Insights from the topographic characteristics of a large global catalog of rainfall-induced landslide event inventories

Robert Emberson[1,2,3], Dalia B. Kirschbaum[1], Pukar Amatya[1,2,3], Hakan Tanyas[4], Odin Marc[5]

[1]NASA Goddard Space Flight Center, Hydrological Sciences Laboratory, United States of America
[2]Universities Space Research Association, Columbia, MD, USA
[3]Goddard Earth Sciences Technology and Research, Columbia, MD, USA.
[4]ITC, University of Twente, Netherlands
[5]Géosciences Environnement Toulouse (GET), UMR 5563, CNRS/IRD/CNES/UPS,Observatoire Midi-Pyrénées, Toulouse, France
Correspondence to: Robert Emberson (robert.a.emberson@nasa.gov)

Abstract. Landslides are a key hazard in high-relief areas around the world and pose a risk to population and infrastructure. It is important to understand where landslides are likely to occur in the landscape to inform local analyses of exposure and potential impacts. Large triggering events such as earthquakes or major rain storms often cause hundreds or thousands of landslides, and mapping the landslide populations generated by these events can provide extensive datasets of landslide locations. Previous work has explored the characteristic locations of landslides triggered by seismic shaking, but rainfall induced landslides are likely to occur in different parts of a given landscape when compared to seismically induced failures. Here we show measurements of a range of topographic parameters associated with rainfall-induced landslides inventories, including a number of previously unpublished inventories which we also present here. We find that average upstream angle and compound topographic index are strong predictors of landslide ~~headscarp~~scar location, while local relief and topographic position index provide a stronger sense of where landslide material may end up (and thus where hazard may be highest). By providing a large compilation of inventory data for open use by the landslide community, we suggest that this work could be useful for other regional and global landslide modelling studies and local calibration of landslide susceptibility assessment, as well as hazard mitigation studies.

1. Introduction

The impact of natural hazards on population and infrastructure is most acute where the footprints of these hazards intersect the locations where people live and buildings are situated. For some hazards like earthquakes and cyclones, the footprints of the hazard can be distributed across wide regions, but for other hazards like landslides the footprint may be significantly more localized. Although the impacts of individual landslides may be localized, large triggering events such as intense rainfall or seismic activity can cause large numbers of landslides across a wide region~~Landslides triggered by intense rainfall or seismic activity, however, can be distributed more broadly~~, the extent of which often mirrors the extent of the intense rainfall and seismic shaking [Marc et al.,2017, 2018, Tanyas and Lombardo 2019]. The individual landslides triggered during these extreme events occur in specific parts of the landscape that are most susceptible to failure. These slopes become critically

unstable due to both preconditioning factors like slope and internal frictional strength, as well as triggering factors like change in fluid pore pressure or seismic acceleration.

A range of studies from around the world have assessed the locations of landslides and used them to construct susceptibility models for local settings [e.g., Emberson et al. 2021, Goetz et al. 2015, Broeckx et al. 2019], across larger regions [e.g., Van Den Eeckhaut & Hervas, 2012, Van Den Eeckhaut et al. 2012,], and globally [e.g., Stanley et al. 2017, Nowicki Jesse et al. 2018, Tanyas et al. 2019]. Comprehensive reviews of landslide susceptibility models [Budimir et al. 2015, Reichenbach et al. 2018] highlight a number of factors that are often considered to be generally important for landslide susceptibility. These include morphological (slope, aspect, roughness), geological (e.g. lithology), land cover, seismic and hydrological factors. Naturally, to study the importance of each of these factors, information on landslide location is essential to both calibrate and validate any susceptibility model that is produced.

Landslide location data can come in different forms and landslide inventory maps are the most useful data source in which the extent of landslide phenomena are systematically documented in a region [Guzzetti et al. 2012]. Unfortunately, the number of digitally available landslide inventories is still rather limited [Wasowski et al. 2011, Guzzetti et al. 2012, Tanyas et al. 2017, Mirus et al. 2020]. As a result, landslide locations in global catalogs are often based on media reports [e.g. Kirschbaum et al. 2015, Froude and Petley 2018], which can limit the accuracy of the defined locations. A review of data in the NASA Global Landslide Catalog [Kirschbaum et al. 2015] suggests that only 33% of landslides have a location known to within 1km resolution, which does not permit assessment of the specific locations where landslides occur within a landscape (e.g. at a hillslope scale). Additionally, global landslide catalogs generally do not include the entire landslide population for a given area. While they may capture many of the landslides that cause damage or fatalities [Petley 2012, Froude & Petley 2018], underestimation of landslide susceptibility may result if systematic biases in reporting are found for certain geographies or terrain parameters.

Landslide inventories are the ideal data source not only to better understand the spatial, temporal and size distribution of landslides but also to conduct more accurate susceptibility, hazard and risk assessments [Guzzetti et al. 2012]. Overall, landslide inventories are categorized as historical and event inventories [Malamud et al. 2004]. Historical landslide inventories include many landslide events over time in a given region. Landslide-event inventories, on the other hand, contain landslides triggered by a specific trigger (e.g., earthquake, rainfall or snowmelt) of known date. In other words, the time of landslide occurrence is unknown in historical landslide inventories and therefore~~, for instance~~, landslide susceptibility models developed based on historical inventories are time-invariant products solely representing geomorphologically landslide-prone hillslopes [Lombardo and Tanyas 2020]. Historic inventories are by definition biased toward frequent, climatic triggers, and are not representative of the long-term average susceptibility to triggers including earthquakes. Whereas, landslide-event inventories are more suitable data sources to develop near-real time products to predict spatial distribution of landslides triggered by a specific event [e.g., Nowicki Jessee et al. 2018].

For specific, large triggering events such as an earthquake or an episode of extreme rainfall, it is possible to relatively accurately define the timing of the event, and if high resolution imagery is found that brackets the dates in question it is also

possible to systematically map the landslides generated by such a trigger [Guzzetti et al. 2012]. Mapping landslides following extreme events has become common, and inventories exist for a large number of earthquakes [Tanyas et al. 2017]. A smaller number of intense rainfall events have also been mapped [Marc et al., 2018], but unlike for earthquakes [Schmitt et al. 2017] no centralized repository of these data exists at present. Location data for landslides triggered by intense rainfall is vitally important to calibrate and validate existing susceptibility models, since the datasets produced are generally considered to be nearly complete. It is also useful to characterize the rainfall required to trigger landslides, and thus help inform local and global hazard models [Emberson et al. 2021, Kirschbaum et al. 2018]. These can then be used to inform exposure and risk assessment estimates [Emberson et al. 2020].

It is important to note that the positions where earthquake-triggered landslides occur on a given hillslope are not necessarily applicable to rainfall-triggered landslides. As shown by previous research [Densmore and Hovius 2000, Meunier et al. 2008], the higher peak ground acceleration in earthquakes at the top of ridges tends to increase landslides in those locations, while increasing water saturation at the base of slopes by intense rain tends to increase landslides lower down the slope [e.g., Rault et al. 2018]. As such, it is imperative to use the appropriate type of landslide inventory to calibrate any model. Finally, recent studies have sought to derive underlying simple topographic rules to understand hazard associated with earthquake-triggered landslides [e.g. Milledge et al. 2019], and it is important that we extend this kind of analysis to rainfall triggered events to provide comparative data.

In this study, we combine 10 existing inventories of landslides triggered by intense rain storms with 6 new inventories mapped using high resolution data for this study. Assessing these landslide-event inventories both individually and in combination, we assess the local topographic characteristics that are most strongly related to where landslides initiate, as well as local forest loss that can be calculated from satellite data. We suggest that these inventories and the associated parameters can be used to calibrate and validate other models of susceptibility and hazard, and will provide valuable information to authors seeking landslide data with high spatial accuracy, as well as supporting characterization of rainfall thresholds for landslide impacts [e.g. Conrad et al. 2021]. Moreover, with a set of simplified rules for landslide hazard, researchers can support hazard assessment in areas where more detailed models may be unavailable. The~~se~~ inventories described here will be available on the NASA Landslide Viewer app (landslides.nasa.gov/viewer) for open access by other researchers.

2. Context & Methods

2.1. Landslide Inventories

It has become common practice to map areas affected by landslide-triggering earthquakes to build a spatially complete picture of landslide impacts [Tanyas et al. 2017], and the inventories that are generated have been used to produce hazard maps [Jibson et al. 2000, Harp et al. 2011], susceptibility models [Garcia-Rodriguez et al. 2008, Xu et al. 2012], guidelines for hazard zonation [Milledge et al. 2019] and global alerting systems [Nowicki Jessee et al. 2018]. Landslide-event inventories are also required to explore the landscape response to tectonic and climatic forcings [e.g., Malamud et al. 2004, Korup et al. 2012, Marc et al., 2016, 2019]. Mapping of landslides in the aftermath of major rainfall events is somewhat less common, since cloud

cover is often a significant impediment in the impacted areas, which may limit clear views from satellites. However, an increasing number of intense rainfall events have now had landslides mapped, with extensive examples in Taiwan [Lin et al. 2011, Chen et al. 2013], Japan and Brazil [Marc et al. 2018] and the Caribbean [van Westen & Zhang 2018].

Several methods exist to generate event-specific landslide inventories. The robustness and accuracy of the final inventory depends on the type and quality of imagery and data available, as well as the method chosen. Synthetic Aperture Radar (SAR) data has been employed to generate inventories of slow-moving landslides [Handwerger et al. 2019, Bekaert et al. 2020], to focus on the kinematics of single slow-moving slides [Hu et al. 2019] and has been used to map landslides occurring in the aftermath of major triggering events [Mondini 2017, Handwerger et al. 2019, Adriano et al. 2020, Burrows et al., 2020, Jung & Yun 2020]. The most widely used technique is to map landslides directly from optical imagery, either from Unmanned Aerial Vehicle (UAV) imagery [Casagli et al. 2017, Rossi et al. 2018], aerial photography [Harp et al. 2004] or satellite observations [Casagli et al. 2017, Martha et al. 2012, Behling et al. 2014]. While satellite observations generally have the lowest spatial resolution and may be impinged by cloud cover, these satellites offer near global coverage and frequent return intervals that generally allows for imagery that brackets the event in question. This is particularly the case for some of the newer commercial satellite constellations. Some rainfall triggered events may occur in locations where cloud cover is so prevalent that it precludes anything other than seasonal assessment of landslide occurrence, such as the Himalayas during the monsoon. However, an increasing number of satellite-generated inventories now exist. The methods used to delineate landslides from optical imagery include manual mapping, where a human determines what is and is not a landslide, or semi-automatic/automatic mapping, where detection algorithms are used to determine landslide locations.

2.2. Methodology

Summarising various previous work, five mapping criteria appear essential for landslide inventories (see Guzzetti et al 2012, Marc & Hovius 2015, Tanyas et al., 2017): (i) manual mapping (or correction) to reduce errors and avoid amalgamation, (ii) high enough imagery resolution for completeness and to avoid amalgamation, (iii) mapping landslides as polygons to allow maximum scientific usage (e.g., area affected, volume of sediment mobilized, frequency-size distributions), (iv) mapping with pre and post-event imagery to focus on landslides with a known trigger, and (v) defined mapping boundary to clarify inventory completeness. For the purposes of this study, we have tried to obtain as many inventories as possible for comparison, while generally satisfying these five essential criteria. Nevertheless, due to varying imagery and mapping techniques, criteria (i) and (ii) are fulfilled with variable quality for the studied inventories (Table 1). More detailed inventories have differentiated the source and deposit areas of landslides, but this often requires field validation. The locations of the inventories are shown in Figure 1. It is important to note that although high resolution imagery can provide more accurate mapping in some cases, it can also be more challenging to ortho-rectify, which can limit the quality of landslide inventories generated [Williams et al. 2018]

As such, we incorporate 10 existing inventories and supplement them with 6 further inventories that we have produced for this study. The details for each of the inventories are described in Table 1. For several of the newly produced inventories, we have

utilized high resolution imagery from Planet Dove satellites [Planet Team, 2017] available through the Commercial Smallsat Data Acquisition (CSDA) Program (https://earthdata.nasa.gov/esds/csdap). Planet imagery represents an important step change for landslide mapping procedures, since the high spatial resolution of approximately 3m of the images is combined with rapid return time for the satellites (of the order of 1 image per day).

For several of the newly generated inventories, we have mapped the landslides manually using GIS software. This resulted in three inventories; two in the Philippines, and one in Thrissur, India (Table 1). The remaining new inventories were generated using the semi-automatic object-based methods of Amatya et al [2019 and 2021]. The algorithmic method was used to reduce the overall time spent mapping some of the larger new inventories. Since algorithmic methods are known to produce artifacts when broadly applied [Pawluszek et al. 2018] and can lead to amalgamation of individual landslides into larger polygons [Marc and Hovius 2015], each of the automatically generated inventories was additionally corrected by manual comparison with pre- and post-event high-resolution imagery in GIS software. ~~We therefore consider each of the newly generated inventories to be of comparable quality to one another~~.

Table 1: Details of landslide inventories analysed in this study.

Location	**Event triggering**	**Date of triggering rainfall**	**Reference**	**Imagery (resolution)**	**Dominant GLiM lithology**
Micronesia (A)	Cyclone	2nd July 2002	Harp et al. 2004	Aerial Photographs (varies)	Basaltic / andesitic / trachytic lava flows
South Taiwan (B)	Typhoon Kalmaegi	15-18 July 2008	Chen et al. 2013, Marc et al. 2018	Landsat 5 (30m)	Sedimentary / metamorphic
Blumenau, Brazil (C)	Prolonged intense rain	20-25 Nov 2008	Marc et al. 2018	Google Earth / Landsat 5 (3-30m / 30m)	Metamorphic
Taiwan (D)	Typhoon Morakot	6-9 Aug 2008	Chen et al. 2013, Chang et al. 2014, Marc et al. 2018	FORMOSAT-2 / Landsat 5 (4m / 30m)	Sedimentary / metamorphic
Teresopolis, Brazil (E)	Local Storm	11-13 Jan 2011	Marc et al. 2018	Google Earth / Landsat 5 / EO-ALI (3-30m / 30m / 30m)	Acid Plutonic / metamorphic
Kii Province, Japan (F)	Typhoon Talas	2-5 Sept 2011	Marc et al. 2018	Aerial photographs/ Google Earth / Landsat 5 (Varies / 3-30m / 30m)	Sedimentary / minor plutonic / metamorphic
Salgar, Colombia (G)	Local Storm	17-18 May 2015	Marc et al. 2018	Sentinel 2 / Google Earth (10m / 3-30m)	Metamorphic
Hiroshima, Japan (H)	Prolonged intense rain	28 Jun - 9 July 2018	Association of Japanese Geographers, 2019	Drone / Aerial imagery (varies)	Plutonic Acidic / Volcanic Acidic / minor siliciclastic
Zimbabwe (I)	Cyclone Idai	15-19 Mar 2019	This study	Planet Dove (3m)	Metamorphic / minor siliciclastic
Itogon, Philippines (J)	Cyclone Mangkhut	15-20 Sept 2018	This study	Planet Dove (3m)	Mixed sedimentary / minor acidic volcanic
Lanao del Norte, Philippines (K)	Tropical Storm Tembin	20-26 Dec 2018	This study	Planet Dove (3m)	Volcanic acidic
Dominica (L)	Tropical Storm Erika	25-28 Aug 2015	van Westen 2016	WorldView 3 (1.8m)	Intermediate Volcanic

Dominica (M)	Hurricane Maria	18-22 Sept 2017	van Westen & Zhang 2018	Pleiades (0.5m)	Intermediate Volcanic
Burundi (N)	Prolonged intense rain	3-5 Dec 2019	This study	Sentinel 2 (10m)	Metamorphic
Thrissur, India (O)	Prolonged intense rain	7-18 Aug 2018	This study	Planet Dove (3m)	Metamorphic / Plutonic acidic
West Pokot, Kenya (P)	Prolonged intense rain	22-25 Nov 2019	This study	Sentinel 2 (10m)	Metamorphic

Beyond the five essential mapping criteria, additional criteria include the differentiation of scar and deposit and the classifications of landslides according to their type/mechanisms. However, these criteria are difficult to fulfill for large event inventories (see in Tanyas et al 2017), especially when based on various sources of optical imagery limiting our ability to differentiate between scar and deposit areas [Casagli et al. 2017].

The mapped inventories combine scars and deposit in the polygon delineation, although in the analysis discussed below we have sought to differentiate these areas. In terms of landslide type we could not systematically classify each landslide polygon. However, we have removed debris flows from the analysis where possible by removing long-runout landslide polygons from each mapped inventory. In general, this mapping identifies rockslides, rock avalanches, shallow soil toppling and slumping failures, but does not capture slow moving landslides where surface changes may be less evident. A focus on these kinds of landslides is warranted since the volume of material mobilized during large storms from such landslides can lead to damaging debris flows and bedload transport impacts [Badoux et al. 2014]. Removing debris flows from the analysis allows us to provide consistent landslide maps that can be used to estimate volumes of mobilised landslide material, for example using global scaling relationships like those defined by Larsen et al. [2010], and permits a focus solely on the topographic characteristics of landslide source regions, rather than the characteristics of preferential runout paths.

We contrast each of the inventories mapped here by comparing the size-frequency distributions of each dataset, shown in Figure 2. For each of the inventories, we show the probability of a landslide within a given area interval, as a way to assess the frequency of small and large landslides across the different datasets.

Each of these landslide events was triggered by extreme rainfall, and although it is not our intention to examine the triggering rainfall in detail in this study it is useful to briefly discuss the characteristics of the rainfall events in question. It is important to note that the date of the triggering rainfall is not identical to the dates the imagery used to map the landslides was obtained. Although we have selected events where the triggering rainfall significantly exceeds historical peak rainfall (and therefore is likely to be the dominant trigger for landslides) some events may have occurred as a result of lesser rainfall before or after. While the new inventories generated for this study that utilise Planet imagery that closely brackets the rainfall events (within 1 week either side), the older inventories may be more subject to this challenge. A detailed analysis of the triggering rainfall associated with several of these inventories is described by Marc et al. [2018], who used local gauge data to characterize the

rainfall intensities. We were unable to find consistent local gauge data for several of the more recent events that are published here for the first time (events in Zimbabwe, Burundi, Kenya, and the two events in the Philippines). We can still use satellite rainfall data as a consistent source of rainfall for each of the events, however. To assess these, we utilize the reprocessed IMERG (Integrated Multi-satellitE Retrievals for GPM) version 6B rainfall product (Huffman et al. 2020), which merges and homogenizes data from NASA's Global Precipitation Measurement (GPM) mission with its predecessor Tropical Rainfall Measurement Mission (TRMM). All of the events considered occurred within the period during which GPM IMERG v06B rainfall data is available (2001-present). Because the satellite rainfall data spatial resolution is relatively coarse, it is not possible to effectively draw comparisons between the landslide polygons and the surrounding data in the same manner as the topographic data. However, we can still characterize the rainfall occurring during each event. We have analysed the total rainfall occurring during each of the events by accumulating the rainfall data over the period of each event indicated in Table 1, and compared this with the calculated historical 99th percentile of daily rainfall as a way to normalize each event to the historical trends. The 99th percentile is calculated empirically based on the GPM IMERG v06B record (2001-2020). Since the length of rainfall period associated with each inventory varies, normalizing by the 99th percentile for a single day provides a consistent normalizing factor for each inventory. Additionally, in Table 2 we show the maximum 3-hour rainfall intensity for each of the events, normalized by the historical 99th percentile of daily rainfall. The normalized total event rainfall and normalized 3-hour rainfall provide a side-by-side comparison of the overall rainfall accumulation and the maximum intensity. The values for both total event rainfall and maximum 3-hour rainfall are calculated as the average across all IMERG grid cells in the area of the inventory. Table 2 summarizes information on the landslide inventory characteristics including the total landslide area, the density of landslides in the mapped area, satellite rainfall, average slope in the mapped area. Despite other studies suggesting links between event total rainfall and the density of landsliding [Chen et al. 2013, Marc et al. 2018], we do not observe clear links between the measured rainfall data and the macro-scale statistics of each landslide inventory. Relations between landslide density and rainfall can be obscured by variations of climatic and/or hydromechanical properties with each study area [Marc et al.2019]. We suggest that exploring the links between rainfall intensity as characterized by satellite measurements and the density of landsliding that results is an important topic for future research.

Table 2: Macro-level characteristics for events discussed, including rainfall statistics. Note that median slope values have been calculated by excluding very low slope values (<1 degree) to remove lakes and oceans.

Event Name	Number of landslides	Total Area (km^2)	Density of landslides (m^2/m^2)	Median slope (degrees)	Total Event rainfall (mm)	Std. Deviation in Event rainfall	Total Event rainfall / Historical 99th percentile	Std. deviation in event rainfall/ historical 99th pctl.	Maximum 3hr rain / Historical 24hr 99th percentile	Std. dev 3hr rain / Historical daily 99th percentile
Micronesia (A)	273	1.949	3.62E-02	12.0	1426.2	85.6	24.7	1.5	11.2	0.4

South Taiwan (B)	429	3.650	3.33E-04	25.5	517.1	103.7	6.6	1.7	2.5	1.1
Blumenau, Brazil (C)	597	5.847	2.11E-03	15.3	72.8	37.0	1.5	0.7	0.3	0.35
Taiwan (D)	10236	205.048	7.43E-03	23.7	1114.0	183.7	14.6	4.0	2.1	1.42
Teresopolis, Brazil (E)	7268	21.560	7.76E-03	20.8	193.2	54.3	5.4	1.6	1.6	1.0
Kii Province, Japan (F)	1901	12.258	1.51E-03	25.4	284.5	46.2	4.3	0.5	1.3	0.6
Salgar, Colombia (G)	131	0.283	4.75E-03	27.44	112.2	34.0	3.5	1.3	1.4	1.8
Hiroshima, Japan (H)	9275	4.542	1.06E-03	14.1	488.2	62.0	10.6	1.1	4.1	0.9
Zimbabwe (I)	1319	2.554	1.62E-03	13.3	321.0	30.0	8.7	0.7	3.6	0.9
Itogon, Philippines (J)	458	0.627	1.04E-03	26.1	179.3	12.5	4.0	0.5	1.5	0.7
Lanao del Norte, Philippines (K)	17	0.195	8.36E-04	16.3	150.6	22.7	5.6	0.9	2.2	1.5
Dominica (L)	1756	10.450	2.48E-03	18.2	169.6	16.2	9.9	1.7	5.2	0.9
Dominica (M)	21379	10.251	1.36E-02	18.2	73.0	11.5	4.0	0.8	2.9	0.8
Burundi (N)	492	1.976	1.12E-02	21.2	54.3	19.0	2.7	1.0	0.6	0.7
Thrissur, India (O)	188	1.130	6.02E-04	14.6	475.0	59.1	10.3	1.9	1.4	0.3
West Pokot, Kenya (P)	338	1.346	4.54E-03	22.4	99.4	7.1	5.0	0.9	2.7	0.6

2.3. Topographic Analysis

We have analyzed the topographic characteristics of landslide locations for the event inventories, using ~~with~~ global satellite datasets to ensure consistency across each site. These datasets are also openly available, which supports replication of these methods and findings by other authors. In Table 3, we show the datasets we have used.

Table 3: Analysis datasets. Explanation of each of the variables in found in the accompanying text.

Dataset	Source / Reference	Parameters from Dataset
SRTM DEM (1-arc	SRTM Non-Void Filled https://doi.org/	Relief (1km radius) Slope

second resolution)	10.5066/F7K072R7	Average upstream angle
		Compound Topographic Index, CTI [Beven & Kirkby 1979, Sorensen et al. 2006]
		Topographic Ruggedness Index, TRI [Riley et al. 1999]
		Topographic Position Index, TPI300 – 300m wavelength [Weiss, 2001]
		~~Topographic Position Index, TPI2000 – 2000m wavelength [Weiss, 2001]~~
Forest Cover (1 arc-second resolution)	Global Forest Change 2000–2018 [Hansen et al. 2013]	Forest loss since 2000
		Forest cover

The DEM and forest loss data are both provided at approximately 1 arc-second resolution, which means we do not have to resample either dataset when conducting a raster-based analysis at this scale~~allows for direct comparison at this scale~~. While this resolution is not as fine as some of the imagery used to map the landslides, which can be 3m or finer, it represents the finest resolution at which these two datasets can be analysed using non-commercial, open datasets at a global extent. We utilize forest loss data derived from Landsat imagery spanning the years 2000-2018. Cells where forest loss is observed on any year from 2000 until the year in which the landslide event occurred are considered as a binary 'True' value for forest loss. This does not consider regrowth of vegetation in places where forest loss was observed many years prior to the event, and as such it is a relatively blunt tool to assess the importance of vegetation to landslide location.

Not all of the topographic parameters are universally used in landslide analysis. Slope is almost universally considered for landslide modelling, but the use of others (in particular Topographic Position Index) is less common [Reichenbach et al. 2018]. Topographic Position Index (TPI) [Weiss 2001] is a quantification of the relative position of a cell within the landscape. It is calculated as the difference in elevation of each cell in a DEM to the mean elevation of a specified neighbourhood around that cell, with the radius of the neighbourhood chosen beforehand (in this case, 300m). Negative values indicate the cell is in a topographic hollow and positive values suggest that it is elevated above its surroundings. The distance over which the neighbourhood comparison is made (TPI wavelength) determines the scale of the features resolved; negative values at long wavelengths indicate a position in a wider valley, while at short wavelengths this would indicate steep narrow gorges. In this study, we focus on short wavelength TPI values since this aligns more closely with the scale of the landslide features. Relief indicates the difference between minimum and maximum elevation in a given window. It is a proxy for both slope as well as the size of hillslopes; higher relief zones have been shown to be associated with landslides in many locations [Reichenbach et al. 2018]. Compound Topographic Index (CTI) is a measure of both slope and the upstream contributing area. It is calculated by the formula ln(a/tan b), where a is the flow accumulation area per pixel and b is the local slope in radians. In some locations CTI is correlated with soil parameters such as thickness [Liang & Chan 2017]. We also calculate Topographic Ruggedness Index (TRI), a measure of the local surface roughness. It is defined as the root mean squared difference in elevation between

a central pixel and each of its eight neighbouring pixels.

Finally, we also analyse the average upstream angle – this is the average angle from the pixel location to every cell that drains into that pixel. It provides a measure of how steep the areas are that feed into each pixel. There is a significant degree of overlap between how some of these parameters are calculated, and we recognize the importance of considering co-linearity.

In order to assess the co-linearity of the variables, we have compared each pair of variables. Pair-plots are shown in the supplementary material (Supplementary figures S1 and S2). Unsurprisingly, strong correlations are observed between slope and average upstream angle, as well as between Topographic Ruggedness and relief. It is also important to note the negative relationship observed between TPI and CTI, confirming that the hollows in the landscape are also locations where saturation state is likely to be higher. Considering these co-linear relationships, it is important to ask which variables are the most effective predictors of landslide locations for the analysed inventories. To analyse the importance of the input variables, we first perform analysis of the influence of each individual variable as a bivariate analysis (section and 3.1) and then use a Generalized Linear Model to explore the effect of co-linearity (section 3.2).

For the assessment of each parameter by itself, we calculate the relative ratios of the distributions for each variable for the topography and the landslide populations. The topography values are calculated for all pixels within the area in which landslides were mapped. Since we lack data on the mapped areas for all of the inventories, we assume that the convex hull (minimum bounding polygon) for the landslide polygons represents the mapping area. This follows the example of other recent studies [Marc et al. 2018, Milledge et al. 2019]. For both the landslide parameter distribution and the parameter distribution for the topography, we divide the values into bins, normalizing by the total size of the distribution. This essentially represents a value-frequency distribution. For each of the bins, we then divide the landslide probability by that of the topography, to obtain a ratio. Using slope as an example, this provides an estimate of the ~~likelihood~~ probability of a landslide occurrence at a given slope value compared to the occurrence of that slope in the landscape. This step is meant to explore the significance of each variable in bivariate structure (Figure 3).

~~Because landslides are triggered as a result of a complex interaction between various factors, we also analyze the inventories using a multivariate regression scheme to consider the interactions between the topographic factors. We do so by fitting a Bayesian Generalized Linear Model (Tibshirani, 1996)(GLM) assuming that the distribution of landslide presence/absence at the pixel level behaves according to a Bernoulli probability distribution, this being repeated for each landslide inventory. From each model we built, we store the information related to the regression coefficients. Because we implement a Bayesian version of a GLM, any Bayesian model provides estimates on the full distribution of each model component (e.g., Lombardo and Mai, 2018). Before fitting the regression model, we apply a mean-zero and unit-variance normalization to all variables (e.g., Lombardo et al., 2018), which are expressed in different ranges and scales. This normalization allows us to better examine the modelling results in terms of contribution of each variable. In this scheme, larger absolute values of the regression coefficients refer to relatively large contribution of variables.~~

~~We also apply a feature selection algorithm to identify the significant and irrelevant variables. For this purpose, we use Least Absolute Shrinkage and Selection Operation (LASSO) technique (Team R, 2013). This method is particularly suggested for landslide susceptibility assessment to reduce the large number of highly correlated predictors without losing parameter interpretability (e.g., Camilo, et al., 2017). GLM fitting with a LASSO implementation is carried out by using the R Friedman et al. (2009) "*glmnet*" library, which was made available by (Hosmer and Lemeshow, 2000). We apply this method and couple it with the 10-fold cross-validation to remove non-informative covariates and to assess the modeling performance based on the area Under the Receiver Operating Characteristic curve (AUC) calculated for each landslide inventory .~~

We characterize the landslides in two ways – first, by calculating the parameter value for the scar area of the landslide~~,~~ and secondly, by calculating the raster values for the entire landslide body. We lack consistent data on the scar area for the landslide inventories in question, so instead we calculate an approximation of the scar area based on the geometry of each individual landslide. We utili~~z~~ze the method of Marc et al. [2018] to extract the scar areas, ~~based on the retrieval on polygon width from perimeter and area and the empirically supported assumption that the scar is 1.5 longer than wide (Domej et al., 2017).~~ which uses the perimeter and area (A) of landslide polygons to calculate aspect ratio of an equivalent ellipse, K, (Marc and Hovius 2015) and the associated width (W), according to the formula $W \cong \sqrt{(4A/\pi K)}$ (Marc et al., 2018). The scar area is defined as $\sim 1.5W^2$ based on a global database of scar aspect ratio (Domej et al 2017).~~This method utilises the perimeter and area (A) of landslide polygons to calculate the width (W) based on aspect ratio K, according to the formula $\sqrt{\quad}$. The scar area is defined as $\sim 1.5W^2$.~~ For each scar, we calculate the average value within the polygon area of each parameter. This avoids bias towards landslides with long runouts and effectively removes the lowest portion of the polygons which may not have the same topographic signal as~~than~~ the source areas. Even if the scar areas are only approximate, focusing on the upper part of the landslides is warranted for an improved understanding of the topographic control on landslide susceptibility.

The second way of characterizing the landslide - assessing the overall area of the landslide - allows us to focus on areas in the landscape that are likely to be hazardous, including areas where landslide material may end up.

To calculate the parameter distribution for the whole landslide body, we first rasterize the polygons of landslide locations to the resolution of the SRTM DEM (1 arc-second). This provides a binary raster of landslide presence. We then assess the parameter values for each of the pixels where landslides are present. This does mean that the largest landslides are most strongly represented in the distribution, but this is intentional as it permits us to focus on all of the areas affected in the landscape. It is important to note this approach – counting all pixels – is not appropriate for statistical susceptibility analysis, since it could lead to highly dependent datasets. For hazard analysis purposes, we feel it is appropriate to consider all pixels, since larger landslides are consistently more damaging, and we seek to capture the entire footprint.

In Figure 3, an example of the comparison is shown for the landslides occurring in Zimbabwe as a result of Cyclone Idai. We show upstream slope as an illustrative parameter. To compare the landslide data with the topography, we split the data into bins, using the same bins for both landslide and topography. The ~~likelihood~~ probability of the landslide and topography values are then compared with one another. To allow for more consistent comparison of inventories with diverse topography, we normalize the landslide and topography data by the median value for the parameter in question prior to splitting the data into

bins [Marc et al. 2018, Milledge et al. 2019]. This specifically results in the normalised conditional probability [Milledge et al. 2019]. For example, in the case of slope, we calculate the median slope value for all pixels within the mapped area for each inventory and divide the each binned interval by the median slope value calculated across the mapped area.

Finally, for each bin, we calculate the confidence interval for the comparison of topography to landslides using the method of Rault et al. ~~([~~2018~~])~~.

We have generated these estimates for each of the variables listed in Table 3 and for each of the landslide inventories listed in Table 1. One of the variables – forest loss – is a binary variable – it is calculated either as forest is lost or not. As such, we can only compare the average value for landslides and for topography at large, to obtain a relative difference in average forest loss value.

Because landslides are triggered as a result of a complex interaction between various factors, we also analyze the inventories using a multivariate regression scheme to consider the interactions between the topographic factors. We do so by fitting a ~~Bayesian~~binomial Generalized Linear Model (GLM) ~~assuming that the distribution of landslide presence/absence at the pixel level behaves according to a Bernoulli probability distribution, this being repeated~~ for each landslide inventory. We also apply a feature selection algorithm to identify the significant and irrelevant variables to feed the GLM. For this purpose, we use Least Absolute Shrinkage and Selection Operation (LASSO) technique (Tibshirani, 1996). This method is particularly suggested for landslide susceptibility assessment to reduce the large number of highly correlated predictors without losing parameter interpretability (e.g., Camilo, et al., 2017). GLM fitting with a LASSO implementation is carried out by using the R (Team R, 2013) "glmnet" library, which was made available by Friedman et al. (2009). We apply this method and couple it with the 10-fold cross-validation to remove non-informative covariates and to assess the modeling performance based on the area Under the Receiver Operating Characteristic curve (AUC) calculated for each landslide inventory (Hosmer and Lemeshow, 2000). From each model we built, we store the information related to the regression coefficients. ~~Because we implement a Bayesian version of a GLM, any Bayesian model provides estimates on the full distribution of each model component (e.g., Lombardo and Mai, 2018).~~ Before fitting the regression model, we apply a mean zero and unit variance normalization to all variables (e.g., Lombardo et al., 2018), which are expressed in different ranges and scales. This normalization allows us to better examine the modelling results in terms of contribution of each variable. In this scheme, larger absolute values of the regression coefficients refer to relatively large contribution of variables.

~~We also apply a feature selection algorithm to identify the significant and irrelevant variables. For this purpose, we use Least Absolute Shrinkage and Selection Operation (LASSO) technique (Tibshirani, 1996). This method is particularly suggested for landslide susceptibility assessment to reduce the large number of highly correlated predictors without losing parameter interpretability (e.g., Camilo, et al., 2017). GLM fitting with a LASSO implementation is carried out by using the R (Team R, 2013) "glmnet" library, which was made available by Friedman et al. (2009). We apply this method and couple it with the 10-fold cross-validation to remove non-informative covariates and to assess the modeling performance based on the area Under the Receiver Operating Characteristic curve (AUC) calculated for each landslide inventory (Hosmer and Lemeshow, 2000).~~

We have combined the results from each individual inventory into a single figure for each of the variables to assess relative differences, as well as which variables are most strongly associated with where landslides are mapped.

3. Results

3.1. Bivariate analysis

The bivariate analyses show that several of the parameters are strong predictors for the location of both the ~~head scarp~~scars and overall area of landslides, and while there is significant variability between the different inventories there are consistent patterns that emerge across all events. In broad terms, we find that rainfall-triggered landslides occur more often in rough, steep terrain (Figures 4, 5, and 7). Results from the Compound Topographic (CTI) and short-wavelength topographic position suggest that these parameters can be used to effectively distinguish between ~~headscarp~~scars and the entire landslide area, with a high probability~~likelihood~~ of ~~headscarp~~scars at low CTI values and at more positive TPI values (i.e., landscape convexities). For all metrics, we find that all studied inventories have approximately equal sampling at the median landscape value. This can be observed in Figure 4-8, where the probability~~likelihood~~ ratio of 1 for almost all inventories occurs at approximately the median value of the parameter for the entire landscape. In other words, the transition from low to high landslide probability~~likelihood~~ is relative to the local landscape median value, not to an absolute value of the considered metric.

For all of the events, there is a general increase in landslide probability at higher slope values (Figure 4). Similarly, a strong increase in landslide probability~~likelihood~~ is observed for average upstream slope angle (Figure 5). The distributions of the different inventories are slightly tighter than for slope, indicating that this may be a more consistently applicable variable. Specifically, we note that both scars and whole landslides are at an equal sampling at the median landscape upstream angle, strongly undersampling and oversampling the gentler and steeper slopes, respectively (i.e., proportionately more landsliding at higher slope values, and less at lower slopes). This can be observed in Figure 5. Consistent trends emerge where results are within a 95th percentile confidence interval, although there is a greater spread of data values where results are not considered statistically significant.

The Compound Topographic Index (sometimes referred to as the~~or~~ wetness index), when tested for landslide scars shows higher probability~~likelihood~~ of landslides for lower CTI values (Figure 6). This trend is negligible for whole landslide areas, with the statistically significant points showing almost no variation in landslide probability~~likelihood~~ with changing CTI. This suggests broadly that CTI is a poor predictor for the areas where landslide hazard may be increased, but a better predictor of the source locations. The relationship between ~~headscarp~~scar probability~~likelihood~~ and CTI is not strongly linked with flow accumulation, despite the role that flow accumulation plays in setting CTI. In Supplementary Figure S3 we show the probability~~likelihood~~ ratios for flow accumulation values, and no clear relationship emerges between probability~~likelihood~~ ratio and flow accumulation. This suggests that the slope component of CTI is more important when considering ~~headscarp~~scar locations, while the flow accumulation factor (observed to be correlated with an increase in probability~~likelihood~~ of whole

landslide areas) may be offset by the slope in the areas where landslides run-out, leading to no overall correlation with CTI and whole landslide area.

Additionally, we observe that the CTI value where landslide ~~headscarp~~scars and topography are equally sampled is approximately the median value, and the fit for each inventory is relatively consistent.

We find that Topographic Ruggedness Index is also a relatively strong predictor for landslide probability~~likelihood~~, with increases in TRI correlated with increases in landslide probability~~likelihood~~ ratio for almost all events (Figure 7), and statistically significant results observed for several of the inventories. For several events, the results are in line with prior work that has shown roughness and related metrics to be correlated with landslide occurrence [Costanzo et al. 2012, Reichenbach et al. 2018]. While the point of equal sampling of landslides and topography is approximately the median value for the inventories analysed, the slope of the relationship diverges somewhat above and below this point. This suggests that the most heterogeneous parts of the landscape may not be as strong a predictor for landslide occurrence as areas of high slope.

There are not strong systematically consistent relationships between relief at a 1km scale and the probability~~likelihood~~ of landslide ~~headscarp~~scars and total areas. Some increase in probability~~likelihood~~ is observed with increasing relief, although this saturates at relief values above the median relief. This suggests suggesting that relief alone is a relatively poor predictor for the source areas of landslides, or that the resolution of relief may be too coarse. In a few cases – Burundi, Typhoon Morakot in Taiwan, and Kii Province in Japan – there is a slightly more clear increasing relationship.

We observe a link between the short wavelength Topographic Position Index (300m assessment radius) and landslide probability~~likelihood~~ ratios for landslide ~~headscarp~~scars in several, but not all (i.e. not Kalmaegi, Morakot and Kii), of the events (Figure 9). Specifically, landslide ~~headscarp~~scars are significantly more likely at positive TPI values (short wavelength landscape convexities like ridges). Although this is the case for the majority of inventories, results from both inventories in Taiwan and Dominica do not exhibit this tendency. This parameter also shows the clearest distinction between the ~~headscarp~~scar areas and the whole landslide areas. The entire landslide area is more likely to be found at negative TPI values (landscape concavities like valley bottom). While TPI at 300m wavelength does not demonstrate quite as consistent relationships as slope or CTI, it remains a strong predictor. In particular, the larger variation between ~~headscarp~~scars and overall landslide areas suggests that short-wavelength may be a valuable way to distinguish between scarps and deposits in a preliminary assessment.

~~The Topographic Position Index assessed over longer wavelengths (2000 m) does not show any significant or consistent links with landslide likelihood ratios, suggesting that TPI assessed on this scale does not effectively resolve the hillslope characteristics associated with landsliding.~~

Since Forest Loss is a binary variable, we do not plot this across multiple bins. However, we calculate the average ratio of forest loss in landslide zones to the overall topography, and we find across all events that the value is 1.98 ±1.38 (1 standard deviation). This implies that forest loss zones are correlated with higher probability~~likelihood~~ of landsliding. We see the largest differences between landslide zones and non-landslide zones in Salgar, Colombia, Hiroshima, Japan, and the landslides

triggered by Typhoon Morakot in Taiwan (forest loss is more than 3.5 times more likely in landslides in these cases). In 4 cases, forest loss was observed as less likely in landslide pixels: in Blumenau, Brazil, Lanao del Norte, Philippines, Thrissur, India, and West Pokot, Kenya. We suggest that while forest loss prior to landslide events is generally higher in landslide locations than in the rest of the landscape, this relationship is not consistent enough to necessarily be a good predictor of landslide location by itself.

3.2. Multivariate analysis

We have used the LASSO method to quantify the importance of the different predictors for both ~~headscarp~~scars and whole landsides, while reducing the influence of co-linearity [Camilo et al. 2017] (Figure 10).

Figures 10A and 10B also show the modelling performances, which are represented by AUC values, varying from "reasonable" (0.6<AUC<0.7) to "adequate" (0.7 < AUC < 0.8) and "outstanding" (0.8<AUC<0.9) based on AUC classes defined by Hosmer and Lemeshow, (2000). The lower AUC values for some events may indicate that we lack some key explanatory variables; crucially, this may include rainfall intensity and duration, although as discussed above we lack consistent high-resolution data to assess this within the spatial domain of the considered inventories. Other relevant parameters may include land cover, lithology and geo-structural parameters, and anthropogenic influence (Reichenbach et al. 2018).

Our findings show that slope, on one hand, is the factor that most frequently appears as significant in GLM run for both the landslide ~~headscarp~~scars and the whole areas, with landslides favoring steeper locations. On the other hand, average upstream angle and CTI in GLM of the ~~headscarp~~scars and average upstream angle, CTI and Topographic Position Index (with 300m radius) in GLM run for the whole area appear as the least commonly observed significant variables. The non-significant or low impact of most of these topographic variables is likely due to the collinearity existing between these variables (e.g., slope and average upstream slope, slope and CTI, CTI and TPI – Supplementary Figures S1 and S2).

The results indicate that except for CTI, all variables have a positive weight on classifying a given grid cell as "landslide presence" instead of "landslide absence" given the choice of predictors. There are only two cases that do not follow this general trend (i.e., the two Taiwanese inventories).

The regression coefficient of TPI calculated for the landslide ~~headscarp~~scar areas in Kalmaegi, Taiwan has a negative sign, unlike all other cases. Overall, as we explained above, different TPI values refer to different subsections of a hillslope. Specifically, positive values refer to ridges or hillslopes, whereas negative values correspond to the valley. As a result, the negative weight of TPI on classifying the landslide presence and absence condition in the GLM is difficult to interpret. This could be caused by interactions between variables. Although we run a variable selection method (i.e., LASSO), TPI could be still interacting with the others. If it shares a similar signal to at least one another variable, the sign of the regression coefficient can be influenced by the interaction.

The negative regression coefficient of TRI obtained for the whole landslide areas in Taiwan, Morakot is the other case that the response of the covariate is different from the other examples. Similar to the Kalmaegi inventory case we presented above, the negative sign could be also caused by the interactions between variables. However, this could be also associated with the

physical properties characterizing whole landslide areas. TRI in Taiwan, Morakot, in particular, might correspond to smooth topography. In either case, TRI does not appear as a significant variable other than the Taiwan, Morakot.

The two inventories where negative CTI values are most significantly associated with landslide incidence – Morakot and Hiroshima – have few commonalities; their lithologies differ, and the mean slope of the affected area in Hiroshima is markedly lower than Morakot. Perhaps the most significant commonality is that the triggering rainfall exceeded the historical maxima by a significant degree (Table 3), which may increase the likelihood of failure resulting from local transient pore pressure increases, rather than due to saturated flow at the base of hillslopes. Excepting these two examples, the results show that the signal of CTI is not contributing to the GLM while classifying landslide presence or absence.

4. Discussion

The primary intention of this study is to assess the critical topographic parameters associated with rainfall triggered landslides, using a large dataset of landslide inventories that includes 6 newly mapped events. Our results are comparable with existing studies [e.g. Marc et al. 2018, Milledge et al. 2019] exploring landslide locations using inventories, while adding more detail and assessment of global variability, which we discuss below. First, however, we explore some of the limitations and assumptions that go into the mapping and analysis of the landslides.

4.1. Uncertainty in landslide mapping and DEM metrics extraction

Firstly, it is important to consider how representative the landslide inventories are. We have attempted to include landslide maps from a diverse set of locations around the world, but this is still only a fraction of landslides that have occurred in the last 2 decades. Some of the inventories, like the landslides occurring due to Typhoon Morakot in Taiwan, are driven by such huge rainfall events that the overall area of landsliding greatly exceeds other examples with lower rainfall. This is one of the key reasons why we have used probability~~likelihood~~ ratios as a metric to assess landslide locations, since it does not consider the overall area of landslides triggered by a given rainfall event. Most of the inventories are drawn from locations with tropical climates [Geiger 1954], and although the pairs from Brazil and Japan sample areas of humid subtropical climates this is not a true representative sampling of climatic regimes. One exception is the landslide inventory from Zimbabwe, which is in a semi-arid climatic region. Although our examples disproportionately sample tropical and subtropical areas, these areas generally experience the highest erosion rates [Milliman & Syvitski 1991], driven by increasingly intense rainfall (e.g., Bookhagen & Strecker 2012). Similarly, while the inventories are drawn from places with diverse lithologies, we do not have datasets from a fully representative set of lithological locations [Hartmann & Moosdorf 2012, Table 1].

Although we have used hand-corrections to reduce the impact of polygon amalgamation from algorithmic mapping methods, ~~While we have attempted to ensure the consistency of landslide inventories by hand-correcting the events~~, some inconsistencies may still exist. One important consideration is that the datasets used here do not distinguish between different landslide types, or distinguish between scar and deposit. While this is consistent across inventories, it is important when considering the results.

In particular, since we do not have constraints on whether mapped landslides are purely shallow soil slides or whether they incorporate deeper bedrock, we cannot determine differences in topographic characteristics associated with each. The change in material properties from soil to bedrock can lead to changes in overall volume mobilized for a given landslide area [Larsen et al. 2010], so inventories where smaller, shallow landslides are a larger proportion of the mapped inventory may have different characteristics. For example, the landslides mapped using aerial photography around Hiroshima, Japan, do not show particularly high probability~~likelihood~~ ratios at very high relief or TRI values, suggesting these landslides generally occur on smaller, less rough hillslopes.

Hand mapping and correcting will help reduce the potential for landslide amalgamation [Marc & Hovius 2015], which is essential in order to estimate width and ~~headscarp~~scars areas from landslide polygon geometry [Marc et al. 2018]. Since we do not have access to the imagery for all of the previously published events, we are not able to correct these events and thus must rely on prior mapping being consistent with our own efforts. Part of the challenge of compiling different events is the different sources of imagery used to create each inventory. For most of the inventories we have mapped as part of this study, the imagery is consistent in terms of resolution, and we have benefited from the rapid return time of Planet Dove satellites to ensure that cloud cover does not mask any of the areas mapped. However, without imagery to clarify it may be possible that parts of the previously published inventories are masked by cloud cover. In addition, the inventories mapped using coarser resolution satellite imagery, such as the part of Taiwan impacted by Typhoon Kalmaegi, may not capture the smallest landslides that resulted. If smaller landslides are preferentially found in certain parts of the landscape this may introduce systematic biases in observed probability~~likelihood~~ ratios.

While the imagery used to compile the different inventories varies in resolution, there do not seem to be consistent, systematic differences between probability~~likelihood~~ ratios that can be explained as a result of small landslides that systematically bias the events. For example, landslides in the Dominica events have similar probability~~likelihood~~ ratios for each parameter compared to the landslides from Typhoon Kalmaegi in Taiwan, despite these datasets having the largest difference in effective imagery resolution.

When considering the whole landslide area, we pixelate the landslides to the resolution of the DEM to highlight the most hazardous parts of the landscape. This pixelation process can introduce a source of systematic error, since if less than half of a cell area is occupied by a landslide polygon it is still considered to be a 'landslide pixel'. Some landslides may be significantly smaller than the SRTM cell resolution if they are mapped using high resolution imagery, but they will still count as a full size pixel for the purposes of analysis as a result of the rasterization process. This introduces a potential source of bias as smaller landslides may make up a larger proportion of analysed pixels than their actual area would represent. Landslides below the approximate area of half an SRTM cell (450 square metres) make up an appreciable proportion of total landslide area in Hiroshima and Dominica (~30%) (Supplementary Figure 12). In these settings, we would expect that the influence of smaller landslides may be over-estimated at the 30m resolution of analysis, although in the majority of other inventories landslides below this cutoff represent less than 5-10% of the total landslide area (and therefore the analytical dataset of landslide pixels).

To address the potential for bias due to over-sampling of small landslides with a coarse resolution DEM, we have resampled the DEM for the events in Hiroshima and Dominica to a resolution of 10m, at which nearly all landslides are captured without size exaggeration. We then recalculate the probability~~likelihood~~ ratios for the landslides and compare the resampled DEM results with those from the original DEM [Supplementary Figure S5]. For slope, average upstream angle, CTI, relief and TRI, only minor differences are observed between the results for a resampled DEM and the original DEM. There are some differences between TPI at 300m resolution, but no consistent relationship seems to emerge. Thus we do not think our results are affected by the coarse rasterization process, although it is likely that accessing higher resolution DEM may alter the result depending on local variable, like slope, CTI or TRI.

Other parameters are often incorporated into landslide susceptibility such as geological factors like soil characteristics or lithology, local land cover type or climatic metrics. Although global data for rainfall, soil type and geological parameters exists, the resolution of these datasets is too low to allow for consistent comparison of landslide and non-landslide areas at the scale of the analysis described here (~100m)~~Characterizing these parameters consistently across different inventories requires a global dataset, but the resolution of globally available, open data for rainfall, soil type, or geological parameters is too low to differentiate between landslide zones and non-landslide zones~~. As such, we have chosen to focus exclusively on topographic factors within our assessment.

4.2. Differences between ~~headscarp~~scar and overall landslide area

While for several parameters, ~~headscarp~~scars and the entire landslide area are similarly sampled at a range of values, for TPI and CTI we see significant differences across a large number of the inventories (Figure 11). ~~Headscarp~~Scars are more likely at lower CTI values, and at more positive TPI values. Positive TPI implies ~~headscarp~~scars are more likely at concave locations in the landscape, while a lower CTI value indicates areas with lower flow accumulation and saturation. This describes parts of the landscape that sit closer to ridges. This broadly supports the assessment above that higher TPI and CTI values may be a way to distinguish between scar and deposit areas. No systematic differences are observed with respect to TRI or average upstream angle (Supplementary Figure 4)

By comparing ~~headscarp~~scars and the overall landslide area, the observations we have made provide informative contrasts with prior work. Similar recent work exploring the characteristics of earthquake-induced landslide inventories suggests that slope angle and upslope contributing area as key determinants of hazard, defined by the entire landslide areas [Milledge et al. 2019]. Our findings are consistent for entire landslide areas, but differ for ~~headscarp~~scar area, which is poorly determined by flow accumulation [Supplementary Figure S3]. One may be surprised by the fact that landslides triggered by intense rainfall have ~~headscarp~~scars uncorrelated with drainage area, while both earthquake and rainfall induced landslides have whole areas strongly related to it. We propose that the whole area relationship mainly reflects runout paths and not hydrological processes, and that the initiation of rainfall induced landslides poorly relates to the surface parallel hydrological flow. This is discussed in more detail below. Our results for drainage are of the whole landslides are quite different from the ones of Milledge et al.,

2019. We suggest that the variability in scar location (higher for EQ, see Rault et al 2018) may explain more diverse behavior in normalized drainage below the median, while the propensity to longer runout (more likely for rainfall induced landslides) may explain that some (not all) cases have probability ratio increasing until very large drainage (Supplementary Figure S3).

The observed differences between ~~headscarp~~scar and overall landslide area can be exploited to refine our understanding of susceptibility and hazard modelling, by focusing on parameters controlling ~~headscarp~~scar areas where landslides initiate (e.g., slope, CTI), and the entire landslide area where landslides impact (e.g., drainage area, TPI), respectively. Such a focus can help support both sets of applications for a more comprehensive landslide hazard information and emphasises the need to distinguish diverse portions of mapped landslides depending on the study objective.

4.3. Implications for triggering mechanisms and landscape evolution

One of the most consistent observations that emerges from this study is that for several parameters (Slope, Average upstream slope, TRI, CTI), the critical point where landslides and topography are equally sampled is approximately the median value for the inventory in question. This is consistent with previous observations on rainfall and earthquake induced landsliding (Marc et al 2018, Milledge 2019). For average upstream slope and CTI, the relationships for different inventories are in fact very similar; this is despite a large variation in the median slope for each of the inventories (Table 2). This suggests that landslide probability~~likelihood~~ is strongly dependent on the median topography, rather than a specific critical angle. This implies two important points: first, that despite important differences in the landscapes observed, consistent hazard relationships can be defined based upon median landscape values, and second, that these diverse landscapes may be in a form of long-term equilibrium with respect to their landslide behaviour.

We suggest that each of the considered landscapes, with its own lithology, vegetation, climate and tectonic forcing, may have evolved such that local hillslopes have slope gradients and hydromechanical properties that set the possibility for landslides on the upper half of the distribution. The ~~E~~evolution of the hillslopes' regolith state, which acts as an important control on landslide susceptibility, under climatic forcing is predicted by geomorphological models of hillslope stability coupled with stochastic rainfall forcing (Dietrich et al., 1995, Iida 1999, 2004). Landscape evolution towards a critical state was also inferred to explain why landsliding in the Kii peninsula better matched relative rainfall anomaly rather than absolute rainfall patterns (Marc et al., 2019).

Alongside the implications for landscape evolution and how to derive susceptibility metrics, our results also offer insight into the mechanisms of landslide triggering in extreme rainfall events. By focusing on the ~~headscarp~~scar areas, we observe that landslides are more likely in locations with lower CTI and higher TPI values – parts of a landscape near ridges with a generally lower propensity for water saturation. This is somewhat in contrast to studies that suggest rainfall-triggered landslides are more likely to occur in areas lower down hillslopes where fluid saturation is greater [Densmore & Hovius 2000, Meunier et al. 2008], possibly because they did not clearly differentiate scar from whole landslide area which are very different relative to these two metrics [Figure 11]. T~~Anyway, t~~he relationship with CTI and drainage area also suggests that modeling landslides

under the assumption of regolith saturation due to slope-parallel, steady state flow (e.g., Montgomery and Dietrich 1995) may be inadequate. Instead, the pore-pressure triggering landslides in extreme rainfall events may rather be controlled by transient, vertical infiltration, and/or preferential flow paths [Iverson, 2000, Montgomery et al., 2009, Hencher 2010, Bogaard and Greco 2016]. This recalls the essential challenge for developing modeling approaches that can account for such complex hillslope hydrology as well as with highly variable hydromechanical properties of the regolith. Nevertheless, we also suggest that future studies should compare the results from this analysis of landslides triggered by extreme rainfall with landslide inventories resulting from longer duration, lower intensity rainfall events, to assess whether the relationship with CTI and TPI changes. Indeed, we might expect that lower intensity rainfall would trigger landslides at parts of the landscape with higher CTI values as steady state saturation may be more widespread. In any future comparative study of low-intensity and high-intensity rainfall events, it will be necessary to carefully select landslide inventories where the imagery used to generate them closely brackets the start and end of the rainfall events, to ensure only landslides triggered by an individual event are analysed.

Finally, for some events, including Typhoon Morakot in Taiwan and Cyclone Idai in Zimbabwe, there is a small decline in relative landslide probability at very high slope values (> 35-40 degrees) [Figure 4]. It is possible that these slopes, which are generally above the angles considered to be critically unstable [Selby 1982, Roering et al. 2001] may represent areas where landslide probability~~likelihood~~ is reduced as a result of non-topographic factors, such as local lithological bedding plane angles [Guzzetti et al. 1996, Santangelo et al. 2015], or thinner soils limiting the availability of material that can be mobilized as landslides [Prancevic et al. 2020]. The decline in landslide probability~~likelihood~~ at the highest slope values may represent the limit to which local pore pressure as a result of extreme rainfall can influence triggering.

5. Conclusions

In this study we have combined 10 existing rainfall-induced landslide inventories from a range of mountainous regions with 6 new inventories mapped as part of this study. We suggest that providing newly mapped inventories is a valuable service for the landslide community at large, and we anticipate that these inventories can provide data to calibrate and validate susceptibility and hazard models both in the specific locations where landslides occurred but also further afield. In addition, we have used moderate resolution open source satellite data to assess the parameters that characterize the location of landslides in these inventories. We find that alongside the previously documented importance of slope and topographic ruggedness, average upstream angle and topographic position are also determinants of landslide probability~~likelihood~~ in a given location. After normalizing the topographic variables by the local landscape median, we find consistent relationships across the different inventories despite the variety of lithological and topographic settings. This suggests that relative metrics should be considered to perform landslide susceptibility analysis and that different landscapes can be at a state of equilibrium with respect to the probability~~likelihood~~ of landsliding. The importance of multiple topographic factors to determine the local landslide probability~~likelihoods~~ highlights the value of high resolution DEM data. While we have used the 1 arc-second resolution SRTM data, higher resolution DEM data are increasingly available. Given that we are able to map landslides at finer and finer

resolution as very-high resolution satellite imagery becomes available, combining these new detailed inventories with DEMs of similar resolution is likely to provide further insights about landslide location within the landscape.

While we have not undertaken a detailed assessment of the rainfall that triggered these landslides, we emphasise that variability in rainfall is likely to explain a significant degree of variability in where landslides occur [e.g., Marc et al. 2019]. Future work should assess each of these inventories with respect to the rainfall that triggered the significant landsliding could yield important insights into the relationship between intense rainfall and landslide occurrence.

Acknowledgements

R.E., P.A. and D.K. are supported by a NASA DISASTERS program grant 18-DISASTER18-0022. All authors were involved in study conceptualisation and writing of the manuscript. R.E., P.A. & O.M. conducted landslide mapping. R.E., H.T. & O.M. conducted data analysis.

Data Availability

All data used in this study is provided in the supplementary material, and all methods are detailed above.

Declaration of Competing Interests

The authors declare that they have no conflict of interest.

References

Amatya, P., Kirschbaum, D., Stanley, T., & Tanyas, H. Landslide mapping using object-based image analysis and open source tools. Engineering Geology, 282, 106000. https://doi.org/10.1016/j.enggeo.2021.106000 (2021)

Amatya, P., Kirschbaum, D., & Stanley, T. Use of very high-resolution optical data for landslide mapping and susceptibility analysis along the Karnali highway, Nepal. Remote Sensing, 11(19), 2284. https://doi.org/10.3390/rs11192284 (2019).

The Association of Japanese Geographers. The 2018 July Heavy rain in West Japan. Available online: http://ajg-disaster.blogspot.com/2018/07/3077.html (accessed on 1 November 2019). (In Japanese).

Badoux, A., Andres, N., & Turowski, J. M. Damage costs due to bedload transport processes in Switzerland. Natural Hazards and Earth System Sciences, 14, 279–294. https://doi.org/10.5194/nhess-14-279-2014 (2014).

Behling, R., Roessner, S., Segl, K., Kleinschmit, B., & Kaufmann, H. Robust automated image co-registration of optical multi-sensor time series data: Database generation for multi-temporal landslide detection. Remote Sensing, 6(3), 2572-2600. https://doi.org/10.3390/rs6032572 (2014).

Bogaard, T. A., & Greco, R. Landslide hydrology: from hydrology to pore pressure. Wiley Interdisciplinary Reviews: Water, 3(3), 439-459. https://doi.org/10.1002/wat2.1126 (2016).

Bookhagen, B., & Strecker, M. R. Spatiotemporal trends in erosion rates across a pronounced rainfall gradient : Examples from the southern Central Andes. Earth and Planetary Science Letters, 327–328, 97–110. https://doi.org/10.1016/j.epsl.2012.02.005 (2012).

Broeckx, J., Maertens, M., Isabirye, M., Vanmaercke, M., Namazzi, B., Deckers, J., … Poesen, J. Landslide susceptibility and mobilization rates in the Mount Elgon region , Uganda, (October 2018), 571–584. https://doi.org/10.1007/s10346-018-1085-y (2019).

Budimir, M.E.A., Atkinson, P.M. & Lewis, H.G. A systematic review of landslide probability mapping using logistic regression. Landslides 12, 419–436 https://doi.org/10.1007/s10346-014-0550-5 (2015).

Burrows, K., Walters, R. J., Milledge, D., & Densmore, A. L. A systematic exploration of satellite radar coherence methods for rapid landslide detection. *Natural Hazards and Earth System Sciences*, *20*(11), 3197–3214. https://doi.org/10.5194/nhess-20-3197-2020 (2020).

Camilo, D. C., Lombardo, L., Mai, P. M., Dou, J., & Huser, R. Environmental Modelling & Software Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized Generalized Linear Model. Environmental Modelling and Software, 97, 145–156. https://doi.org/10.1016/j.envsoft.2017.08.003 (2017).

Casagli, N., Frodella, W., Morelli, S., Tofani, V., Ciampalini, A., Intrieri, E., … Lu, P. Spaceborne , UAV and ground-based remote sensing techniques for landslide mapping , monitoring and early warning. Geoenvironmental Disasters, 4(9), 1–23. https://doi.org/10.1186/s40677-017-0073-1 (2017).

Chang, K., Chiang, S., Chen, Y., & Mondini, A. C. Modeling the spatial occurrence of shallow landslides triggered by typhoons. Geomorphology, 208, 137–148. https://doi.org/10.1016/j.geomorph.2013.11.020 (2014).

Chen, Y., Chang, K., Chiu, Y., Lau, S., Lee, H., & County, T. Quantifying rainfall controls on catchment-scale landslide erosion in Taiwan. Earth Surface Processes and Landforms, 382, 372–382. https://doi.org/10.1002/esp.3284 (2013).

Conrad, J. L., Morphew, M. D., Baum, R. L., & Mirus, B. B. HydroMet: A New Code for Automated Objective Optimization of Hydrometeorological Thresholds for Landslide Initiation. Water, 13(13), 1752. https://doi.org/10.3390/w13131752 (2021).

Costanzo, D., Rotigliano, E., Irigaray, C., Jiménez-Perálvarez, J. D., & Chacón, J. Factors selection in landslide susceptibility modelling on large scale following the gis matrix method: Application to the river Beiro basin (Spain). Natural Hazards and Earth System Science, 12(2), 327–340. https://doi.org/10.5194/nhess-12-327-2012 (2012).

Dietrich, W. E., Reiss, R., Hsu, M. L., & Montgomery, D. R. A process-based model for colluvial soil depth and shallow landsliding using digital elevation data. Hydrological processes, 9(3-4), 383-400. https://doi.org/10.1002/hyp.3360090311 (1995).

Densmore, A. L., & Hovius, N. Topographic fingerprints of bedrock landslides. Geology, 28(4), 371–374. https://doi.org/10.1130/0091-7613(2000)28<371:TFOBL>2.0.CO;2 (2000).

Froude, M. J., & Petley, D. N. Global fatal landslide occurrence from 2004 to 2016. Natural Hazards and Earth System Science, 18, 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018 (2018).

García-Rodríguez, M. J., Malpica, J. A., Benito, B., & Díaz, M. Susceptibility assessment of earthquake-triggered landslides in El Salvador using logistic regression. Geomorphology, 95, 172–191. https://doi.org/10.1016/j.geomorph.2007.06.001 (2008).

Geiger, R. Klassifikation der Klimate nach W. Köppen. In Landolt-Börnstein – Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik, alte Serie. (pp. 603–607). Berlin: Springer. (1954).

Goetz, J. N., Brenning, A., Petschko, H., & Leopold, P. Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers and Geosciences, 81, 1–11. https://doi.org/10.1016/j.cageo.2015.04.007 (2015).

Goetz, J. N., Guthrie, R. H., & Brenning, A. Geomorphology Integrating physical and empirical landslide susceptibility models using generalized additive models. Geomorphology, 129(3–4), 376–386. https://doi.org/10.1016/j.geomorph.2011.03.001 (2011).

Guzzetti, F., Cardinali, M., & Reichenbach, P. The Influence of Structural Setting and Lithology on Landslide Type and Pattern. Environmental and Engineering Geoscience, II(4), 531–555. https://doi.org/10.2113/gseegeosci.II.4.531 (1996).

Guzzetti, F., Cesare, A., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. Landslide inventory maps: New tools for an old problem. Earth Science Reviews, 112(1–2), 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001 (2012).

Handwerger, A. L., Fielding, E. J., Huang, M., Bennett, G., Liang, C., & Schulz, W. H. Widespread Initiation, Reactivation , and Acceleration of Landslides in the Northern California Coast Ranges due to Extreme Rainfall. Journal of Geophysical Research: Earth Surface, 124, 1782–1797. https://doi.org/10.1029/2019JF005035 (2019).

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., … Townshend, J. R. G. High-Resolution Global Maps of 21st-Century Forest Cover Change. Science, 342(6160), 850–853. https://doi.org/10.1126/science.1244693 (2013).

Harp, B. E. L., Reid, M. E., Michael, J. A., & Report, U. S. G. S. O. Hazard Analysis of Landslides Triggered by Typhoon Chata'an on July 2, 2002, in Chuuk State, Federated States of Micronesia. USGS Open-File Report 2004-1348. https://doi.org/10.3133/ofr20041348 (2004).

Harp, E. L., Keefer, D. K., Sato, H. P., & Yagi, H. Landslide inventories: The essential part of seismic landslide hazard analyses. Engineering Geology, 122(1–2), 9–21. https://doi.org/10.1016/j.enggeo.2010.06.013 (2011).

Hartmann, J., & Moosdorf, N. The new global lithological map database GLiM: A representation of rock properties at the Earth surface. Geochemistry, Geophysics, Geosystems, 13(12), 1–37. https://doi.org/10.1029/2012GC004370 (2012).

Hencher, S. R. Preferential flow paths through soil and rock and their association with landslides. Hydrological processes, 24(12), 1610-1630. https://doi.org/10.1002/hyp.7721 (2010).

Hong, Y., Adler, R., & Huffman, G. Evaluation of the potential of NASA multi-satellite precipitation analysis in global landslide hazard assessment. Geophysical Research Letters, 33(October), 1–5. https://doi.org/10.1029/2006GL028010 (2006).

Hu, X., Bürgmann, R., Lu, Z., Handwerger, A. L., Wang, T., & Miao, R. Mobility, Thickness, and Hydraulic Diffusivity of the Slow - Moving Monroe Landslide in California Revealed by L - Band Satellite Radar Interferometry Journal of Geophysical Research: Solid Earth. Journal of Geophysical Research: Earth Surface, 124, 7504–7518. https://doi.org/10.1029/2019JB017560 (2019).

Iida, T. A stochastic hydro-geomorphological model for shallow landsliding due to rainstorm. Catena, 34(3-4), 293-313. https://doi.org/10.1016/S0341-8162(98)00093-9 (1999).

Iida, T. Theoretical research on the relationship between return period of rainfall and shallow landslides. Hydrological Processes, 18(4), 739-756. https://doi.org/10.1002/hyp.1264 (2004).

Jibson, R. W., Harp, E. L., & Michael, J. A. A method for producing digital probabilistic seismic landslide hazard maps. Engineering Geology, 58, 271–289. https://doi.org/10.1016/S0013-7952(00)00039-9 (2000).

Kirschbaum, D. B., & Stanley, T. Satellite-Based Assessment of Rainfall-Triggered Landslide Hazard for Situational Awareness. Earth's Future, 6(3), 505–523. https://doi.org/10.1002/2017EF000715 (2018).

Kirschbaum, D. B., Stanley, T., & Zhou, Y. Spatial and temporal analysis of a global landslide catalog. Geomorphology, 249, 4–15. https://doi.org/10.1016/j.geomorph.2015.03.016 (2015).

Korup, O., Görüm, T., Hayakawa, Y., Without power? Landslide inventories in the face of climate change. Earth Surf. Process. Landforms 37, 92–99. https://doi.org/10.1002/esp.2248 (2012)

Larsen, I. J., Montgomery, D. R., & Korup, O. Landslide erosion controlled by hillslope material. Nature Geoscience, 3(4), 247–251. https://doi.org/10.1038/ngeo776 (2010).

Liang, W., & Chan, M. Spatial and temporal variations in the effects of soil depth and topographic wetness index of bedrock topography on subsurface saturation generation in a steep natural forested headwater catchment. Journal of Hydrology, 546, 405–418. https://doi.org/10.1016/j.jhydrol.2017.01.033 (2017).

Lin, C., Chang, W., Liu, S., Tsai, T., Lee, S., Tsang, Y., … Tseng, C. Landslides triggered by the 7 August 2009 Typhoon Morakot in southern Taiwan. Engineering Geology, 123(1–2), 3–12. https://doi.org/10.1016/j.enggeo.2011.06.007 (2011).

Lombardo, L., & Mai, P. M. Presenting logistic regression-based landslide susceptibility results. Engineering Geology, 244, 14–24. https://doi.org/10.1016/j.enggeo.2018.07.019 (2018).

Lombardo, L., Optiz, T., & Huser, R. Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. Stochastic Environmental Research and Risk Assessment, 32, 2179–2198. https://doi.org/10.1007/s00477-018-1518-0 (2018).

Lombardo, L., Tanyas, H., Chrono-validation of near-real-time landslide susceptibility models via plug-in statistical simulations. Eng. Geol. 278, 105818. https://doi.org/https://doi.org/10.1016/j.enggeo.2020.105818 (2020).

Malamud, B. D., Turcotte, D. L., Guzzetti, F., & Reichenbach, P. Landslide inventories and their statistical properties. Earth Surface Processes and Landforms, 29(6), 687–711. https://doi.org/10.1002/esp.1064 (2004).

Marc, O., & Hovius, N. Amalgamation in landslide maps: effects and automatic detection. Natural Hazards and Earth System Science, 15, 723–733. https://doi.org/10.5194/nhess-15-723-2015 (2015).

Marc, O., Stumpf, A., Malet, J., Gosset, M., Uchida, T., & Chiang, S. Initial insights from a global database of rainfall-induced landslide inventories: the weak influence of slope and strong influence of total storm rainfall. Earth Surface Dynamics, 6, 903–922. (2018).

Marc, O., Gosset, M., Saito, H., Uchida, T., & Malet, J.-P. Spatial Patterns of Storm-Induced Landslides and Their Relation to Rainfall Anomaly Maps. Geophysical Research Letters, 46(20), 11167–11177. https://doi.org/10.1029/2019GL083173 (2019).

Marc, O., Meunier, P., & Hovius, N. Prediction of the area affected by earthquake-induced landsliding based on seismological parameters. Nat. Hazards Earth Syst. Sci., 17(7), 1159–1175. https://doi.org/10.5194/nhess-17-1159-2017 (2017).

Marc, O., Hovius, N., Meunier, P., Gorum, T., & Uchida, T. A seismologically consistent expression for the total area and volume of earthquake-triggered landsliding. Journal of Geophysical Research: Earth Surface, 121(4), 640-663. https://doi.org/10.1002/2015JF003732 (2016).

Martha, T. R., Kerle, N., Van Westen, C. J., Jetten, V., & Kumar, K. V. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. ISPRS journal of photogrammetry and remote sensing, 67, 105-119. https://doi.org/10.1016/j.isprsjprs.2011.11.004 (2012).

Meunier, P., Hovius, N., & Haines, J. A. Topographic site effects and the location of earthquake induced landslides. Earth and Planetary Science Letters, 275(3–4), 221–232. https://doi.org/10.1016/j.epsl.2008.07.020 (2008).

Milledge, D. G., Densmore, A. L., Bellugi, D., Rosser, N. J., Watt, J., Li, G., & Oven, K. J. (2019). Simple rules to minimise exposure to coseismic landslide hazard, Nat. Hazards Earth Syst. Sci., 19, 837–856. https://doi.org/10.5194/nhess-19-837-2019 (2019)

Milliman, J., & Syvitski, J. P. M. Geomorphic Tectonic Control of Sediment Discharge to Ocean – The Importance of Small Mountainous Rivers Geomorphic / Tectonic Control of Sediment Discharge to the Ocean: The Importance of Small. The Journal of Geology, 100(5), 525–544. https://doi.org/10.1086/629606 (1991).

Mirus, B. B., Jones, E. S., Baum, R. L., Godt, J. W., Slaughter, S., Crawford, M. M., ... & McCoy, K. M. (2020). Landslides across the USA: occurrence, susceptibility, and data limitations. Landslides, 17, 2271-2285.

Montgomery, D. R., Schmidt, K. M., Dietrich, W. E., & McKean, J. Instrumental record of debris flow initiation during natural rainfall: Implications for modeling slope stability. Journal of Geophysical Research: Earth Surface, 114(F1). https://doi.org/10.1029/2008JF001078 (2009).

Montgomery, D. R., & Dietrich, W. E. A physically based model for the topographic control on shallow landsliding. Water resources research, 30(4), 1153-1171. https://doi.org/10.1029/93WR02979 (1994).

Nowicki Jessee, M. A., Hamburger, M. W., Allstadt, K., Wald, D. J., Robeson, S. M., Tanyas, H., … Thompson, E. M. A Global Empirical Model for Near-Real-Time Assessment of Seismically Induced Landslides. Journal of Geophysical Research: Earth Surface, 123, 1835–1859. https://doi.org/10.1029/2017JF004494 (2018).

Pawluszek, K., Borkowski, A., & Tarolli, P. Sensitivity analysis of automatic landslide mapping : numerical experiments towards the best solution. Landslides, 15, 1851–1865. https://doi.org/10.1007/s10346-018-0986-0 (2018).

Petley, D. Global patterns of loss of life from landslides. Geology, 40(10), 927–930. https://doi.org/10.1130/G33217.1 (2012).

Prancevic, J. P., Lamb, M. P., McArdell, B. W., Rickli, C., & Kirchner, J. W. Decreasing landslide erosion on steeper slopes in soil-mantled landscapes. Geophysical Research Letters, 47(10), e2020GL087505. https://doi.org/10.1029/2020GL087505 (2020).

Rault, C., Robert, A., Marc, O., Hovius, N., & Meunier, P. Seismic and geologic controls on spatial clustering of landslides in three large earthquakes. *Earth Surface Dynamics*, *7*(3), 829–839. https://doi.org/10.5194/esurf-7-829-2019 (2019).

Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. A review of statistically-based landslide susceptibility models. Earth-Science Reviews, 180(November 2017), 60–91. https://doi.org/10.1016/j.earscirev.2018.03.001 (2018).

Riley, S. J., DeGloria, S. D., & Elliot, R. A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity. Intermountain Journal of Sciences, 5(1–4), 23–27. (1999).

Roering, J. J., Kirchner, J. W., Sklar, L. S., & Dietrich, W. E. Hillslope evolution by nonlinear creep and landsliding: An experimental study. Geology, 29(2), 143–146. https://doi.org/10.1130/0091-7613(2001)029<0143:HEBNCA>2.0.CO;2 (2001).

Rossi, G., Tanteri, L., Tofani, V., Vannocci, P., Moretti, S., & Casagli, N. Multitemporal UAV surveys for landslide mapping and characterization. Landslides, 15, 1045–1052. https://doi.org/10.1007/s10346-018-0978-0 (2018).

Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., Ardizzone, F., & Rossi, M. Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory. Landslides, 14(2), 547–558. https://doi.org/10.1007/s10346-016-0739-x (2017).

Santangelo, M., Marchesini, I., Cardinali, M., Fiorucci, F., Rossi, M., Bucci, F., & Guzzetti, F. A method for the assessment of the influence of bedding on landslide abundance and types. Landslides, 12, 295–309. https://doi.org/10.1007/s10346-014-0485-x (2015).

Schmitt, R.G., Tanyas, H., Nowicki Jessee, M.A., Zhu, J., Biegel, K.M., Allstadt, K.E., Jibson, R.W., Thompson, E.M., van Westen, C.J., Sato, H.P., Wald, D.J., Godt, J.W., Gorum, T., Xu, C., Rathje, E.M., Knudsen, K.L. An open repository of earthquake-triggered ground-failure inventories, Data Series. Reston, VA. https://doi.org/10.3133/ds1064 (2017)

Selby, M. J. Controls on the Stability and Inclinations of Hillslopes formed on hard rock. Earth Surface Processes and Landforms, 7, 449–467. https://doi.org/10.1002/esp.3290070506 (1982)

Stanley, T. A., & Kirschbaum, D. B. A heuristic approach to global landslide susceptibility mapping. Natural Hazards, 1–20. https://doi.org/10.1007/s11069-017-2757-y (2017).

Tanyas, H., van Westen, C., Allstadt, K., Nowicki Jessee, A., Görüm, T., Jibson, R. W., … Hovius, N. Presentation and Analysis of Earthquake-Induced Landslide Inventories. Journal of Geophysical Research: Earth Surface, 122, 1991–2015. https://doi.org/10.1002/2017JF004236 (2017).

Tanyaş, H., & Lombardo, L. Variation in landslide-affected area under the control of ground motion and topography. Engineering Geology, 260, 105229. https://doi.org/10.1016/j.enggeo.2019.105229 (2019).

Tanyas, H., Rossi, M., Alvioli, M., van Westen, C. J., & Marchesini, I. A global slope unit-based method for the near real-time prediction of earthquake-induced landslides. Geomorphology, 327, 126-146. https://doi.org/10.1016/j.geomorph.2018.10.022 (2019).

Tibshirani, R. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society, Series B (Methodological), 58(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x (1996).

Van Den Eeckhaut, M., & Hervás, J. Geomorphology State of the art of national landslide databases in Europe and their potential for assessing landslide susceptibility, hazard and risk, 140, 545–558. https://doi.org/10.1016/j.geomorph.2011.12.006 (2012).

Van Den Eeckhaut, M., Hervás, J., Jaedicke, C., Malet, J.-P., Montanarella, L., & Nadim, F. Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data. Landslides, (9), 357–369. https://doi.org/10.1007/s10346-011-0299-z (2012).

van Westen CJ (Photographer), Zhang J (Photographer). Landslides and floods triggered by Hurricane Maria (18 September, 2017) in Dominica. UNITAR-UNOSAT. (2018).

Weiss, A. Topographic Position and Landforms Analysis. In ESRI User Conference, San Diego, CA. (2001).

Williams, J. G., Rosser, N. J., Kincey, M. E., Benjamin, J., Oven, K. J., Densmore, A. L., Milledge, D. G., Robinson, T. R., Jordan, C. A., and Dijkstra, T. A.: Satellite-based emergency mapping using optical imagery: experience and reflections from the 2015 Nepal earthquakes, Nat. Hazards Earth Syst. Sci., 18, 185–205, https://doi.org/10.5194/nhess-18-185-2018, 2018.

Xu, C., Dai, F., Xu, X., & Hsi, Y. GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. Geomorphology, 145–146, 70–80. https://doi.org/10.1016/j.geomorph.2011.12.040 (2012).

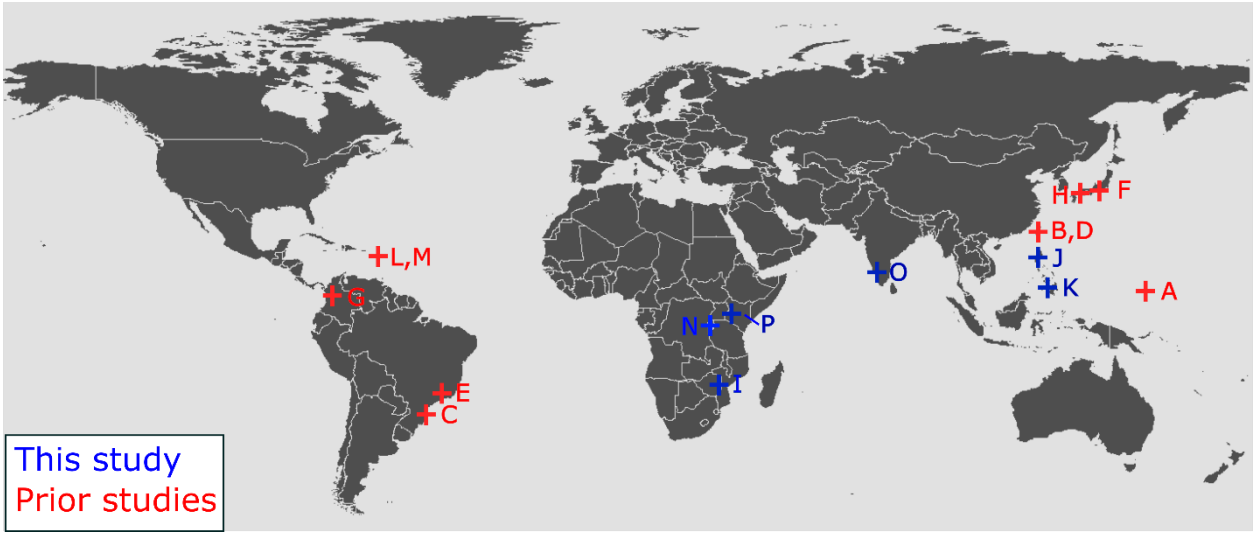


Figure 1: Locations of landslide inventories considered in this study. Locations labelled in red have been published previously, while those in blue are presented for the first time here. Satellite images of the newly mapped landslide inventories can be found in the Supplemental material. Table 1 contains the details of each of the inventories, organised in alphabetical order.

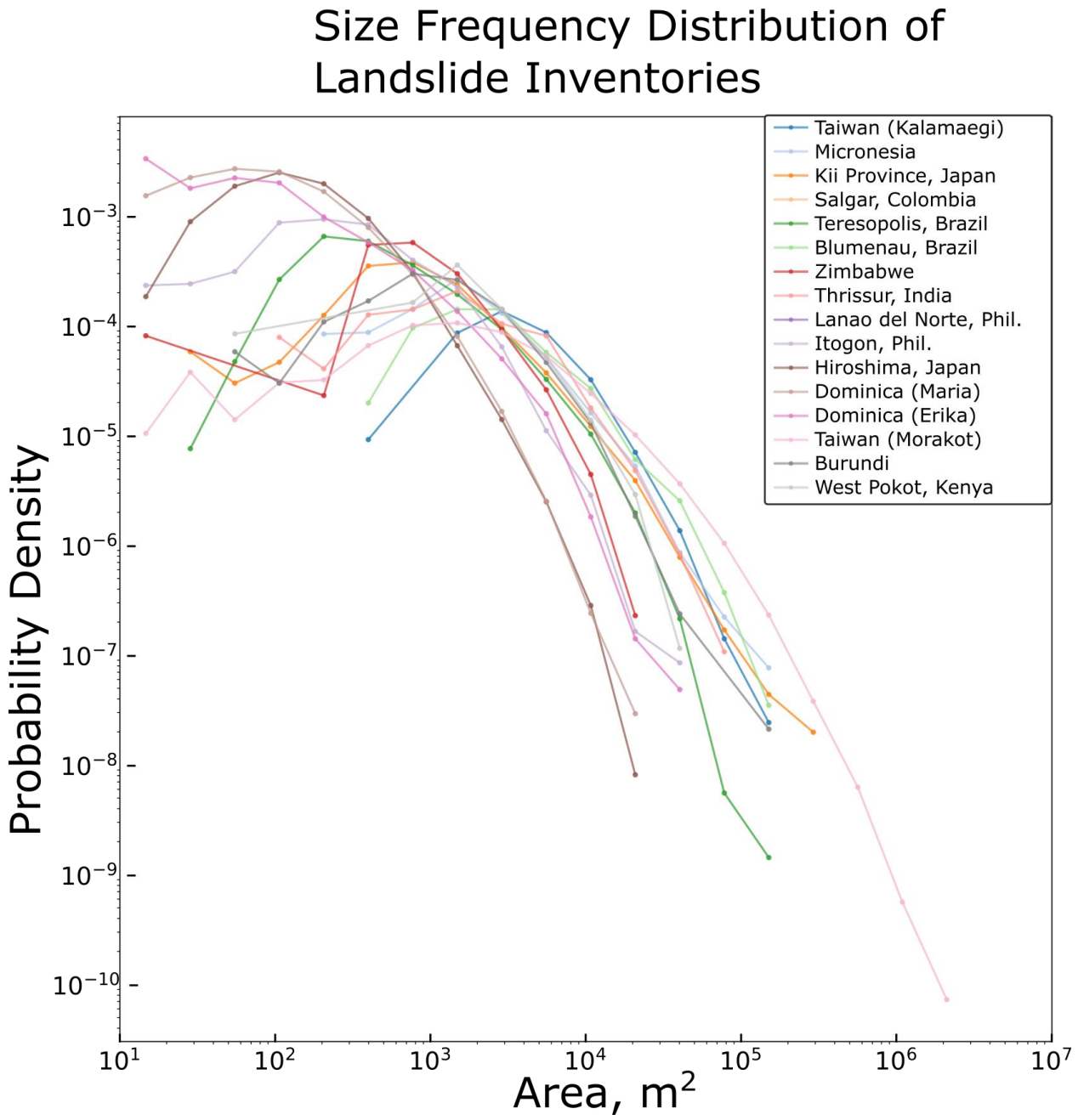
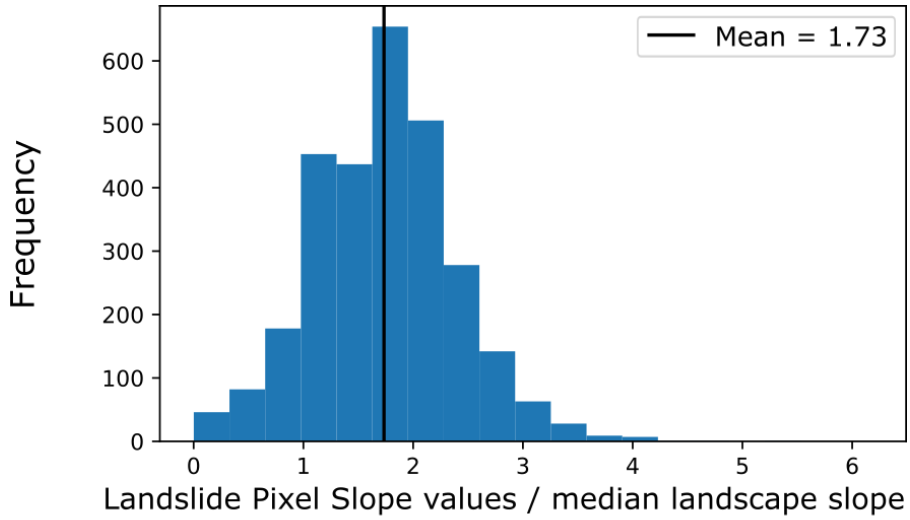
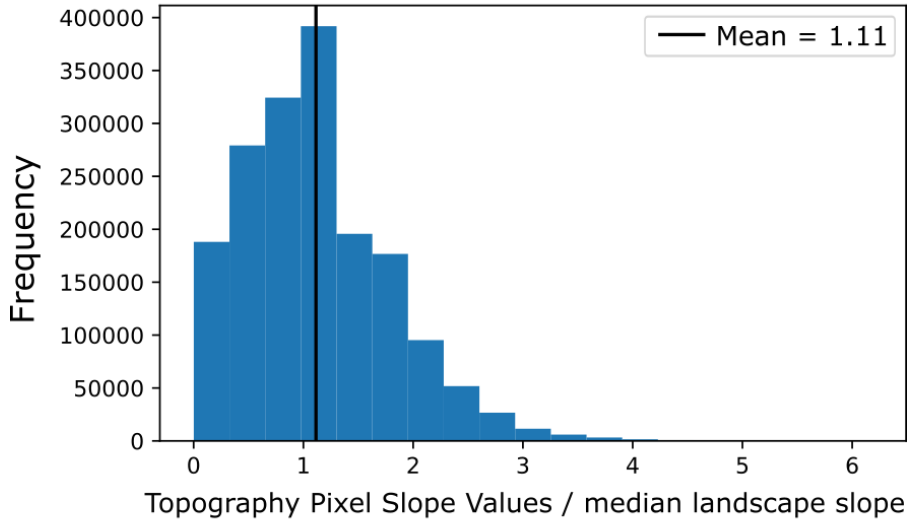


Figure 2: Probability density for landslides in each event inventory, obtained as the number of landslides with areas falling into logarithmic bins (consistent bins for all inventories), $N_{[A:A+dA]}$, normalized by the bin width, dA, and the total number of slides for the event, N_{tot} (see Malamud et al., 2004).


Mean = 1.73
Frequency
600
500
400
300
200
100
0
0
1
2
3
4
5
6
Landslide Pixel Slope values / median landscape slope


Mean = 1.11
Frequency
400000
350000
300000
250000
200000
150000
100000
50000
0
0
1
2
3
4
5
6
Topography Pixel Slope Values / median landscape slope

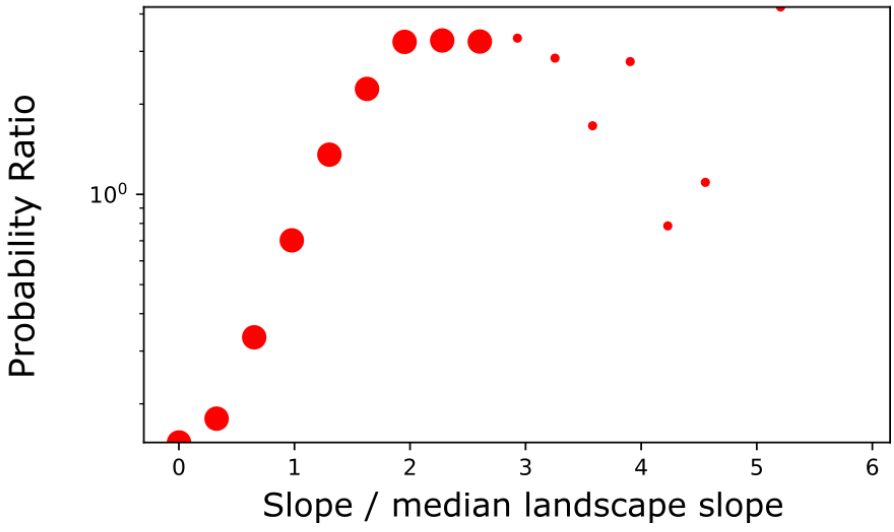

Probability Ratio
10^0
0
1
2
3
4
5
6
Slope / median landscape slope

Figure 3: Example of landslide – topography ratio comparison for landslides in Zimbabwe triggered by Cyclone Idai. This shows the distribution of values for landslides (A) and topography (B) for slope. The lower part of the figure (C) shows the relative ratio of the two distributions for the parameter. The black lines in parts A and B illustrate the mean value of the data. In the lower figure (C), the size of the points depends on whether the difference between landslide and topography exceeds the calculated confidence interval for that bin interval; larger points indicate significant deviation (p>0.95), and small points indicate the difference is not significant at that probability level.

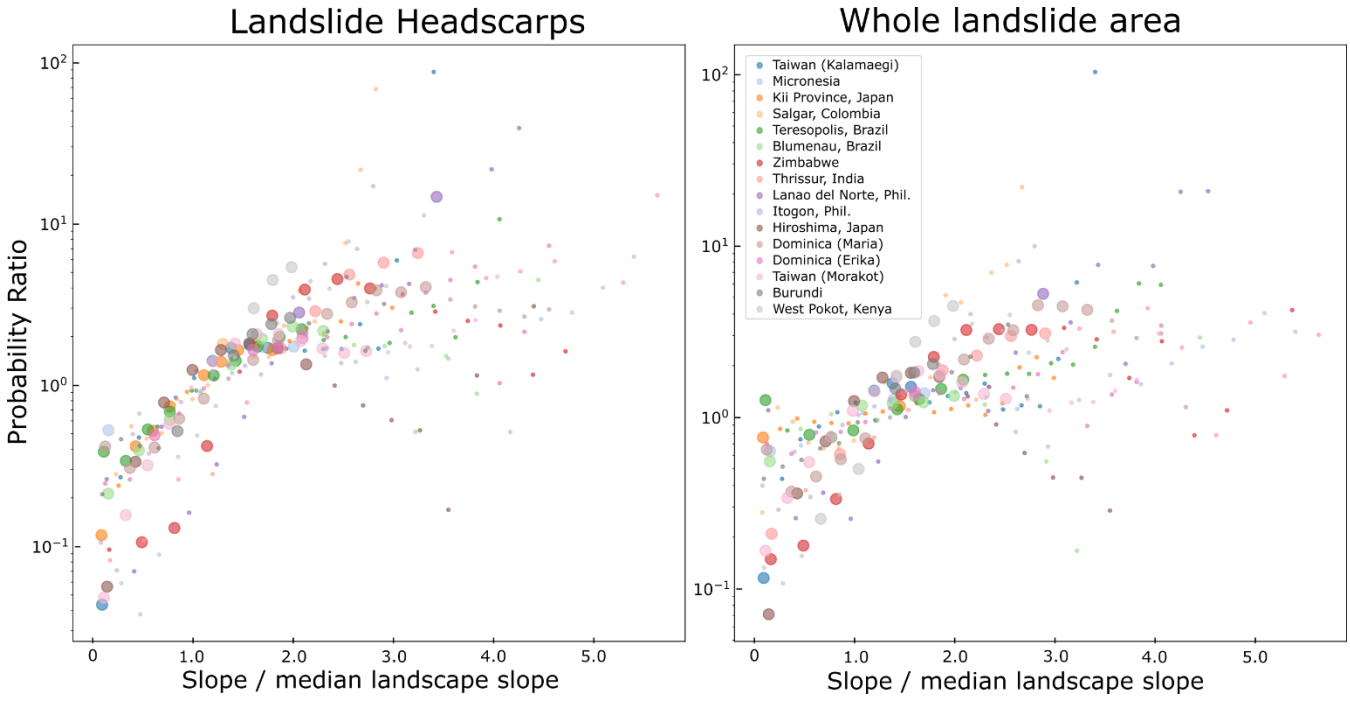


Figure 4: Landslide probability~~likelihood~~ ratio against the slope normalized by the median of the local landscape, for the scar area (left) and the whole landslide area (right). The size of the points depends on whether the difference between landslide and topography exceeds the calculated confidence interval for that bin interval; larger points indicate significant deviation (p>0.95), and small points indicate the difference is not significant at that probability level.

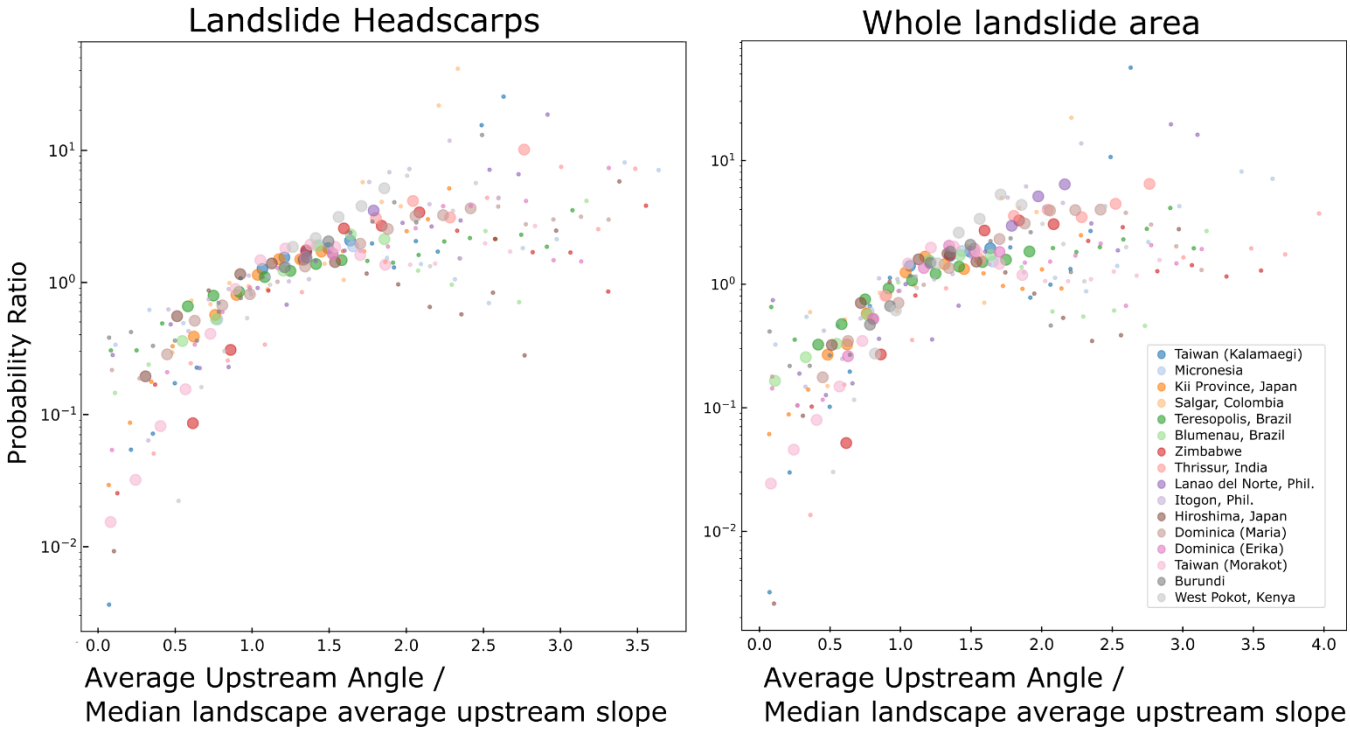


Figure 5: Landslide probability~~likelihood~~ ratio against the average upstream angle normalized by the median of the local landscape, for the scar area (left) and the whole landslide area (right). The size of the points depends on whether the difference between landslide and topography exceeds the calculated confidence interval for that bin interval; larger points indicate significant deviation (p>0.95), and small points indicate the difference is not significant at that probability level.

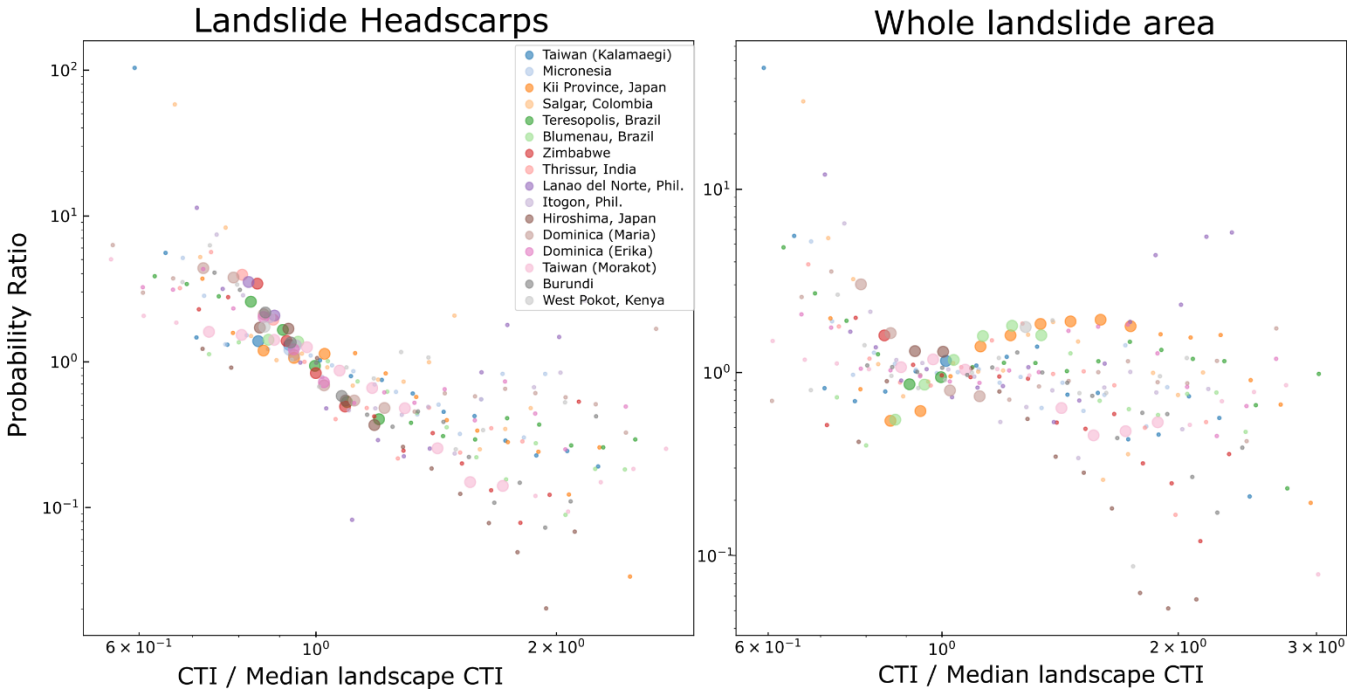


Figure 6: Same as Figure 4 but for the Compound Topographic Index (CTI). Note that the x-axis is a logarithmic scale.

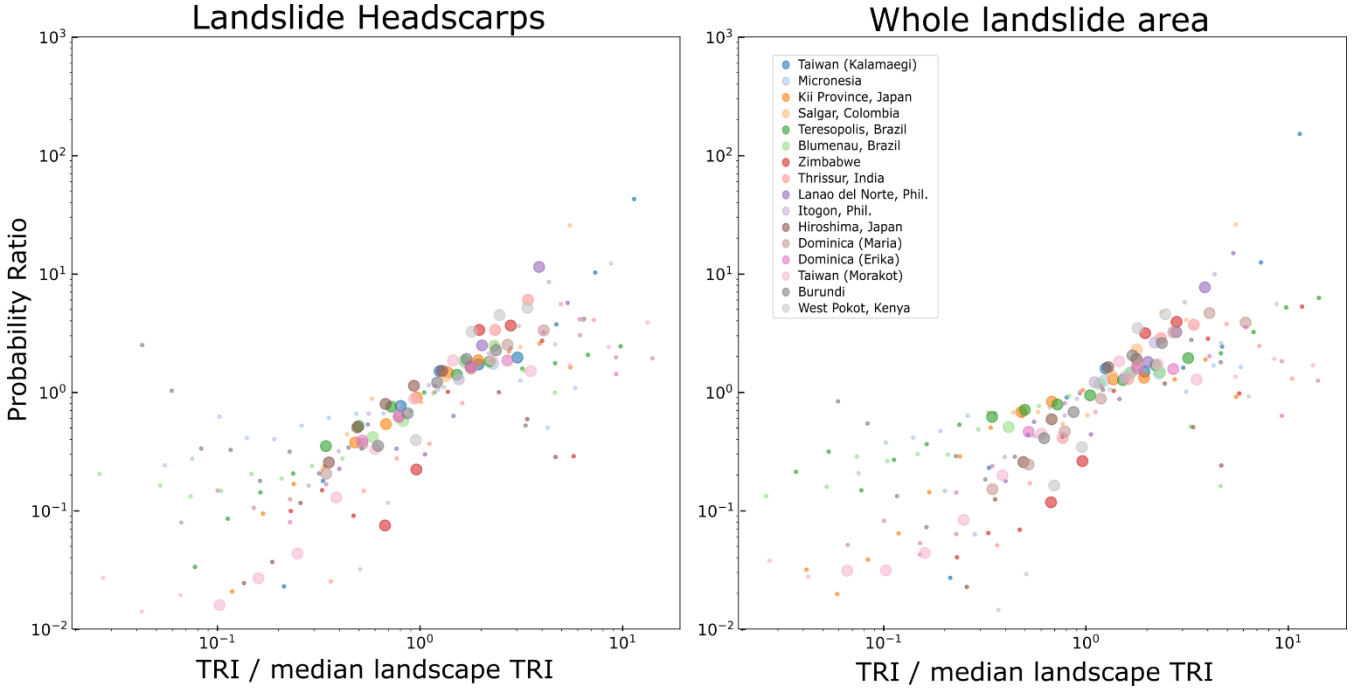
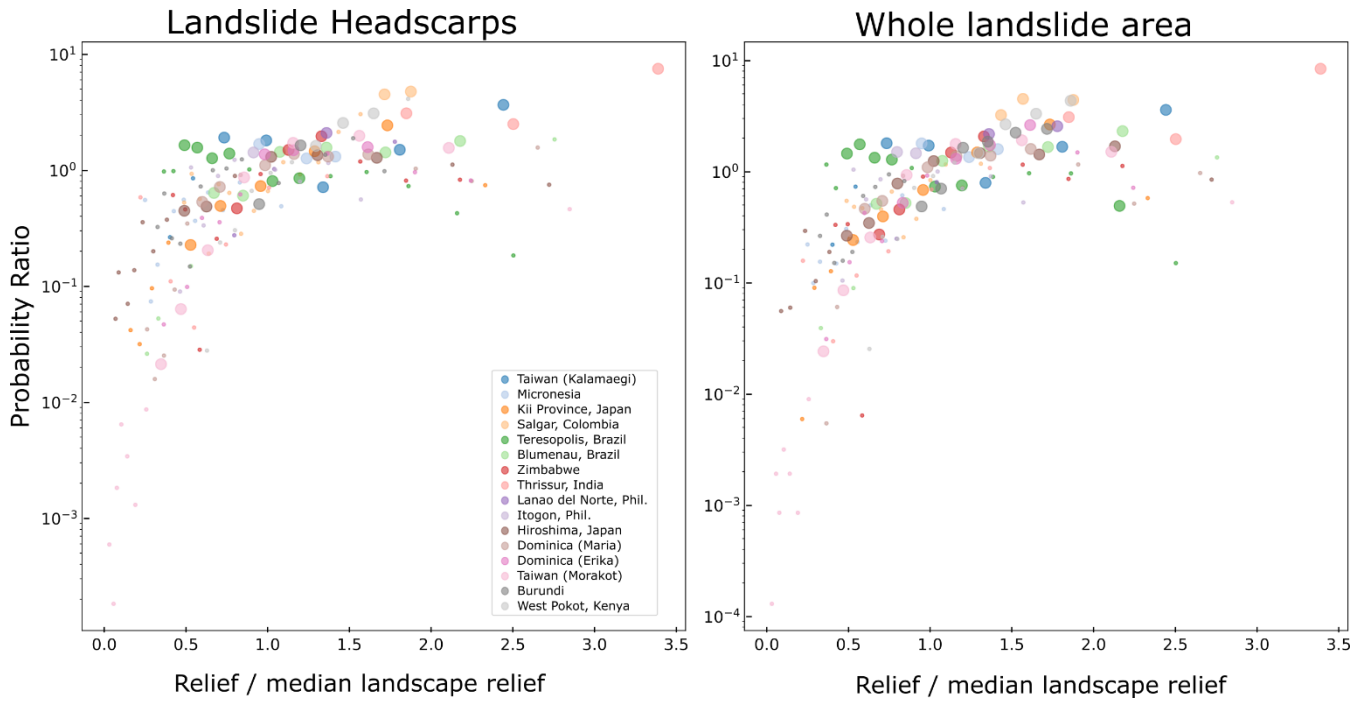


Figure 7: Same as Figure 4 but for Topographic Ruggedness Index (TRI).



870

Figure 8: Same as Figure 4 but for Relief in 1km radius of each cell.

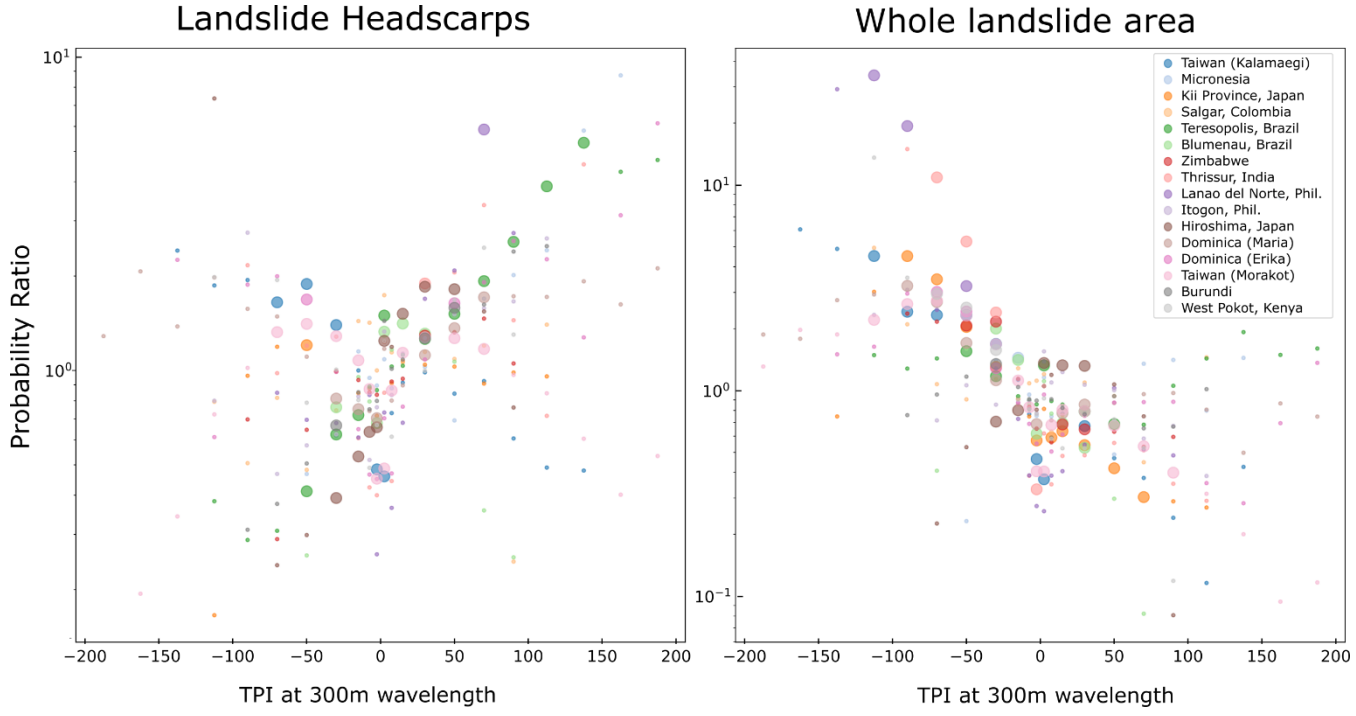


Figure 9: Same as Figure 4 but for Topographic Position Index (however, here we omit normalization given that TPI is a zero centered variable) with an analysis window radius of 300m.

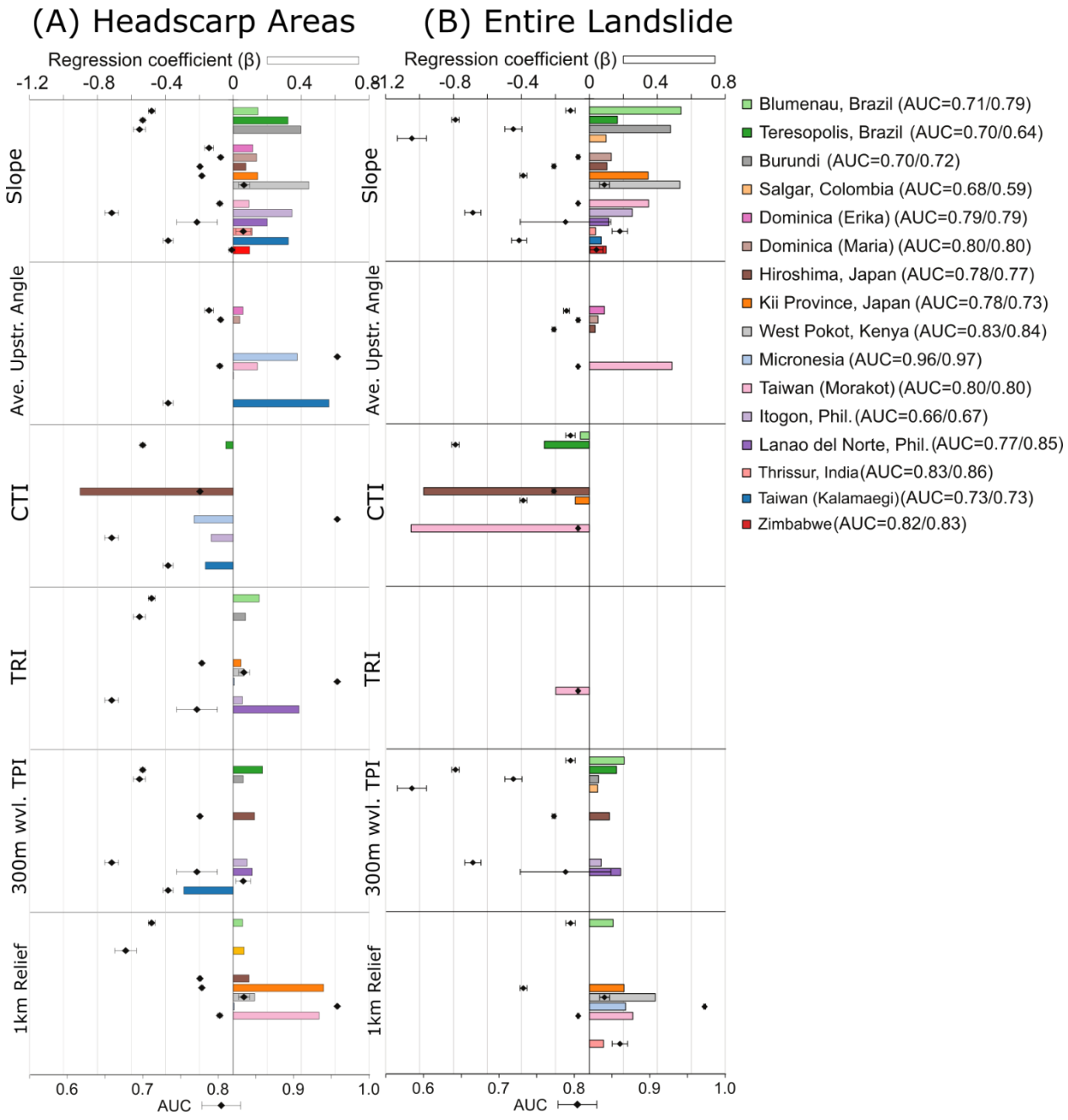


Figure 10: Figure showing regression coefficients and corresponding AUC values calculated from fitted GLM for each landslide inventory, considering scar (a) or whole landslide area (b). The error bar around corresponds to the variation in ~~95% confidence intervals of~~ the AUC distribution across 10-fold cross-validation replicates. Regression coefficient are not shown for predictors considered not significant by the LASSO methods. The AUC values shown in the legend are for ~~headscarp~~scars (prior to slash) and entire landslides (second value).

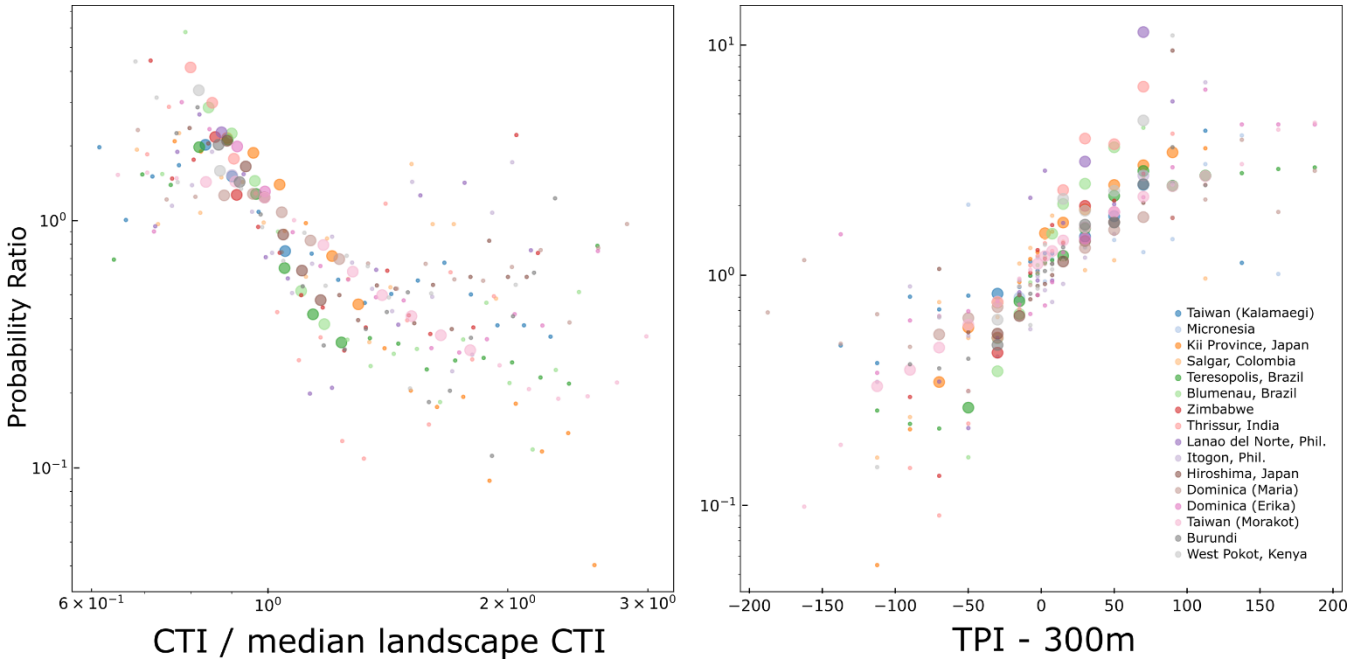


Figure 11: Probability~~Likelihood~~ ratio of landslide ~~headscarp~~scar areas compared with entire landslide area. It is important to note that while figures 3-9 contrast the landslide areas (scar or entire mapped area) with the topography, this figures shows the ratio of probabilities for scar and whole landslide area. Specifically, this shows the probability of a scar at a given CTI value, normalized by the probability of that CTI value, divided by the probability of the entire landslide at a given CTI value, normalized by the probability of that CTI value. Higher values indicate that ~~headscarp~~scars are more likely than whole landslide areas, at that parameter value. Left figure shows the result for CTI; right figure for short wavelength TPI.