>
</tr>
<tr>
<td>RR (•)</td>
<td>-0.08</td>
<td>-0.01</td>
<td>-0.15</td>
<td>0.03</td>
<td>0.11</td>
<td>0.02</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 3. Performance of Satellite precipitation products based on rainfall frequency indicators

<table>
<thead>
<tr>
<th>Indices</th>
<th>Description</th>
<th>Gauge</th>
<th>TRMM 3B42 v7</th>
<th>CHIRPS</th>
<th>PERSIANN CDR</th>
<th>GLDAS 2.1</th>
<th>CFSv2</th>
<th>IMERG</th>
<th>ERA 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_{50}</td>
<td>Rainy days &gt;0 mm</td>
<td>1259</td>
<td>1691</td>
<td>1256</td>
<td>2732</td>
<td>3086</td>
<td>2835</td>
<td>2842</td>
<td>3520</td>
</tr>
<tr>
<td>R_{100}</td>
<td>Heavy rainy days &gt;10 mm</td>
<td>397</td>
<td>307</td>
<td>424</td>
<td>138</td>
<td>377</td>
<td>617</td>
<td>383</td>
<td>879</td>
</tr>
<tr>
<td>R_{200}</td>
<td>Very heavy rainy days &gt;20 mm</td>
<td>132</td>
<td>87</td>
<td>101</td>
<td>9</td>
<td>79</td>
<td>199</td>
<td>126</td>
<td>250</td>
</tr>
<tr>
<td>R_{300}</td>
<td>Even heavier rainfall days &gt;30 mm</td>
<td>49</td>
<td>29</td>
<td>25</td>
<td>0</td>
<td>22</td>
<td>84</td>
<td>42</td>
<td>78</td>
</tr>
<tr>
<td>R_{500}</td>
<td>Extremely heavy rainfall &gt;50 mm</td>
<td>9</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>22</td>
<td>6</td>
<td>21</td>
</tr>
</tbody>
</table>

Figure 4. Intensity comparison between satellite-based and gauge-based precipitation based on the cumulative 30 day rainfall

4.2 Prospective of satellite and model-derived information in landslide hazard assessment

4.2.1 Mean soil moisture response to rainfall and landslide events
Figure 6 indicates the GPM based IMERG precipitation spatially averaged over all landslide precipitation footprints and over the modelled catchments. It shows also the temporal dynamics of the satellite estimates $S_{\text{top}}$ and the model-derived soil moisture time series $S_{\text{root}}$ spatially averaged over all landslide ROIs and $S_{\text{uz}}$ compared to the modelled soil moisture at the root zone top 50 cm $S_{\text{root}}$ and top 2 m $S_{\text{uz}}$. The in situ soil moisture observations from the automatic weather station AWS is also rescaled (on secondary axis) to be compared with modelled and satellite derived soil moisture. Regardless of the difference in measuring depth (5 cm, 50 cm & 2 m), the time response to precipitation and overestimation of soil moisture, the satellite $S_{\text{top}}$ and model-derived soil moisture time series $S_{\text{root}}$ and $S_{\text{uz}}$ broadly reproduce the most important temporal variation as recorded by in situ soil moisture sensors (Fig. 6 and Fig. A1). This indicates their usefulness for landslide hazards assessment as an alternative to the sparse in situ AWSs. The spatial averaging of soil moisture across all research areas of interest (ROIs) was undertaken to have an insight on the critical ranges of soil moisture that induce landslides in Rwandan climate conditions. The spatially averaged $S_{\text{top}}$, $S_{\text{root}}$ and $S_{\text{uz}}$ soil moisture dynamics and the linked landslide hazards occurrence are presented in Fig. 6. The average $S_{\text{top}}$, $S_{\text{root}}$ and $S_{\text{uz}}$ of all ROIs, indicate general similarities in terms of landslide predisposing but also reveal systematic differences between response time influenced by the soil moisture recording depth. For example, it is obvious that the $S_{\text{top}}$ (5 cm) responds faster than $S_{\text{root}}$ (50 cm) and $S_{\text{uz}}$ (2 m). It is clear that the majority of landslides occurs when the soil moisture levels positively deviate (by about 0.1) from the long-term mean up to a critical level for landslide initiation. It is also evident that the critical level for landslide occurrence is a function of more or less fixed when other geological and geomorphological condition are kept constant and it is reached more or less easily depending on the prior rainfall received and expressed in terms of antecedent soil moisture and the time lag between the landslide triggering rainfall and the soil moisture response hydrologic response.
Figure 5. Satellite and model-derived information and landslide activities (a) GPM based IMERG precipitation [mm] spatially averaged over all landslide precipitation footprints (b) satellite derived soil moisture $S_{\text{top}}$ [-] spatially averaged over all landslide ROIs and in situ soil moisture AWS [-] on secondary y-axis; (c) GPM based IMERG precipitation [mm] spatially averaged over the landslide precipitation footprint located in the modelled catchments, (d) modelled soil moisture at the root zone top 50 cm $S_{\text{root}}$ [-], modelled soil moisture top 2 m $S_{\text{uz}}$ [-] and in situ soil moisture AWS [-] on secondary y-axis. The dashed horizontal lines represent the long term mean soil moisture and the red triangles stand for the landslide events.
4.2.2 Single variable landslide meteorological and hydrological thresholds and prediction capabilities

Figure 7 and Table 4 show the derived landslide meteorological and hydrological thresholds and their predictive capabilities in terms of true positive rate TPR and false positive rate FPR. The discriminatory power of each of the tested variables was evaluated with a receiver operating characteristic (ROC) curve and the area under the curve (AUC) statistical metrics as shown in Fig. 2b. Among the tested landslide triggering meteorological variables $E, D, E/D, R_{D1}, R_{D2},$ and $R_{D3},$ the cumulative 3 day rainfall $R_{D3}$ and event rainfall volume $E$ showed the highest discriminatory power with $AUC \approx 0.7271$ and hence the highest impact on landslide initiation. However, the event rainfall mean intensity $E$ normalised over the event duration $D$ as $E/D$ indicated low capability ($AUC \approx 0.53$) to distinguish landslides from no landslides. This stresses the importance of using the recent cumulative rainfall with a fixed duration and thus highlighting the highest impact of $R_{D3}$ on landslide initiation process and its relevance on landslide hazard assessment and prediction compared to $E$ that need to be normalised. Contrarily to the gauge-based cumulative rainfall thresholds (Uwihirwe et al., 2020), the satellite based cumulative rainfall on the day of landslide $R_{D1}$ was not impactful to landslide initiation ($AUC=0.35–0.38$). This may be due to the inaccuracies between the landslide occurrence and the reporting time, and additionally also due to the satellite revisiting time and or period which may introduce inaccuracies in timing.

Figure 7c and d indicate that the wetness state of soil prior to the cumulative rainfall $R_{D3}$ have the most significant impact on landslide occurrence as indicated by their $AUC=0.72–0.76$. Contrarily, Fig. 2a and b show that the wetness state of the soil prior to the landslide triggering event $E$ has no significant impact on landslide occurrence ($AUC=0.6566–0.67$). This is to say that the antecedent soil moisture conditions prior to the longer triggering rainfall event $E$ are not relevant for landslide initiation in the study area conditions. Among other factors, the duration of the triggering condition plays a major role in determining the relevance of the antecedent soil moisture on landslide occurrence. The shorter the duration of the triggering conditions, the higher the relevance of the antecedent soil moisture on landslides initiation. Highly permeable soils are less sensitive to antecedent soil moisture conditions because of the high gravity driven drainage and or deep percolation. With a tropical climate, evaporation process may also rapidly take away the antecedent soil moisture content of the top soil due to the longer timescale of the minimum inter-event time IET and the landslide triggering event $E$.

The thresholds definition metrics, TSS and Rad, resulted in quite comparable landslide thresholds as summarised in Table 4. Moreover, it was noticed that the defined satellite precipitation thresholds are more similar to the ones defined using gauge based precipitation. For example, the optimum landslide threshold event rainfall volume $E$ defined from satellite precipitation varied between $44\text{mm}–47\text{mm}$ and $64\text{mm}–60.7\text{mm}$ (Table 4) while gauge-based threshold $E$ varied from $46\text{mm}$ to $67\text{mm}$ (Uwihirwe et al., 2021). Similarly but with a quite minor difference, the defined satellite-based $E/D$ thresholds $16–17.5\text{mmd}^{-1}$ seemed quite similar to gauge-based thresholds $7–13\text{mmd}^{-1}$ found in Uwihirwe et al. (2020 and 2021). Nevertheless, the
single variable threshold E/D being the most informative, showed quite low prediction capability in terms of TPR~56–60 % with elevated rate of false positive FPR ~ 43–54 % i.e. incorrect predictions of landslide and thus being less effective for a robust early warning system development. Contrarily, the single variable thresholds defined from the cumulative 3 day rainfall \( R_{D3} \) outperforms other tested triggering conditions with highest prediction capability in terms of true positive rate TPR=79–81 %. The same holds for the soil moisture in the root zone (50 cm deep) \( S_{e_root} \) that consistently showed the highest performance. Nevertheless, despite the high true positive rate from these single variables thresholds, the resulting elevated rate of false positives FPR=36–42 % still constrain their use for the development of a robust landslide early warning system. It has to be noted that the threshold defined from the antecedent soil moisture specifies the critical levels below which the impact of pre-wetting state of the soil is considered unimportant for landslide occurrence. On the contrary, once these thresholds are exceeded, the pre-wetting state of the soil has significant impact on landslide occurrence and has to be considered while defining the landslide hydro-meteorological threshold models.
Figure 6. Receiver operating characteristics ROC curves, area under the curves AUC and optimum landslide thresholds defined by the true skill statistic TSS (square shaped marker) and radial distance Rad (cycle shaped marker) using: a) Event rainfall and satellite (VdS) based top 5 cm soil moisture $S_{\text{Se,top}}$ from all ROIs b) Event rainfall and modelled root zone soil moisture of the top 50 cm $S_{\text{Se,root}}$ and top 2 m $S_{\text{Se,uz}}$ from ROIs located in the Wflow model catchment c) Cumulative 1, 2 and 3 day rainfall (R$_D$) and satellite (VdS) based top 5 cm soil moisture $S_{\text{Se,top}}$ from all ROIs and d) cumulative 1, 2 and 3 day rainfall (R$_D$) and modelled root zone soil moisture of the top 50 cm $S_{\text{Se,root}}$ and top 2 m $S_{\text{Se,uz}}$ from ROIs located in the Wflow model catchment.
4.2.3 Landslide hydro-meteorological thresholds and implication for warning

With respect to the high rate of false positives resulting from the single variable thresholds, we have tested whether the incorporation of antecedent soil moisture information to the triggering rainfall conditions improves the landslide prediction capability. The optimum single variable hydrological and meteorological thresholds have been therefore combined into hydro-meteorological thresholds following the cause-trigger concept in a bilinear framework as shown in Fig. 82 and Fig. 98. Figure 87 illustrates the first category of landslide hydro-meteorological thresholds defined based on the maximum possible rainfall event E combined with different variables of antecedent soil moisture. The derived thresholds resulted into quite elevated rate of false alarms FPR once used as single variable thresholds (single lines).
In contrast to the classical precipitation thresholds, the combination of hydro-meteorological thresholds in a bilinear framework provide an improvement in terms of reduced rate of false alarms by about 30% \( [\text{Se}_{\text{root}}(t-1) - \text{E}]/\text{D} \), 13% \( [\text{Se}_{\text{root}}(t-1) - \text{E}]/\text{D} \), and 35% \( [\text{Se}_{\text{root}}(t-1) - \text{E}]/\text{D} \) respectively as compared to the ones obtained from the exclusive use of single variable precipitation based \( \text{E}/\text{D} \) thresholds. The intention of adopting the bilinear hydro-meteorological threshold in spite of precipitation thresholds is to minimize the rate of incorrect prediction of landslides FPR while improving or at least keeping unchanged the rate of true alarms TPR. This was only achieved by using the bilinear hydro-meteorological thresholds defined using antecedent soil moisture at the root zone \( [\text{Se}_{\text{root}}(t-1) - \text{E}]/\text{D} \) that performs better (TPR=66%) than the traditional precipitation threshold \( \text{E}-\text{D} \) (TPR=50%). However, this category still suffers from the low landslide warning capability (max TPR=66%) and is thus not satisfactory for a robust early warning system development. The lower performance was attributed to the timescale of the triggering events. Apparently, the effect of the antecedent soil moisture lasts for a limited period of time and subsequently decays towards zero and below.

The inter-event time IET and the timescale of the rainfall events \( \text{E} \) are not constant and vary in durations with high probability to landslides induction. Consequently, the incorporation of the wetness state of the soil prior to the landslide triggering events \( \text{E} \) did not lead to a significant improvement of the landslide prediction in Rwanda conditions. We therefore explored other landslide hydro-meteorological thresholds that use the triggering meteorological conditions with short and constant timescale as shown in Fig. 98. These consider the cumulative one, two and three day rainfall \( R_{01}, R_{02}, \) and \( R_{03} \) while extending the timescale of the predisposing conditions up to one, two or three days prior to the landslide triggering conditions. Figure 98 portrays the optimum bilinear hydro-meteorological threshold models defined from this second category. The 3-day cumulative rainfall \( R_{03} \) was the most impactful trigger of landslide with an optimum threshold \( \sim 15.4 \text{mm} / \text{day} \) defined by both TSS &Rad and resulted into 79–81% of TPR (Table 5) much higher than predicted by the first category. Similarly, the antecedent soil moisture threshold \( \text{Se}_{\text{root}}(t-3) \) was able to predict \( \sim 84\% \) of landslides. However, this true prediction i.e true alarms is also associated with high rate of false alarms \( \sim 40–42\% \). The combination into hydro-meteorological thresholds \( [\text{Se}_{\text{root}}(t-3) - R_{01}] \) decreased the rate of false alarms up to \( \sim 22\% \) with about 72% of true alarms (Fig. 98b) and thus being more satisfactory than other hydro-meteorological threshold models and much better than the traditional \( \text{E}-\text{D} \) model (TPR=50%).
Figure 7. Landslide hydro-meteorological thresholds and prediction capabilities: (a) Event intensity-Antecedent 5cm top soil moisture thresholds \( \text{Se}_{\text{top},-1} > 0.56; \ AND \ E/D > 16.1 \text{mm}^{-1} \) (b) Event-duration E-D thresholds \( D > 2.5 \text{ days}; \ AND \ E > 44.9 \text{ mm} \) defined using precipitation foot prints from all landslide locations; (c) Event intensity-Antecedent 50 cm top soil moisture threshold \( \text{Se}_{\text{root},-1} > 0.56; \ AND \ E/D > 16.1 \text{mm} \); (d) Event intensity-Antecedent 2 m top soil moisture threshold \( \text{Se}_{\text{root},-1} > 0.84; \ AND \ E/D > 16.1 \text{mm}^{-1} \); (e) Event-duration E-D thresholds \( E > 60.7 \text{ mm}; \ AND \ D > 2.5 \text{ days} \) defined using precipitation foot prints and landslides located in Wflow modelled catchments; (f) Bilinear threshold values and prediction capabilities
Figure 8.
Figure 9. Landslide hydro-meteorological thresholds and prediction capabilities: (a) Cumulative 3–day rainfall $R_{D3}$ and antecedent 5 cm top soil moisture thresholds $S_{e_{top}}(t-3) > 0.73$, AND $R_{D3} > 15$ mm d$^{-1}$ defined using precipitation foot prints from all landslide locations; (b) Cumulative 3–day rainfall $R_{D3}$ and antecedent soil moisture of the root zone $S_{e_{root}}(t-3) > 0.73$, AND $R_{D3} > 15$ mm d$^{-1}$; (c) Cumulative 3–day rainfall $R_{D3}$ and antecedent soil moisture of the top 2 m $S_{e_{top}}(t-3) > 0.89$, AND $R_{D3} > 15$ mm d$^{-1}$ defined from the Wflow model catchment (d) Bilinear hydro-meteorological threshold values and prediction capabilities
4.2.4 Prospective of the satellite-based hydro-meteorological thresholds, advances and limitations

This study reveals the high capability of the NASA GPM-based IMERG product to reproduce rainfall patterns which are consistent with the gauge-based precipitation and thus more suitable for landslide hazard assessment thresholds than sparsely distributed rain gauges in Rwanda. However, this research also points out that IMERG satellite-based product overestimates the number of rainy days whose daily rainfall is between 0–10 mm and thus the mean annual totals. This may not only lead to differences between satellite- and gauge-based landslide thresholds defined under same locations but also to the statistical bias especially when probabilistic methods are used for landslide threshold definition. To address this constraint and be able to exploit the usefulness of IMERG precipitation in landslide hazards assessment thresholds, we objectively used 10 mm as a threshold to define a rainy day for IMERG precipitation data. This threshold was defined based on the frequency indicator metric- adopted as one of the techniques of bias evaluation between ground and satellite-based rainfall. For gauge-based rainfall, 2 mm is generally considered as a threshold to define a rainy day and have been defined based on the mean daily potential evaporation (Marino et al., 2020; Peres et al., 2018).

Although the threshold definition of a rainy day (10 mm) may have led to the omission of some rainfall information and thus shortening the event duration, this approach improved the similarities between the satellite-based and gauge-based landslide hazard assessment thresholds, and thus used as a bias correction between the two sources of rainfall data. Similarly, other bias correction methods have been adopted by other researchers to ensure for the high accuracy between ground- and satellite-based rainfall data (Bhatti et al., 2016; Vernimmen et al., 2012). However, the defined satellite-based event/duration E/D thresholds 16–17.5 mm d$^{-1}$ were quite higher than previously defined gauge based-thresholds ~7–13 mm d$^{-1}$. Contrarily the defined thresholds from the recent cumulative 2 and 3 day rainfall were much smaller than defined from gauge based data (Uwihirwe et al., 2020, 2021). These differences are probably due to the predefined threshold (10 mm) that probably omits some rainy days. This also led to shortened event duration and hence slightly higher E/D. Nevertheless, the landslide triggering conditions defined based on the E/D reveals poor discriminatory power to distinguish landslide from no landslides (AUC~0.53) and thus not impactful on landslide initiation. The linked landslide thresholds also underperform in terms of landslide prediction capabilities measured by the resulting low rate of true positives TPR~56-69 %. Similarly, the landslide hydro-meteorological thresholds that included the rainfall event E/D as a trigger resulted into poor landslide warning performance TPR max ~66 %.

It is agreed that the consequences of offering false alarms (FPR) are less harmful on the short-term than missed alarms (FNR) which implies that the best threshold should maximize the TPR (true alarms) while minimizing the FNR (missed alarms). However, the thresholds in Fig. 7b and 7e are classical thresholds E-D relying exclusively on rainfall (trigger), leading to the high rate of missed alarms and thus less important for a robust LEWS development. Similar to this study, previous studies (Bogaard and Greco, 2018; Marino et al., 2020; Minus et al., 2018; Peres et al., 2018; Thomas et al., 2019; Uwihirwe et al., 2020, 2021).
indicated that the consideration of the prior subsurface hydrological conditions reduce the number of missed alarms FNR as well as the number of false alarms FPR relative to the exclusive use of rainfall-only thresholds. In Fig. 8a, 8b and 8c, we integrated the hydrological information (i.e. antecedent soil moisture) in landslide thresholds to improve the rate of TPR and reduce the rate of FNR and FPR. The main goal of hydro-meteorological thresholds (cause-trigger) is to maximize the true alarms TPR i.e. minimize the missed alarms FNR but at the same time reducing the false alarms FPR. The used statistical metrics (TSS and RAD) are also in line with this concept aiming at maximizing the TPR while minimizing the FPR. Once TPR is maximized, the FNR is also minimized though difficult and or impossible to have a perfect threshold model with zero FNR and FPR.

The poor performance of the rainfall event-based thresholds concept is due to uncertainties from multiple sources. We hypothetically used the rainfall events as landslide triggering conditions, defined as individual periods of continuous rain interrupted by at least two dry day periods referred to as the minimum inter-event time (IET). Nevertheless, this definition needs further exploration to be standardised to avoid uncertainties. According to Adams et al. (1987); Hong et al. (2017), the IET is defined as the minimum period of time that separates two consecutive rainfall events and is considered as the period for which the effects of the antecedent soil moisture or precipitation index may last. This is to say that the antecedent soil moisture and or antecedent precipitation index have no significant effect on landslide initiations once the rainfall events and IETs are well defined. However, the IET, the period during which the effect of antecedent soil moisture becomes null, depends on a number of site-specific factors (soil properties, land use/land cover, potential evaporation etc.) and is thus difficult to be standardized. Another drawback associated with the use of rainfall event concept may be linked to the transient timescales of the triggering events that bring about difficulties to fix the appropriate time to give an alert or an early landslide warning to the threatened community.

Beholding the constraints associated with IET, rainy day and rainfall events definition, we explored the shorter scaled triggering rainfall conditions that include the cumulated rainfall with constant duration 1, 2, 3 days (RD1, RD2, RD3). The cumulative 3 day rainfall RD3 showed the highest impact on landslide initiation AUC ~0.72 and true positive alarms TPR~79–81 %. Although the meteorological trigger-based thresholds RD3, have resulted into high rate of true alarms, they lack the concrete physical significance and are also challenging for a robust landslide early warning system due to the linked high level of erroneous alarms i.e false positives FPR~40–42 %. To account for the pre-wetting state of the soil, the antecedent soil moisture conditions have been considered. These antecedent soil moisture conditions from the top 5 cm, 50 cm and 2 m, Stop, Sroot, Sroot respectively showed significant impact on landslide predisposal AUC=0.71–0.76. Moreover, with exception to the Sroot, the hydrological landslide thresholds 0.56 [Stop], 0.73 [Sroot] defined from these soil moisture conditions revealed high landslide warning capability with true alarms TPR ~75–85 %. These hydrological thresholds indicate the critical pre-wetting state above which any additional amount of rainfall > 11-15mm is highly likely to trigger landslides. We therefore combined
both landslide hydrological predisposing and meteorological triggering conditions following the cause-trigger concept into bilinear hydro-meteorological thresholds framework. This approach improved the landslide prediction capabilities in terms of reduced rate of false alarms (FPR~22 %) and increased true alarms (TPR~72 %) as compared to the approaches that consider the maximum probable rainfall event (max TPR~66 % and FPR~41 %). In other words, once combined with the pre-wetting hydrological conditions, the cumulative few days rainfall have significant impact on landslide initiation and warning as compared to the longer and no constant triggering conditions. Furthermore, the incorporation of the antecedent wetness state of the terrain not only improved the landslide warning capabilities but also provide accurate insights into landslide alert time as compared to the use of transient time scale associated with the rainfall event concept.

Among the tested pre-wetting conditions, the incorporation of the antecedent soil moisture modelled at the root zone $S_{\text{root}}$ was the most impactful for landslide initiation and thus the most useful in landslide hazard assessment thresholds in Rwanda. The finer spatial resolution of the hydrological model-derived soil moisture together with the consideration of the specific climate and hydrogeological characteristics of the model catchments could be a possible explanation of the positive impact of soil moisture assimilated at the root zone. This could also be explained by the less exposure of the root zone to the solar heat and evaporation processes as compared to the near surface $S_{\text{top}}$. The probable less prone to the gravity driven drainage and deep percolation due to the soil texture, vegetation and organic matter at the root zone could also be an explanation. Moreover, the soil depth involved in shallow (0.5–2 m) and deep landslides (>2 m) (Greco et al., 2018) is much thicker than $S_{\text{top}}$ (5cm) currently measured by the satellite based soil moisture technologies and this is more captured by the hydrological modelling approaches (Wflow). An overestimation of soil moisture by satellite (VdS) and the distributed hydrological model (Wflow) was also noted and attributed to the similar overestimation of satellite-based precipitation, an important element in soil moisture estimation. Therefore, more reliable algorithms that addresses the reliance between the satellite and in situ based information could thus improve the performance and enhance data accuracy needed for landslide hazard assessment.

The adopted bilinear threshold framework, indicating the distribution of data points in a 2D space, reflects the relationship between the landslide causal and triggering conditions. The adopted bilinear threshold framework, indicating the distribution of data points in a 2D space, reflects the relationship between the landslide causal and triggering conditions despite other linked constraints and limitations (Conrad et al., 2021). We objectively used the bilinear thresholds framework because the majority of positive classes were clustered in upper right corner of the 2D threshold space. Although, this format proved to be suitable for landslide hydro-meteorological thresholds definition (Mirus et al., 2018a; Thomas et al., 2019; Uwihirwe et al., 2020, 2021), it is clear that for meteorological threshold model E-D, the commonly used power law function would have resulted into high rate of true alarms at the expense of a very high rate of false alarms. However, the other formats could also be useful depending on the distribution of the positive classes in the 2D space. The adopted bilinear framework is in line with the goal of the hydro-meteorological cause-trigger based thresholds concept that prioritize the minimization of false alarms while at
least keeping unchanged the rate of true alarms. Additionally, in some cases, single variable thresholds lead to high prediction capabilities in terms of elevated rate of true alarms and with quite low rate of false alarms and could be adopted especially for hydrologically based thresholds that consider the long-term wetting process of the soil until the landslide day.

Regardless of the good performance of soil moisture as landslide hydro-meteorological threshold, the incorporation of pre-wetting state of soil in landslide hazard assessment thresholds using groundwater levels, $h_{(t-1)}-E/D$ (TPR=54–64% and FPR=6–11%) (Uwihirwe et al., 2021) with low rate of false alarms, performed higher than using root zone soil moisture $S_{e(t-1)}-E/D$ (TPR=66% and FPR=44%) due to the elevated rate of false alarms.

Despite the need to have relatively large samples (about 200 landslide events) to have precise estimation of threshold parameters (Peres and Cancelliere, 2021), the landslide inventory used for this study counts for only 32 hazardous landslides. Although, the reliance on this limited sample size is likely to lead to a bias towards larger landslide events and those with impact to society, this landslide inventory is the most comprehensive currently available in the study area.

Ideally, one would have a landslide inventory of about 200 landslides events in order to have a precise estimation of threshold parameters (Peres and Cancelliere, 2021). However, the landslide inventory used for this study counts for only 32 hazardous landslides. Although, the reliance on this limited sample size is likely to lead to a bias towards the larger landslide events and those with impact to society, this landslide inventory is the most comprehensive currently available in the study area.

5 Conclusion

This research aimed to evaluate the potential of satellite-based measurements of precipitation and soil moisture as well as hydrological model-derived soil moisture information for landslide initiation thresholds in Rwanda. The GPM-based IMERG rainfall product was found to be a good spatially distributed source of rainfall data for landslide hazard assessment especially in data scarce areas like Rwanda. The satellite and model-derived soil moisture time series broadly reproduce the most important trends of the in situ soil moisture. Regardless of different depths of data records and slightly overestimation of soil moisture by satellite and model-derived techniques, it was concluded that they follow the in situ observed temporal variation and are thus potentially useful for landslide hazard assessment initiation thresholds definition. The purpose of incorporating the antecedent soil moisture in landslide hazard assessment was to account for the physical effect of the pre-wetness state of soil, responsible for the predisposal of the slopes to near failure, prior to the landslide triggering conditions. Two categories of landslide triggering conditions have been considered to assess the potential value of including the antecedent soil moisture information. The category that considers the cumulative 3 day rainfall was the most impactful and thus more useful for landslide hazard assessment in Rwanda rather than the rainfall event-based trigger. Although the area under the curve (AUC=0.71–0.76) statistical metric indicated the significant impact of all tested antecedent soil moisture variables prior to the
triggering conditions, the antecedent soil moisture modelled from the root zone $S_{\text{root}}$ performed best. Similarly, The classical thresholds E-D relying exclusively on rainfall (trigger) performed lower with high rate of missed alarms (50%) and thus less important for a robust early warning system development. Contrary, the hydro-meteorological thresholds that incorporate the antecedent soil moisture $S_{\text{root}}$ and the recent 3-day cumulative rainfall $R_{\text{D3}}[S_{\text{root}}(t-3)-R_{\text{D3}}]$ outperforms other threshold models with high rate of true alarms (72%) and low rate of false alarms (20%) and thus can be very useful for landslide hazard assessment and early warning system development in Rwanda.

Data availability. The landslide inventory used for this research can be accessed at https://doi.org/10.4121/15040446.v1

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

Author contributions. JU collected the in situ data, conducted the statistical analysis, conceptualise and prepared the manuscript storyline. ARP collected and prepared data from satellite precipitation products, organised the soil moisture data and conducted some statistical analysis. HW and JS prepared and availed the high resolution satellite-based soil moisture data. FSW actively participated in hydrological modelling of soil moisture, highly contributed to the correction and perfection of the manuscript. MS corrected the manuscript storyline, shaped the discussion and contributed to the perfection of the manuscript. TAB initiated the research idea, proposed the research approaches, created the research network and collaboration, verified the data preparation and statistical analysis, shaped the manuscript storyline and contributed to the perfection of the manuscript.

Acknowledgements. We are so thankful to Dr. Elise Monsieur, Dr. Arthur Depicker and Dr. Olivier Dewitte for sharing part of the Landslide Inventory for Rwanda as part of the central section of the Western branch of the East African Rift (LIWEAR) project. We are thankful to the Rwanda meteorological agency for offering access to the in situ meteorological and hydrological datasets used for validation of other source of data.

References


Appendix A
Figure A1. Selected examples of satellite and model-derived soil moisture compared to in situ recorded soil moisture at 20 cm soil depth (AWS): a) $\theta_{top}$ and in situ $\theta_{20cm}$ soil moisture time series at Gacurubwenge station; b) model-derived soil moisture in the root zone $\theta_{root}$ and in situ soil moisture $\theta_{20cm}$ at Kibisabo station; c) satellite derived $\theta_{top}$ and model-derived soil moisture in the root zone $\theta_{root}$ with vertical red lines indicate the timing of the landslide occurrence time.