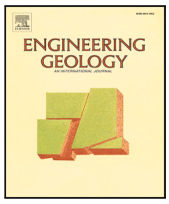
Engineering Geology 345 (2025) 107866

Contents lists available at ScienceDirect

Engineering Geology

journal homepage: www.elsevier.com/locate/enggeo



Towards automatic delineation of landslide source and runout

Kushanav Bhuyan [c,d,*], Kamal Rana [e], Ugur Ozturk [f,g], Lorenzo Nava [a,b,c], Ascanio Rosi [c], Sansar Raj Meena [c], Xuanmei Fan [d,**], Mario Floris [c], Cees van Westen [h], Filippo Catani [c]

[a] Department of Earth Sciences, University of Cambridge, Cambridge, UK
[b] Department of Geography, University of Cambridge, Cambridge, UK
[c] Machine Intelligence and Slope Stability Laboratory, Department of Geosciences, University of Padova, Padova, 35129, Veneto, Italy
[d] State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, People's Republic of China
[e] Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, 14623, NY, USA
[f] Helmholtz Centre Potsdam - GFZ German Research Centre for Geosciences, Potsdam, 14473, Brandenburg, Germany
[g] Institute of Environmental Science and Geography, University of Potsdam, Potsdam, 14473, Brandenburg, Germany
[h] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Twente, The Netherlands

ARTICLE INFO

Dataset link: https://github.com/kushanavbhuyan/Delineating-failure-kinematics

Keywords:
Landslide source
Landslide runout
Topology
Morphology
Topological data analyses
Volumes

ABSTRACT

Mapping landslide-depleted source areas is pivotal for refining predictive models and volume estimations, yet these critical regions are often conflated with landslide runouts, leading to sub-optimal assessments. The source (or scarp) areas are typically the regions where the actual failure occurs, providing crucial information on the initiation mechanisms and the nature of landslide propagation. Catering to this objective, we built a method based on a landslide's topology and morphological information to delineate the source and runout margins. We develop and test this method in geomorphologically distinct regions such as Dominica, Turkey, Italy, Nepal, and Japan (Niigata) to showcase the model's robust adaptive capacity. The model can demarcate the source and runout zones from landslide planforms found in inventories with accuracy deviations under 15%–20%. While distinguishing landslide source and runout areas, the model also considers triggering information and movement types. We also deploy the model in Chile, Japan (Hokkaido), Colombia, Papua New Guinea, and China. In these new regions, we found the mean area of the scarp to be consistently under 30% of the total landslide area. We additionally showcased the application of our model to the area–volume scaling of the coseismic landslides triggered by the 2018 Hokkaido Eastern Iburi Earthquake (M_W 6.6) in Japan. Our analysis revealed that area–volume fitting using the landslide source areas instead of the total landslide planforms or polygons improves the linear fit from R^2=0.49 to R^2=0.81. Our work could improve diverse landslide analysis, such as hazard and runout models, and facilitate a deeper understanding of landslide behaviour.

1. Introduction

Existing landslide inventories insufficiently document kinematic zones of the recorded landslides, i.e., source and runout/deposition areas (Van Westen, 2016). Most inventories typically record only essential data such as the landslide's geographical coordinates (Martha et al., 2021), at times, also date occurrence (Behling et al., 2016) especially if the data is linked to an event, such as an earthquake (Tanyaş et al., 2017). In exceptional cases, types of movement, such as flows, slides, and rock falls, are also documented (Trigila et al., 2007) or can be estimated if landslide polygons are available (Amato et al., 2021; Bhuyan et al., 2024).

The mechanical properties of landslide sources (or depletion zones) and runout or accumulation areas are inherently different. The source area, defined as the surface of the rupture, is controlled by a complex interplay of factors such as soil composition, pore water pressure, and geological features (Crosta et al., 2003). As a function of the hydraulic and slope gradient, the runout area is more concerned with transporting landslide mass and its eventual deposition and accumulation (Meyrat et al., 2022).

The lack of separation impinges on the reliability and precision of subsequent analytical tasks, most notably in susceptibility and hazard assessments (Ozturk et al., 2021; Dahal et al., 2024b). Such predictive models, which ascertain the probability of future landslide events, are

* Corresponding author at: Machine Intelligence and Slope Stability Laboratory, Department of Geosciences, University of Padova, Padova, 35129, Veneto, Italy.
** Corresponding author.
E-mail addresses: kushanav.bhuyan@phd.unipd.it (K. Bhuyan), fxm_cdut@qq.com (X. Fan).

https://doi.org/10.1016/j.enggeo.2024.107866
Received 17 July 2024; Received in revised form 8 December 2024; Accepted 9 December 2024
Available online 19 December 2024

Table 1
Landslide data and relevant information. Additional information into their usage in model development is also provided.

Study areas	Triggers	Number of landslides[a]	Types of landslide failures	Mapping scale/resolution	Purpose
Dominica	Rainfall, 2017	~9900	Debris slides and debris flows	Mapped in 0.5 m resolution	Proof-of-concept (training and testing)[b]
Nepal	Earthquake, 2015	~20,000	N/A	Not mentioned	Proof-of-concept (training and testing)[b]
Turkey, Bartin	Historical	~1200	Slides, Earth flows, and Slide evolving to earth flows	1:25000	Proof-of-concept (training and testing)[b]
Italy, Belluno	Rainfall, 2018	42	N/A	1:6000	Adaptability (with ground-truth)[b]
Japan, Niigata	Earthquake, 2004	~200	N/A	Not mentioned	Adaptability (with ground-truth)[b]
Japan, Hokkaido	Earthquake, 2018	~560	N/A	Not mentioned	Real-world deployment (no ground-truth)
Papua New Guinea	Earthquake, 2018	~1150	N/A	Mapped in 3 m-resolution	Real-world deployment (no ground-truth)
Colombia	Earthquake, 2019	~800	Slides and Flows	Not mentioned	Real-world deployment (no ground-truth)
Chile	Historical	~440	Debris flows	Mapped in 0.5 m-resolution	Real-world deployment (no ground-truth)
China, Wenchuan	Earthquake, 2008	~8900	Slides and Flows	Mapped in sub-metre resolutions	Real-world deployment (no ground-truth)

[a] All landslides were pre-processed to remove faulty polygons. Faulty polygons may include: single source/scarp polygons with no runout, polygons under 1000 m^2, and multi-polygonal geometry.
[b] These inventories include the separated zones of source and runout, enabling model testing and evaluation of the predictive accuracy.

predominantly contingent upon the geomorphological and geotechnical characteristics of the source area instead of the entire landslide region (Keck et al., 2023). Additionally, a lack of distinct demarcation between these zones could lead to erroneous estimations of the velocities and trajectories (Marinelli et al., 2022), affected areas (Dahal et al., 2024a), and volumes (Yunus et al., 2023; Ju et al., 2023), thereby compromising the efficacy of mitigation strategies. For instance, applying the same preconditioning factors across the landslide source and runout areas in statistical models can result in flawed volume estimates, consequently impacting calculations of sediment transport capacity, deposition patterns, and mass balance assessments (Francis et al., 2022; Hovius et al., 2011; Parker et al., 2011; Yunus et al., 2023), along with kinematic trajectory simulations and hazard quantification (Keck et al., 2023). Recent studies have also demonstrated the use of power-law models to establish volume-related correlations, providing insights into the empirical relationships between landslide volume and source characteristics (area, width, length, and depth) for engineering applications (Ju et al., 2023). But again, information regarding the source characteristics is required for developing such models, which can be challenging to obtain in regions where landslide inventories are incomplete or lack sufficient precision. Thus, the source area separation is not merely a taxonomic exercise but has significant implications for a gamut of landslide studies, from fundamental geotechnical analyses to advanced hazard and risk assessments.

Previous efforts to delineate the kinematic zones of landslides initially involved manually examining traces or signatures of transitions between the source and runout areas (Guzzetti et al., 2012), besides detailed geomorphological and geotechnical field surveys. Subsequently, semi-ellipsoidal approximations were used where the scarp region was defined using empirical measurements with a major axis length of say, 'x' metres (Cruden and Varnes, 1996). However, this region-specific approach has the risk of generalizing complex morphologies into simpler geometrical forms, potentially losing crucial morphological details (such as terrain steepness and scarp deformations). More recently, a preliminary machine learning-based attempt involved image segmentation of satellite imagery (Lai, 2020). However, spectral responses in remote sensing images often obscure the fine delineation between different kinematic zones. Hence, both the semi-ellipsoidal and image segmentation approaches can lose essential morphological information, such as surface rupture, either by oversimplifying the landscape or failing to capture the subtle differences accurately. Most importantly, both approaches are also very limited in terms of providing a scaleable solution, applicable in other regions. To date, no other studies besides those mentioned have provided a comprehensive framework for delineating the transition zones between landslide source and runout areas. Importantly, no existing work has integrated a morphological framework that accounts for the varying dimensions and features of different landslides to address this issue. Such a framework would be well-suited to capture the nuanced characteristics of landslides, offering a scalable and adaptable solution across diverse terrains.

Incorporating morphological information of landslides obtained via their shapes, such as perimeter and area (Taylor et al., 2018), into automated regimes, such as tree-based learning algorithms, have already shown promising results in classifying landslide-triggering mechanisms (e.g., Rana et al., 2021, 2022). More advanced approaches extracting topological properties of the landslides beyond mere geometric information (Amato et al., 2021) have even led to accurately classifying failure movement types of landslides (Bhuyan et al., 2024). This underscores the potential for utilizing detailed landslide polygonal information to gain deeper and more nuanced insights into landslide zones. However, their utility in providing affluent solutions to separate the zones within landslide propagation has not been explored, yet.

In this study, we use landslide inventories from the literature, encompassing both event-based triggers (rainfall and earthquake) and historical inventories. We develop a data-driven model guided by morphological and topological information that automatically delineates the landslide kinematic zones, i.e., source and runout. We combine topology, a branch of mathematics that deals with transforming shapes into a high-dimensional feature space, with standard landslide morphological attributes such as travel distance, angle, height, and empirically estimated velocities. Note that the terms 'scarp', 'source zones', and 'source area' will be used interchangeably throughout the paper. We perform three experiments to gauge the model's performance. First, we choose a set of regions termed 'proof-of-concept' regions upon which the model is designed. Second, we test its performance on two unseen

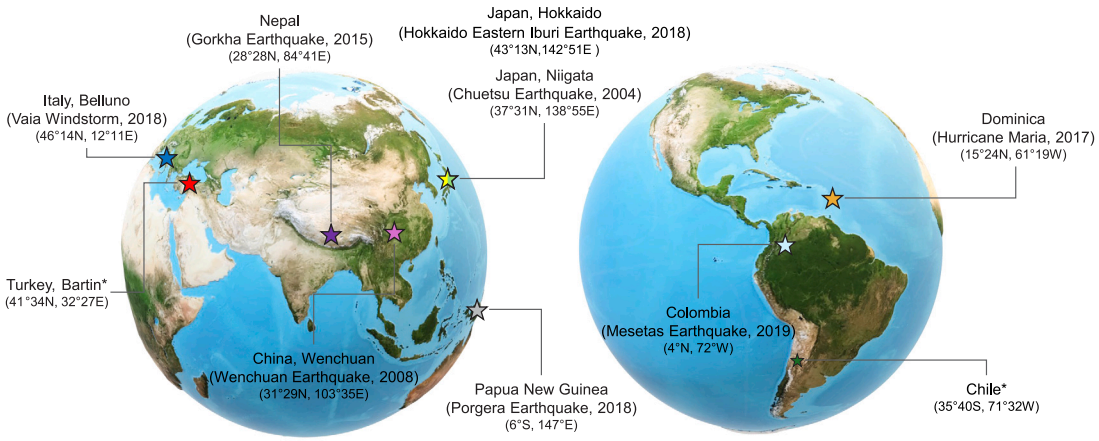


Fig. 1. 3D perspective globe maps for the respective study areas marked by their triggering events and centroid coordinates. * refers to historical inventories where pre-dominant triggers are unknown.

regions using the model trained on the 'proof-of-concept' regions. Finally, the third experiment is driven more along the lines of practical usage in real-case scenarios where the model is applied to regions (inventories) with no ground truth. Aside from these experiments, we also demonstrate area–volume scaling for the region of Hokkaido, Japan which was struck by the M_W 6.6 earthquake in 2018. The underlying contribution of this work is the development of a method that combines topological and morphometric data with machine learning to automate large-scale delineation of landslide source and runout zones, enabling the documentation of previously undocumented inventories and enhance landslide mapping efforts. This method is adaptable across diverse terrains, making it highly valuable for both regional and large-scale applications, especially in hazard and runout modelling, and volume estimations.

2. Landslide databases

We have selected five distinct case study areas from Dominica (van Westen and Zhang, 2018), Italy (Belluno) (Meena et al., 2022), Nepal (Roback et al., 2018), Japan (Niigata) (Sekiguchi and Sato, 2006), and Turkey (Bartin) (Görüm, 2019) with different landslide triggers and geomorphic/climatic settings for model development (Table 1). The subsections below highlight the diversity of the study areas to ensure thorough testing of the model's robustness and applicability and help demonstrate its versatility. These case studies have been previously documented in the literature, underscoring their importance in analysing landslide behaviour. As such, we discuss these five case studies moving forward. Following model development and validation, we used other sets of inventories such as from Chile, Colombia, Japan-Hokkaido, Papua New Guinea, and China-Wenchuan, for deployment. Details of these inventories are in Table 1 and their locations can be seen in Fig. 1.

The statistics of the morphological properties for the compiled landslide inventories show variations (Figs. 2 and 3, see further explanation of the properties in Section "Topo-morphological landslide properties"). Rainfall-induced landslide inventories such as those from Dominica consist of relatively smaller landslides than co-seismic inventories from Nepal and Japan (Niigata) (Fig. 2-A). At the same time, despite similar triggers, properties such as the source-to-total length ratio show considerable variations due to differences in lithologies and climate (Figs. 2-B, 3). Overall, these regions and their respective landslides are quite diverse, which is reflected in the unique distributions of these exploratory statistical plots.

2.1. Dominica: Hurricane Maria 2017

Geologically, the young volcanic Caribbean island of Dominica hosts nine volcanic complexes and is primarily composed of volcanic rocks (Rouse, 1990). The island's rugged terrain features peaks over 1400 m AMSL and many hillslopes are steeper than 30°. Dominica receives a particularly high amount of rainfall at higher elevations, with the central peaks receiving over 7500 mm annually (Benson et al., 2001). Four main soil types—smectoid clays, kandoid clays, allophane latosolics, and allophane podzolics—constitute the lithology of the island, largely distributed according to rainfall variability (Rouse et al., 1986; Rouse, 1990). Smectoid clays with high shear strength and soil impermeability exhibit low landslide activity in less rainy areas. Given their strong soil/rock interface properties, Kandoid clays and allophane latosolics require a high amount of rainfall to induce landslides.

Prone to shallow translational slides, Allophane podzolics demand exceptionally intense rainfall for slope failure. This varied soil composition and response to rainfall explain the differing landslide frequencies across the island. For example, Hurricane Maria triggered 9900 landslides in 2017, mostly debris slides followed by debris flows, affecting more than 1.3% of the island's area. The hurricane caused extensive damage, totalling USD 930.9 million, with the transportation, housing, and agriculture sectors suffering the most significant impacts (Government of the Commonwealth of Dominica, 2017). Following Hurricane Maria, the Global Facility for Disaster Reduction and Recovery funded the CHARIM (Caribbean Handbook on Risk Information Management) initiative (CHARIM, 2018). This project aimed to develop an extensive natural hazard evaluation and as part of this effort, a detailed landslide inventory covering the entire island was created by the University of Twente in the Netherlands (van Westen and Zhang, 2018). This inventory was made available through UNITAR-UNOSAT (https://www.unitar.org/).

2.2. Italy, Belluno: Vaia Windstorm 2018

The province of Belluno is located in northeastern Italy. The province's altitude ranges from 42 to 3325 m a.s.l. featuring the Dolomite Alps and Venetian Prealps. The Dolomite Alps contain a Hercynian crystalline basement with micaschists, phyllites, and Permian ignimbrites, whereas the Venetian Prealps are characterized by Jurassic–Cretaceous sedimentary formations (Doglioni, 1990). This province is prone to slope instability, affecting about 6% of its area. Landslides are predominantly located in the northwestern and southeastern sectors, with the dominant types being slides, rapid flows, slow flows, and shallow soil slips (Iadanza et al., 2021). Belluno's continental climate, with a mean annual temperature of 7 °C and precipitation

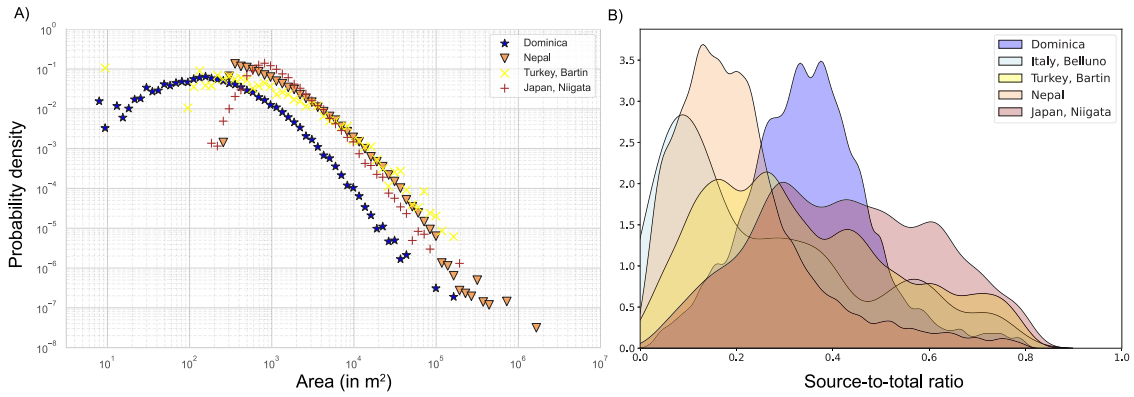


Fig. 2. The landslide inventories' (A) size frequency (for Dominica-Rainfall, Nepal-Earthquake, Turkey-Historical, and Japan (Niigata-Earthquake) and (B) the probability density distributions. The probability density distribution is with respect to the source-to-total landslide length ratio. The data for Italy, Belluno contains only 40 landslides and, thus, is not shown in (A).

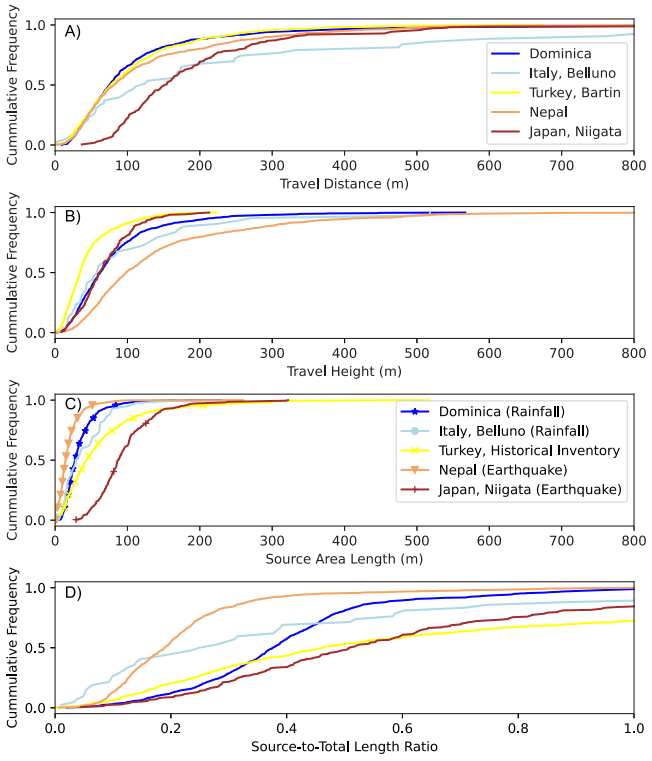


Fig. 3. Morphological properties of the landslide inventories, (A) travel distances, (B) travel heights, (C) the source area lengths, and (D) source-to-total length ratios.

of 1284 mm, peaks in spring and autumn. The past 27 years have increased temperature and rainfall intensity due to climatic changes, elevating the frequency of landslides (ARPAV, 2021). For example, the Vaia Windstorm of 2018 caused tremendous damage to forest areas and was responsible for triggering over 500 landslides in the Belluno province. Out of the 500 landslides (Meena et al., 2022), we manually delineated the source and runout areas for 42 landslides using aerial photography.

2.3. Nepal: Gorkha earthquake 2015

Nepal is located along the Himalayan mountain belt and experiences high seismic activity. The region's steep terrain, intense monsoon rainfall, and dense population exacerbate the landslide risk, often triggered by seismic activity and heavy rainstorms. For instance, the M_W 7.8 Gorkha earthquake on April 25, 2015, affected a 140 km segment of the Main Himalayan Thrust, causing around 20,000 landslides without a surface break (Roback et al., 2018). This event also demonstrated a high density of landslides in areas with heavy rainfall, indicating that climate might influence subsurface weathering, particularly in the drier, higher elevations (Gallen et al., 2015). The 2015 event resulted in nearly 9000 deaths and extensive economic damage. The earthquake also created more than 69 landslide dams, posing additional risks to downstream populations. Co-seismic landslides were uniformly distributed throughout the major lithological sequences of both the Greater and Lesser Himalayan rock formations, with no particular geological units at a regional scale showing a concentration of landslides. However, at a local level, specific sub-units like "Other Proterozoic Crystalline" and "Other Metamorphic" experienced a greater frequency of landslides, especially near the epicentre of the earthquake (Roback et al., 2018). The inventory used in this study was obtained from Roback et al. (2018) which included over 20,000 landslides with the source and runout areas separated.

2.4. Japan, Niigata: Chuetsu earthquake 2004

Located along the western edge of the Pacific Ring of Fire, South-North elongated Japan has a subarctic climate in the north with warm summers and snowy winters. Southern Japan exhibits a more subtropical climate with humid summers and moderate and rainy winters. The Japanese islands consist of various geological units. Those facing oceanic plates are younger, with more volcanic material, while those facing the Sea of Japan are mainly faulted and folded sedimentary deposits. The country is among the world's most geologically active regions, frequently affected by earthquakes and extreme rainfall events. For example, the Mid-Niigata Prefecture earthquake (M_W 6.8) in October 2004 hit the western coasts of the main island, Honshu (Sassa, 2005), triggering 362 landslides. Most of the landslides induced by the Mid-Niigata Prefecture earthquake occurred within the sand, silty sand, or silt layers. The inventory was obtained from Sekiguchi and Sato (2006), comprising over 200 distinct post-processed source and runout zones of the respective landslides.

2.5. Turkey, Bartin: Historical inventory

Located in Turkey's rugged Western Black Sea region, the Bartin's topography is characterized by elevations ranging from 195 m in the river floodplains to 900 m (Görüm, 2019). Hillslopes vary from a gentle 0° to a steep 78°. The area is primarily composed of two sedimentary lithologic units. The first comprises Quaternary alluvial deposits found in valley fills, while the second consists of Cretaceous sandstone and mudstone formations, the primary substrates for landslide occurrences (Duman et al., 2005).

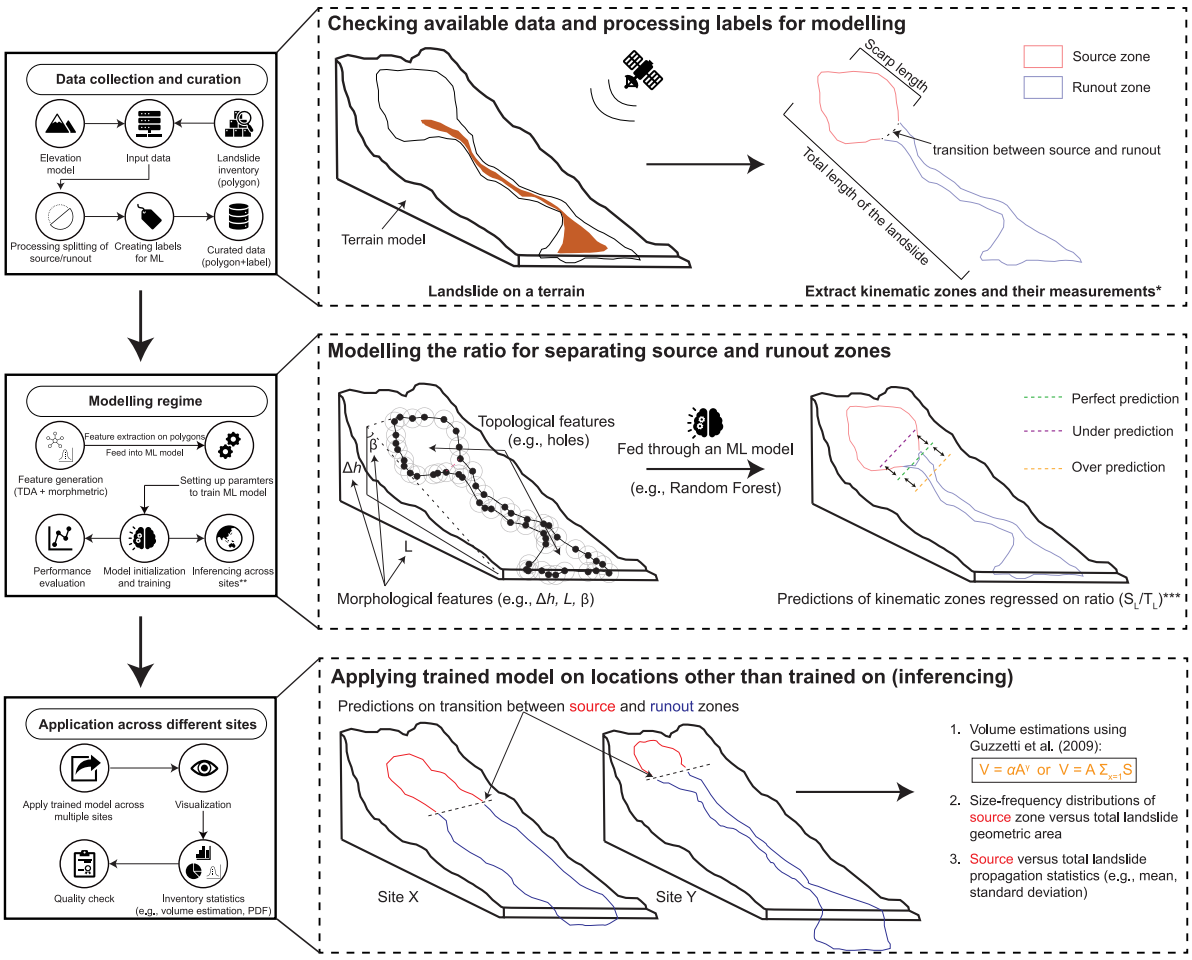


Fig. 4. Flowchart of the methodology adopted in the study. The flowchart depicts phases of data collection and curation, the modelling regime, and the eventual application of the trained model on other sites for inferencing (locations where the model has not been trained). Δh, L, and β indicate the travel height, travel distance, and travel angle, respectively. Area–volume scaling is based on empirical relationships based on equations introduced for global applications (Guzzetti et al., 2009). * The kinematic zones are already divided in the data obtained from the literature. Here, we merely perform quality checks on the inventories and extract information such as respective area, travel height, distance of travel, etc. ** Inferencing refers to the application or usage of a trained model in a location where the model has never seen the region. *** The ratio refers to the label that is fed to the regression-based model for the training regime (see Fig. 7).

The region, characterized by a maritime temperate climate, receives an annual average rainfall of 980 mm, peaking between October and December. The lowest mean monthly rainfall is in August, 55 mm, and the highest is in December, 125 mm. Landslide incidents are predominantly triggered by abrupt and intense, or prolonged rainfall (Can et al., 2005), with the highest frequency of landslide occurrences between May and August, a period not typically associated with the high rain accumulations. Most landslides (68%) are situated within forest-covered terrain. Within the specific test area downstream of the Ulus River, forest stand densities range from low (<10%) to medium (40%–70%). The landslide inventory in Turkey was sourced from Görüm (2019), documented in 2019. These landslides were charted using airborne Light Detection And Ranging (LiDAR) data, encompassing a total of 900 landslides, categorized based on the system developed by Cruden and Varnes (1996).

3. Methods

3.1. Topo-morphological landslide properties

The problem of identifying the source area from the entire landslide polygon is rooted in the landslide's morphology via its polygonal shape (both two- and three-dimensional). Topology measures the complex shapes of data using techniques from algebraic topology (Rana et al., 2022; Bhuyan et al., 2024), while morphology measures the morphometric characteristics of the landslide, allowing a detailed analysis of how the landslides evolve spatially and dynamically. Hence, we adopted these frameworks to extract the topological and morphometric information of landslides. Then, we used these properties to train a machine learning algorithm to address a regression challenge—separating source and runout zones of a landslide (see Fig. 4).

The core concept of TDA (Tauzin et al., 2021) is persistent homology, which identifies lasting geometric and topological features by using a method called simplicial complexes to extract shape-based information from point cloud data (Fig. 5). A simplicial complex is an assemblage of simplices, foundational elements of the graph's higher-dimensional analogues (Carlsson, 2009). An n-dimensional simplex arises from the connection of $n+1$ affinely independent points. For instance, a simple point represents a 0-dimensional simplex, a line segment joining two points constitutes a 1-dimensional simplex, and a filled triangle formed by three non-linear points exemplifies a 2-dimensional simplex. The Vietoris–Rips complex characterizes the data's shape as a simplicial complex using a parameter ϵ (Munch, 2017). Its fundamental principle is to connect any pair of points in the point cloud dataset if their distance is less than ϵ, with ϵ increasing incrementally. Thus, to thoroughly comprehend all structures within the data, one must consider all values of ϵ greater than 0. In the end, crucial structures such as holes and connected components emerge and die with changes in ϵ within a landslide polygon. These structures relay

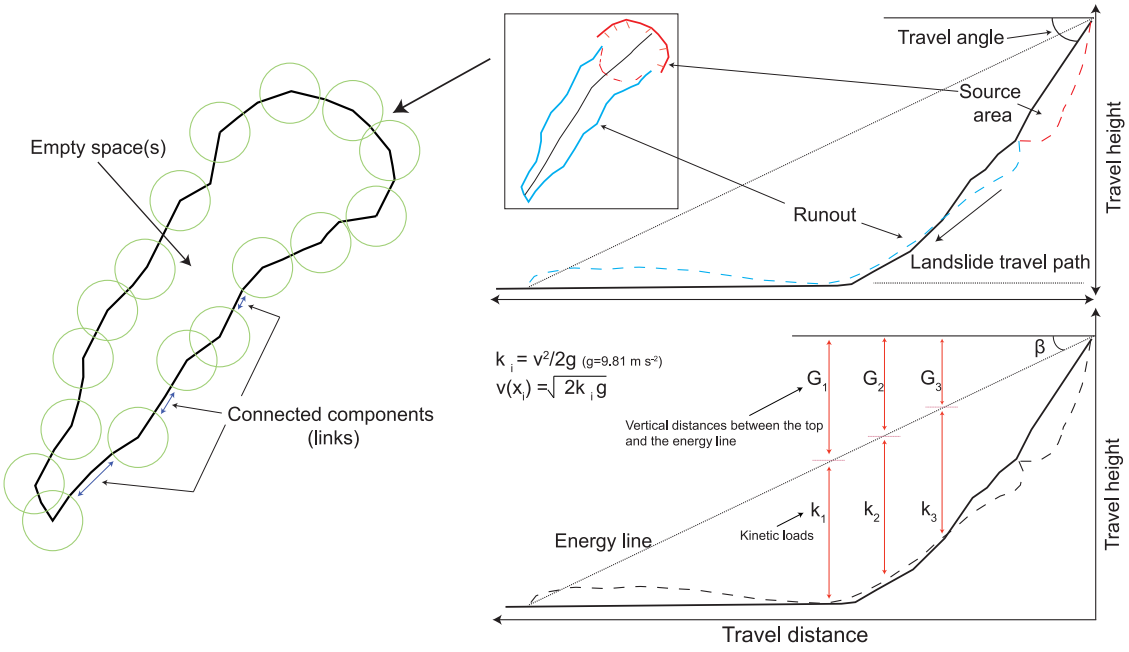


Fig. 5. The schematic illustrates the varying topological and morphological properties that are employed in the machine learning regime. Topologically, the empty spaces, termed holes, and connected components define the basis for topological analysis. The green circles denote the disks that are generated over the vertices of the polygons, which increase with an increase in ϵ until structures such as holes and connected components collapse. For the morphology, properties such as travel distance, travel angle (β), travel height are calculated. Furthermore, using said properties, potential velocities are calculated for multiple locations within a landslide body. k_i refers to kinetic loads acting at different points along the landslide path under the influence of gravity, g (=9.81 m/s^2). G are the vertical distances between the top of the path and the energy line. $v(x_i)$ represents the velocity of the landslide body at different points as a function of the travel distance. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

morphological information such as the size of the landslide indicating (i) the degree of compactness: related to the scarp definition, (ii) the sinuosity of the transport zone describing kinematic propagation related to the runout, and (iii) slope variations distinguishing sharp or gentle changes in the hillslope related to the elevation change and distance of runout (Bhuyan et al., 2024). We calculated various measures quantifying the topological properties of the landslides based on these structures—average lifetime, number of points, Betti curve-based measure, Wasserstein amplitude, and Bottleneck amplitude. These properties are then fed into a machine-learning algorithm to separate source and runout areas of landslides.

In addition to leveraging high-dimensional TDA to decipher insights from the shapes of landslides, we extend to encompassing properties that describe the physical attributes of the landslide process. These attributes contain various factors, such as the travel angle (or angle of descent/propagation), travel distance, travel height, and an estimation of the energy expressed as velocity (see these properties in Fig. 5), which serves as a kinetic indicator. These properties are derived from the geometry of the landslide, and for the estimation of landslide velocity, our methodology adopts a simplistic gravitational model.

The energy model is based on the Heim energy line (Heim, 1932), which we used to estimate potential velocities within each landslide (a few samples are shown in Fig. 6). It is important to note that these velocity estimates are not corroborated with actual velocity data for the landslides in our inventories. They are theoretical constructs designed to extract the dynamic differences within a landslide's occurrence (or body), particularly when distinguishing between source and runout zones. However, the velocity model has been validated and used effectively in previous works, such as Lucas et al. (2014) and Marinelli et al. (2022).

We calculated the average of the velocity vectors within the landslides' polygonal boundaries. Identifying kinetic characteristics may shed light on dynamic behaviours in each landslide. This distinction is based on the different velocity vectors in these zones, offering a detailed insight into landslide behaviour. Our approach, therefore, combines the concepts of topological analysis with landslide morphology to offer a comprehensive view of landslide characteristics, improving our understanding of their general kinematic propagation stages.

3.2. Topo-morphological-aided machine learning regime

We systematically extracted a combination of topological and morphological properties, which were then integrated into the Random Forest algorithm. We then explored these amalgamated properties, enabling us to quantify and evaluate the landslides' shape. The primary objective was to ascertain the extent to which these properties/features contribute to the discrimination of the source region from the main body of the landslide. We eliminated the features that have high collinearity with others and those that have less feature importance via a feature importance analysis. We identified and prioritized the top ten features that significantly enhance the model's predictive accuracy and efficacy.

An essential element in the application of machine learning techniques is the preparation of labelled data, which is crucial for the training of supervised algorithms like Random Forest (Breiman, 2001). We aimed to identify the source region via the propagating length of a bounding box aligned with the scarp. Utilizing the ratio between the propagation length of the bounding box encasing the scarp region (in red, Fig. 7) and that of the bounding box encompassing the entire landslide (in blue, Fig. 7), we fed this information through a regression-based model (Eq. (1)). This model regresses on the cost function in Eq. (2). This approach allows for a more nuanced understanding of landslide scarp delineation.

The model can be conceptualized in the following manner:

$$\text{Source-to-total length ratio} = \text{RandomForestRegressor}(X;\Theta) \quad (1)$$

Here, X represents the array of input features, which includes topological and morphological data, while Θ signifies the set of parameters within the Random Forest model. The primary goal is to minimize the error between the predicted ratios and the observed ratios of

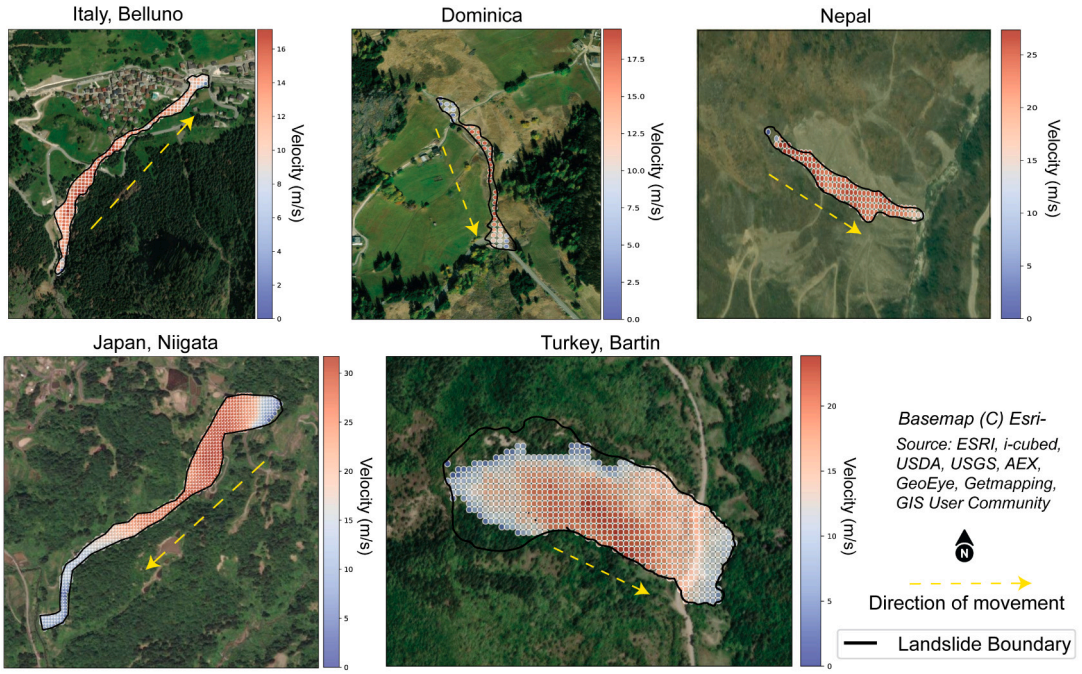


Fig. 6. An example of the estimated velocity based on the Heim energy line (Heim, 1932). Different locations within the landslide body show varying velocity vectors.

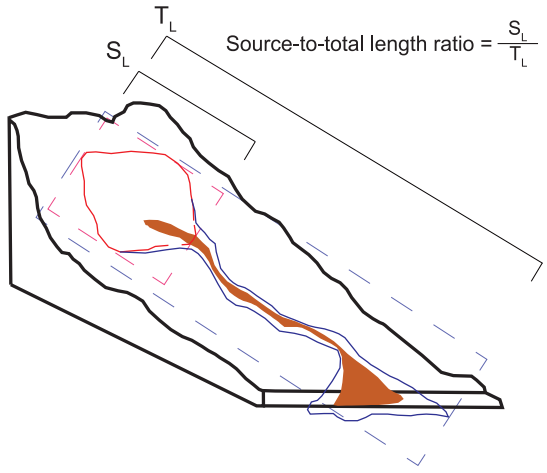


Fig. 7. An illustration of the ratio between the source length (S_L) and the total landslide length (T_L) that is used in the machine learning algorithm.

source-to-total length in the training data, expressed as a cost function J(Θ):

$$J(\Theta) = \frac{1}{N}\sum_{i=1}^{N}(\text{Predicted Ratio}_i - \text{Observed Ratio}_i)^2 \quad (2)$$

where,

Predicted Ratio$_i$ is the ratio of source-to-total length predicted by the model for the$_i$th data, and

Observed Ratio$_i$ is the actual ratio of source-to-total length observed in the training data.

Once the model is trained and the scarp-to-landslide length ratio (1) is predicted, we convert this back to the predicted propagation length of the scarp region by multiplying the predicted ratio with the total length of the entire landslide. This approach allows for precisely predicting the source region's extent within the overall landslide polygon, ensuring that the model's output is directly applicable.

3.3. Evaluation scheme

We suggest a new metric to evaluate the performance of the predictions that rely on the observed and predicted propagation lengths of the scarp bounding boxes with respect to the total length of the bounding box encompassing the entire landslide (Eq. (3)). A larger landslide is likely to result in a larger error, and similarly, a smaller landslide could lead to a smaller error. To make the results comparable, we offer the below standardization equation to evaluate the model's performance.

Deviation metric

$$= \frac{|\text{Observed scarp length}_{O_L} - \text{Predicted scarp length}_{P_L}|}{\text{Total length}_{T_L}} * 100\% \quad (3)$$

where,

O_L is the observed length of the bounding box encompassing the source area;

P_L is the predicted length of the scarp;

T_L is the total length of the bounding box encompassing the whole landslide.

The integration of morphological and topological analysis-based machine learning approaches yields two primary contributions. First, the methodology extends beyond shape analysis (offered by morphological analysis) to encompass the physical attributes of landslides, such as travel angle, travel distance, travel height, and estimated velocity. This approach allows for a transition from static, deterministic values to dynamic, spectral metrics (e.g. dominant frequency components) that more accurately reflect the complex physical mechanisms governing landslide behaviour. Second, we incorporate morphological analysis with topological analysis that enables capturing nuanced variations within landslide shapes and properties across different terrains, ensuring generalized accuracy in landslide kinematic zone characterization across the varying inventories, meaning a transferable model.

4. Results

4.1. Topo-morphological proxies to separate kinematic zones

The importance of topology is explained by the shape (a landslide polygon in our case) properties, which help approximate the kinematic

propagation of a landslide. Coupled with morphological properties such as the travel distance, we are able to calculate the landslide's different kinematic zones/stages, i.e., erosion/depletion at the source area and transportation with deposition as runout. Information regarding the purpose of both topological and morphological properties for source area separation can be found in Table 2. We extracted this information from a combination of digital elevation models (DEM). For regions with access to high-resolution elevation models i.e., in Dominica and Italy, Belluno (Tarquini et al., 2007), 5-metre and 10-metre resolution DEMs, respectively were used (obtained via LiDAR and aerial surveys); while for the rest—Nepal, Japan (Niigata), and Turkey (Bartin), 30 m resolution Shuttle Radar Topography Mission (SRTM) DEM (Farr and Kobrick, 2000) were utilized. The landslide polygon provides an approximation of the landslide boundaries, as they are often derived by standard surveying methods with suitable accuracies, such as 1–2 m. In the experiments performed below, we combine all the data from Dominica, Nepal, Japan (Niigata), Italy (Belluno), and Turkey (Bartin) to understand the importance of these topo-morphological properties (Fig. 8).

A critical property that helps gauge landslide kinematics is compactness and sinuosity, which capture the propagation as a landslide progresses downslope. This information is effectively encapsulated by the topological properties—Average Lifetime of Holes (AL_H) and Bottleneck Amplitude of Holes (BA_H), where the former represents a hole's average size and estimates information related to an empty space, thereby the compactness within a landslide body, and the latter represents the maximum lifetime of holes within a landslide body, thereby quantifying the maximum empty space occupied in the landslide.

An analysis based on the two properties through partial dependence plots (Figs. 8-B–K) demonstrates a clear correlation between the topological properties of AL_H and BA_H, and the ratio of scarp length to total landslide length. Higher AL_H values, reflecting the persistence and size of holes, suggest a more substantial and enduring scarp (therefore, more compactness), possibly due to extensive deformation. Likewise, increased BA_H values indicate a wider variation in void size over the landslide's 'lifespan', pointing to a more irregular and extensive scarp geometry.

In terms of BA_H, greater amplitude might reflect a scarp that has undergone extensive erosion or weathering, leading to an uneven distribution of holes and, hence, a larger scarp. This relationship underscores the sinuosity as a key feature in understanding the topological properties within a landslide's structure. The sinuosity reveals how the internal topology and the geomorphological context coalesce to shape the landslide's morphology and influence the source-to-total length ratio.

Another important property is that of slope variations within landslide bodies. Although travel height does capture this information in a morphological sense (measuring the vertical extent), topologically, properties like Wasserstein Amplitude of Connected Components (WA_C) and Average Lifetime of Connected Components (AL_C) measure the slope's internal variations. The WA_C measures the largest slope changes within the landslide outline by quantifying the longest lifetimes of connected components. An increase in WA_C signifies that the landslide possesses pronounced topographic transitions or critical slope alterations. More substantial slope variations elevate WA_C values, which often correspond to a more pronounced scarp area in relation to the total landslide mass. For example, the landslide scarp has developed in areas of marked topographic relief, potentially due to underlying geological heterogeneities or pronounced erosional processes that have shaped the slope over time.

The trend observed with AL_C—a decrease in AL_C values associated with an increase in the source-to-total length ratio (Fig. 8-A)—suggests that in areas where the scarp (the initial break or detachment zone of a landslide) is larger, the features within the landslide, such as empty spaces, do not persist for long periods. A smaller AL_C value means that the landslide's internal structure (through a topological lens), especially near the scarp, is changing or collapsing more quickly. Hence, larger scarps may mean that the landslide material is less stable or more heterogeneous. In such conditions, the internal features of the landslide are more transient and less consistent over time, leading to a quicker alteration or breakdown of these features. This change contributes to a larger scarp, as the instability in these areas results in more extensive initial failures or detachments.

Finally, the Betti Curve of Holes (BC_H) and the Betti Curve of Connected Components (BC_C) are topological metrics that analyse the structural complexity of landslides. BC_H measures the number of holes, indicating the degree of fragmentation within the landslide. BC_C quantifies the number of connected links within a landslide, reflecting how the mass is segmented or unified. In the partial dependence plots (Fig. 8-B), a simultaneous decrease in both BC_H and BC_C, along with a reduction in the source-to-total length ratio, is observed. This decrease suggests landslides with fewer internal discontinuities that are more cohesive. Correspondingly, these landslides exhibit relatively bigger scarps compared to their total size, indicating a cohesive movement pattern, while in contrast, scenarios with higher values of BC_H and BC_C are indicative of landslides featuring smaller scarps and a greater degree of internal fragmentation.

Morphological properties of landslides derived from DEMs can also provide important contextual information about the kinematic propagation stages of landslides. As seen in Fig. 8, travel distance is deemed of the highest importance in predicting the source area ratio. Travel distance essentially reflects the extent to which the landslide material moves from its point of origin. Accordingly, we observe a 'negative' influence of travel distance on the source area length (Fig. 8).

Shorter travel distances in landslides typically suggest a rapid loss of momentum or a quick deposition of materials. Common in sliding mechanisms, the failure plane allows for a sudden release and descent of material but with limited horizontal displacement. In these scenarios, the material tends to come from a larger, more cohesive source area, as the landslide involves a substantial mass moving almost as a block.

The association of longer travel distances with smaller sources of failure can be linked to the dynamics of flow-type landslides, where materials enter channels. Flow-type landslides such as debris flows are characterized by their fluid-like movement and can originate from relatively smaller source areas yet travel significant distances (due to channelization) compared to other landslide types, for instance, in mountainous regions with narrow valleys. In such steep landscapes, a small initial failure in a steep headwater area can transform into a large and fast-moving debris flow.

Both travel height and angles have a 'positive' role in predicting the ratios, meaning higher values of the two properties predicted larger source areas. Higher values of travel height and angles indicate larger source areas in landslide events. A greater travel height typically suggests significant mass mobilization with high energy and momentum (Cruden and Varnes, 1996; Hungr et al., 2014), often correlating with larger source areas.

The angle of descent in landslides describes the steepness of the slope over which the landslide travels. Steeper angles often lead to faster and more extensive landslides, thereby indicating a more comprehensive area of origin. The steepness affects the kinetic energy and velocity of the moving mass, contributing to a broader dispersion of materials and, thus, a larger source area.

Initial calculated velocities of approximately 1–7 m/s depict larger source ratios (i.e., source-to-total length ratio); however, higher velocities (over 8 m/s) depict medium-sized source zones (under 0.38). The range of velocities between 1 and 7 m/s suggests an association with landslides with a substantial mass and cohesive nature. These relatively slower landslides are typically sliding-type failures with greater friction and resistance encountered by their larger volume. With higher velocities, it can be argued that landslides originating from smaller, more localized areas can often gain significant speed due to steeper slopes (which was also observed in the travel angle plot where a ratio of 0.38 occurred around 40 degrees of slope) and less resistance, allowing material to travel further and faster.

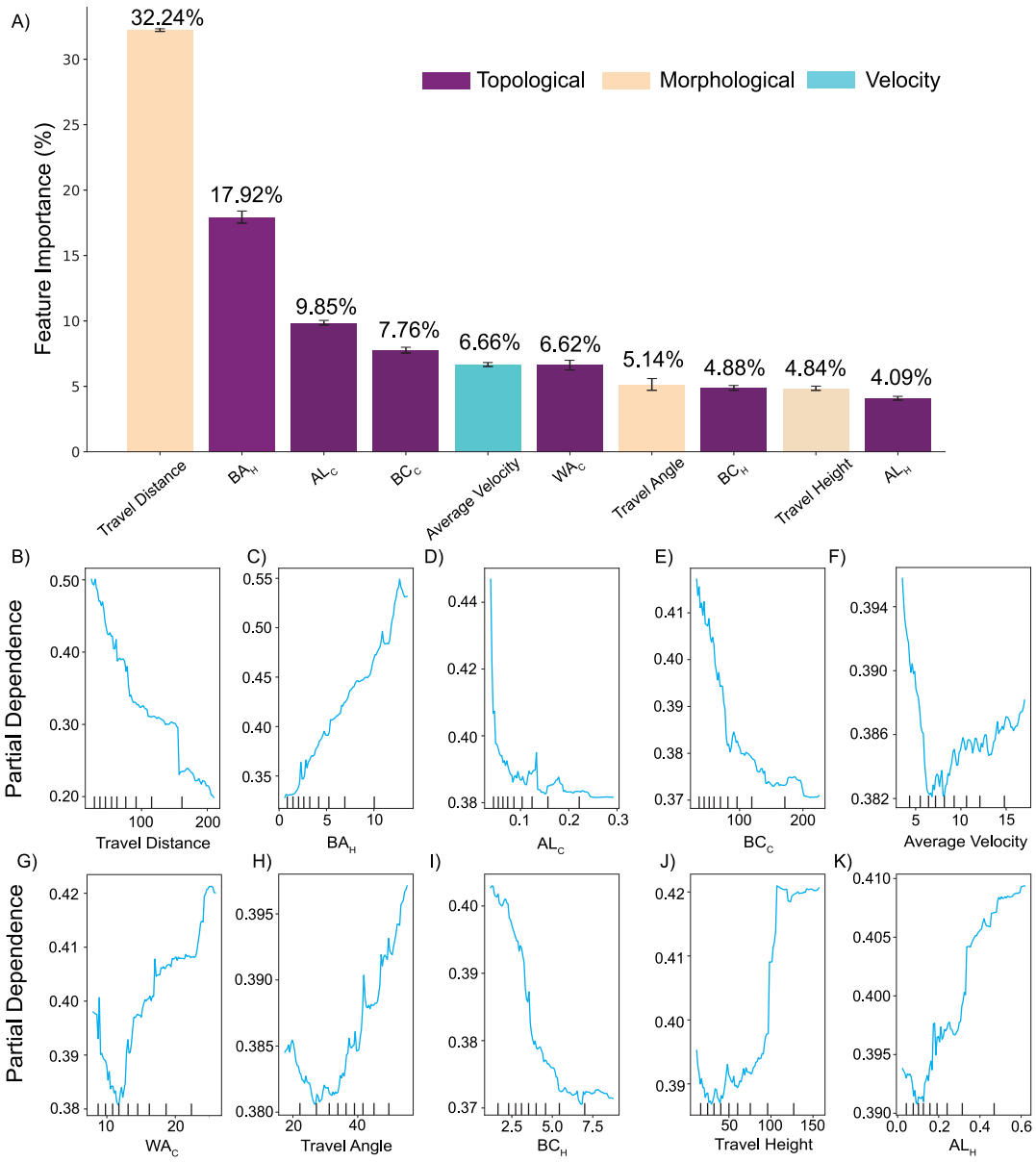


Fig. 8. Sub-plot (A) shows the probability density distribution of the top 10 most optimal features. We calculate these importance metrics (out of a total of 34 features) based on the combined training of the model using the landslide inventories of Dominica, Nepal, Japan (Niigata), Italy (Belluno), and Turkey (Bartin). The standard deviation for each feature is reported (obtained after 100 runs of 10-fold model training). The topological properties (in decreasing order) are: Bottleneck Amplitude of Holes (BA_H), Average Lifetime of Connected Components (AL_C), Betti-Curve of Connected Components (BC_C), Wasserstein Amplitude of Connected Components (WA_C), Betti-Curve of Holes (BC_H), and Average Lifetime of Holes (AL_H); the morphological properties are: Travel Distance, Travel Angle, and Travel Height; and Average Velocity (velocity). (B) to (K) shows the partial dependence plots of each feature or property (x-axis) and their corresponding response to the prediction ratio (y-axis).

4.2. Proof-of-concept (Dominica, Nepal, Turkey)

We trained and evaluated the model for the landslides in the regions of both event-based and historical inventories: Dominica, Nepal, and Turkey (Table 1, approximately 30,000 landslide samples). The model was trained using 10 k-folds and yielded an average deviation metric (Eq. (3)) of under 15% with a standard deviation of ~8%. Also, the mean deviation metric varied with respect to different regions (Fig. 9); however, this can be explained by—the varying scarp distances of landslides specific to each region, different triggers, the type of landslide failure movement, topography, and underlying geological conditions. The predictions from the model against the observed ground truth can be seen in Fig. 10.

The prediction of scarp length against the observed scarp length lies close to the 45-degree line, along with low variations (e.g., Fig. 9). The probability distribution of the deviation metric shows that 93.9% of all data falls under a 25% deviation metric (Fig. 9-A). A total of 59 landslide polygons showed variations above 50%. The 95% confidence interval of the mean for this metric ranges from 7.3% to 7.7%, indicating a narrow deviation range.

Dominica exhibited a mean deviation metric of 10.37%, accompanied by a narrow 95% confidence interval for the mean, ranging from 9.86% to 10.88%. Nepal showed even more promising results, with a mean deviation metric of 7.48%; showcasing the best results with a 95% confidence interval of the mean extending from 7.05% to 7.91%, implying a slightly more accurate prediction capability. Turkey, with the highest mean deviation metric of 12.30%, also displays considerable prediction accuracy. The confidence interval of the mean for Turkey ranged from 11.09% to 13.52%, indicating the lowest accuracy among the three regions, yet still capturing a significant level of precision in the overall dimensions of the landslide regions.

Table 2
Summary of landslide topological and morphological properties, their purposes, and their method of calculation.

Topo-morphological properties	Purpose	Calculation method
Travel Distance (m)	Indicates the extent of landslide material movement	Measured as the horizontal distance from the landslide's origin to the furthest point reached.
Travel Height (m)	Measures the vertical extent of a landslide, with greater heights suggesting significant mass mobilization	Calculated as the vertical difference between the highest and lowest points of the landslide.
Travel Angle (°)	Describes the landslide angle of reach, affecting kinetic energy and the extent of the landslide	Calculated using the arctangent of the ratio of travel height to travel distance.
Estimated Velocities (m/s)	Reflects landslide dynamics by indicating the velocity range, which correlates with the size of the source area and the mobility of the landslide mass	Derived from kinematic equations, considering the slope angle and kinetic energy.
Average Lifetime of Holes (AL_H)	Measures the average compactness of the landslide body	Computes the mean lifetime of holes within a landslide polygon, a structure observed when implementing TDA.
Bottleneck Amplitude of Holes (BA_H)	Quantifies the maximum compactness present within the landslide	Computes the maximum lifetime of holes within a landslide polygon using TDA.
Wasserstein Amplitude of Connected Components (WA_C)	Assesses the largest slope changes, indicating critical slope alterations within a landslide	Analysed through TDA, quantifies the connectivity of connected components – structures formed from the vertices of landslide polygons – by applying the Wasserstein amplitude.
Average Lifetime of Connected Components (AL_C)	Tracks the average slope variations, with bigger values suggesting quicker changes in slope	Calculated by evaluating the persistence/lifetime of connected components indicating how long features like slope variations last as ϵ change.
Betti Curve of Holes (BC_H)	Indicates the overall compactness and sinuosity of the landslide	Calculated by counting the number of holes and their sizes using TDA.
Betti Curve of Connected Components (BC_C)	Measures internal fragmentation of landslides that highlight combinations of large, medium, and no slope changes	Derived from TDA, this curve records the number of connected components or clusters and their sizes within the landslide.

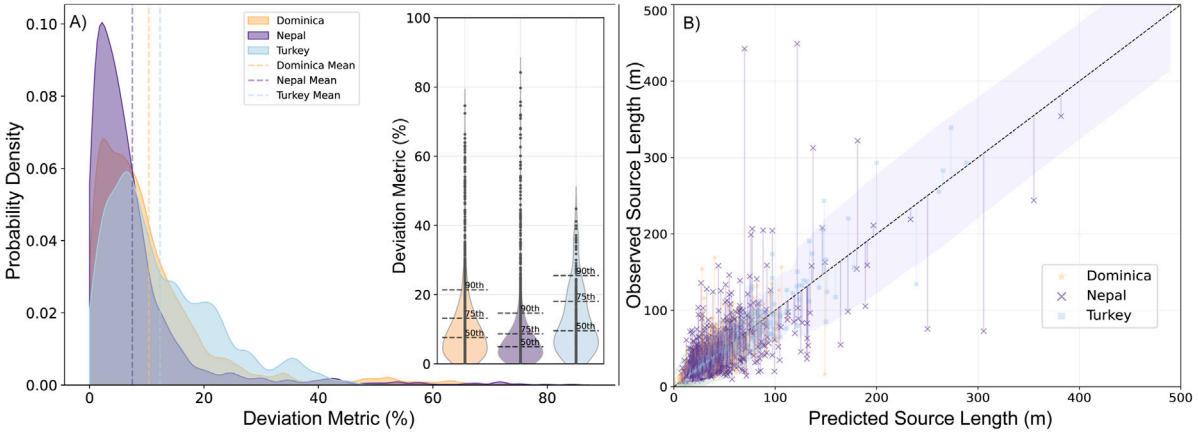


Fig. 9. Deviation metric of predictions from the observations for the proof-of-concept regions. (A) the probability density distribution of the proof-of-concept countries, where the x-axis reports the deviation metric (in %) and the y-axis reports the probability density. (A) documents the mean deviation and the 50th, 75th, and 90th percentiles in violin plots. (B) the scatter plot of the observed and predicted source length dimensions (in metres) while fitting a kernel density in green with an error margin (in shaded blue) across all the test samples in the proof-of-concept regions.

Some results, such as those from Nepal, exhibit both under- and over-predictions, though these cases are minimal. Among the 3852 test landslides, less than 1% (0.9%) showed significant under-predictions (32 cases) or over-predictions (5 cases). We attribute this low uncertainty to the under-representation of deep-seated landslides in the dataset, making it difficult for the model to accurately predict these events. Additionally, landslides with source lengths greater than 150–200 metres were underrepresented in the training data (about 0.64%), further contributing to the model's challenges. As a result, the model, which was primarily trained on shallow events, tends to treat deep-seated landslides as having similar shallow properties, leading to under-predictions of the source-to-total length ratio and source length. This facet is further discussed in Section 6.3.

Overall, these results illustrated the robustness of our methodology in handling variations among samples (of landslides) across varying geomorphological and climatic settings.

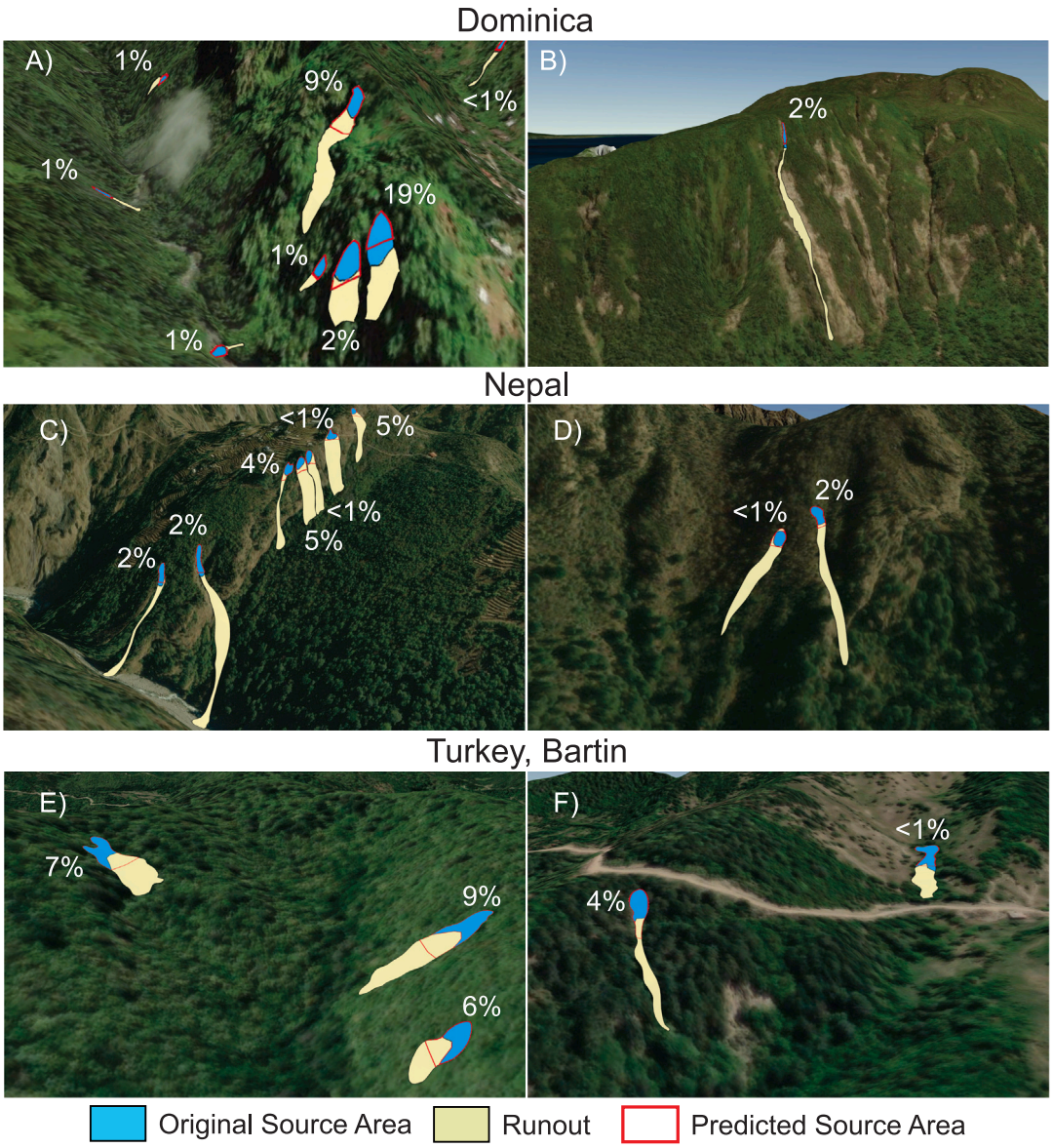


Fig. 10. Model predictions for the proof-of-concept countries, where (A), (C), and (E) show clusters of different landslides, and (B), (D), and (F) show specific landslide examples, such as slides and flows. Each landslide is accompanied by the respective deviation metric (Eq. (3)), highlighting the prediction accuracy. Base map source: World Imagery, credits: Esri, Maxar, Earthstar Geographics, and the GIS User Community (Esri, 2023).

4.3. Separation based on failure movements (Dominica and Turkey)

Inventories from Dominica and Turkey also contained landslide failure movement types. The Dominican inventory predominantly consisted of debris slides and debris flows, whereas the Turkish inventory encompassed a variety of landslide types, including slides (mostly rotational), earth flows, and slides evolving to flows (termed 'Slide Earth Flows'). We pragmatically explored how the distinct failure types could influence model performance. This exploration aimed to deepen the understanding of the interplay between failure types and the predictive accuracy of analytical methods.

To explain the combinations, any reference to the single failure type means that the model is trained and tested on the same type with 60% to 40% splits via a 10-fold model. For instance, flow in Fig. 11 for Dominica means that we used 60% of the flow type failures in training the model, and we tested the model on the remaining 40%.

Some of the landslides are complex, including multiple types, such as slides transforming to flows. Hence, we suggested another, more challenging testing scheme. For example, we trained the model instance with slides and tested it with flow-type failures. These are referenced with two types of failures, one of which is slide-flow. Finally, when a failure type is followed by 'all', it indicates that the training was conducted using sub-samples of the specified failure type and then tested on all remaining samples of the same and other failure types.

We found that cross-failure-type tests performed relatively well (e.g., training on slides and testing on flows in Dominica) with deviation metrics generally under 20%. Training and testing on similar types of failures resulted in an overall good performance with low deviation metrics < 11%. Similar failure types would bear comparable topological and morphological properties (Bhuyan et al., 2024), thereby consisting of a similar range of ratios that are predicted by the model.

Another testing scheme included training on one set of failure types and testing on all remaining data (including all types). This restriction

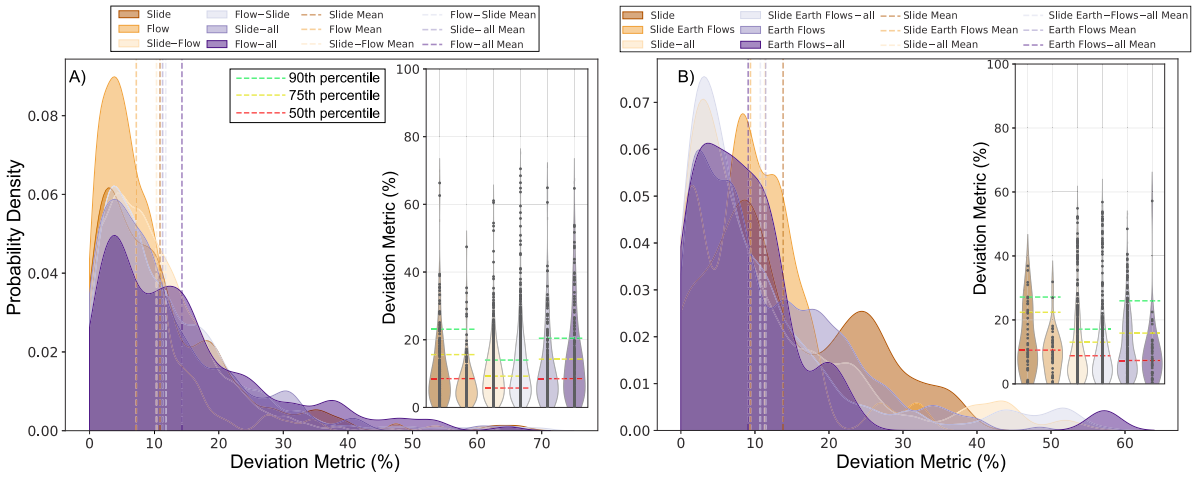


Fig. 11. Probability density distribution of the various failure movement types and their combination in (A) Dominica and (B) Turkey, where the x-axis reports the deviation metric (in%) and the y-axis reports the probability density. (A) and (B) show the mean deviation and the 50th, 75th, and 90th percentiles in violin plots. A single type of landslide failure movement refers to those training/testing combinations exclusively conducted in the respective category. For example, if a colour is referenced as slide-flow, this means that the model is trained on slides and tested on flows.

also resulted in < 10% deviation metric. The model does struggle at times to predict across different failure types but still within acceptable margins of under 15% deviation metric.

4.4. Model adaptability (Italy, Belluno and Japan, Niigata)

The Vaia storm in 2018 in Italy's Belluno province and the 2004 earthquake in Japan's Niigata prefecture were two notable examples (with ground truth) with single dominant triggering mechanisms for testing the method's application beyond the proof-of-concept countries. The model, developed using data from the proof-of-concept countries, was directly tested on these regions in Italy and Japan (Fig. 13). We observe in Fig. 12-B that the error margin is limited around the 45-degree line, with the predictions closely resembling the actual source area lengths. Compared to the previous scheme, where testing and training are performed in the proof-of-concept data, the mean deviation metrics are higher, on average 25% (Fig. 12-A). The reduced prediction accuracy is expected as the trained model (trained on the proof-of-concept data) contains no landslides from Italy and Japan. Hence, we observe higher deviation metrics for certain individual landslides as seen in Fig. 13-C.

In the case of Italy, it is noteworthy that, unlike Japan, the samples to test the model were quite limited (only 42, which were manually crafted based on aerial photography). Referring to Fig. 12-B, most predictions are under 100 m (both x- and y-axes). The high mean deviation metric of 18% is due to a few samples being predicted with a larger margin of error, caused by the diverse nature of landslides not represented in the original training data. This results in the overall mean skewing heavily on the right side of the scatter plot.

5. Real-world application

After validating the model performance and adaptability across diverse sites, we use it to delineate kinematic zones on landslide inventories that are available in the literature but do not contain such information. We choose open-source inventories and treat them as real-world applications of the proposed method for separating the source areas from the total landslide bodies. Details regarding these inventories can be found in Table 1. Of course, some predictions will harbour uncertainties introduced by (1) how the inventories were generated in the first place, (2) what data was used to map and delineate the landslides, (3) prior expertise while mapping the landslides, and (4) the quality of the elevation model. We release the model's output openly for users. Overall, we considered two experiments: First, deploying the model on inventories that are widely distributed geographically and therefore highly diverse in terms of typology, characteristics, and mapping methods; Second, deploying on a multi-temporal inventory.

5.1. Spatial deployment

We considered landslide inventories of the 2018 Hokkaido Eastern Iburi co-seismic (M_W 6.6) inventory (Wang et al., 2019; Dou et al., 2020; Yunus et al., 2023), the 2018 Papua New Guinea, Porgera co-seismic (M_W 7.5) inventory (Tanyaş et al., 2021), the 2019 Mesetas, Colombia co-seismic (M_W 6.0) inventory (Delgado et al., 2021), and the Chile historical debris flow inventory (Goetz et al., 2021). Fig. 14 illustrates the model's predictions applied to these inventories. We also report the morphological statistics (see Fig. A.2) that are calculated on these respective inventories. Predictions for Papua New Guinea and Colombia were made using the SRTM elevation model, for Chile using a 12-metre ALOS-PALSAR elevation model, and for Hokkaido using a 1-metre LiDAR-based elevation model.

We calculated the mean and standard deviation of the area for the source (μ_S, σ_S) and the entire landslide body (μ_T, σ_T), respectively. We present these statistics for Colombia based on the failure movements already identified in the inventory by Delgado et al. (2021), Fig. A.2.

Regarding the 2018 Hokkaido earthquake event, we validate our results using the DEM difference (post- and pre-event LiDAR; 1-metre resolution) obtained from Yunus et al. (2023), see Fig. 15-A. Japan, Hokkaido has a mean scarp area of μ_S=1643.1 m^2, which is 30% of the mean total landslide area (μ_T=5546.5 m^2).

In the case of Papua New Guinea, Porgera, the mean scarp area accommodated μ_S=2614.2 m^2, which constitutes 13% of the mean total landslide area (μ_T=19440.2 m^2). The relatively low percentage coverage may be attributed to underestimating the scarp regions in some rock/debris avalanches (Tanyaş et al., 2021). The definitions of scarp areas for such failure movements are often more complex.

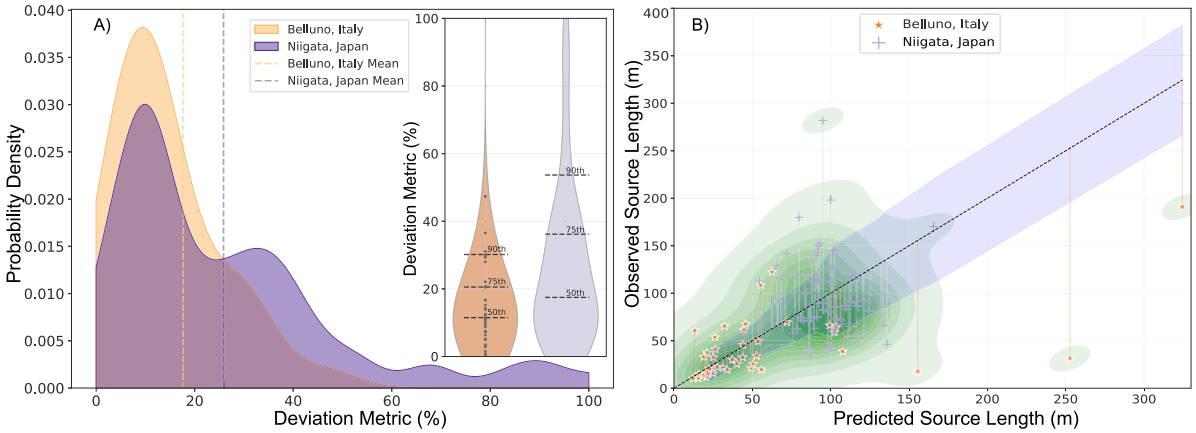


Fig. 12. Deviation metric of predictions from the observations for the application regions. Probability density distribution of Italy (light orange) and Japan (purple) in (A). (A) inlet violin plots highlights the mean deviation and the 50th, 75th, and 90th percentiles in violin plots. The two regions are tested for transferability as an application with 42 and 1200 samples, respectively. (B) compares observed and predicted source lengths (in metres) fitted to a kernel density shown in green with an error margin (in shaded blue). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

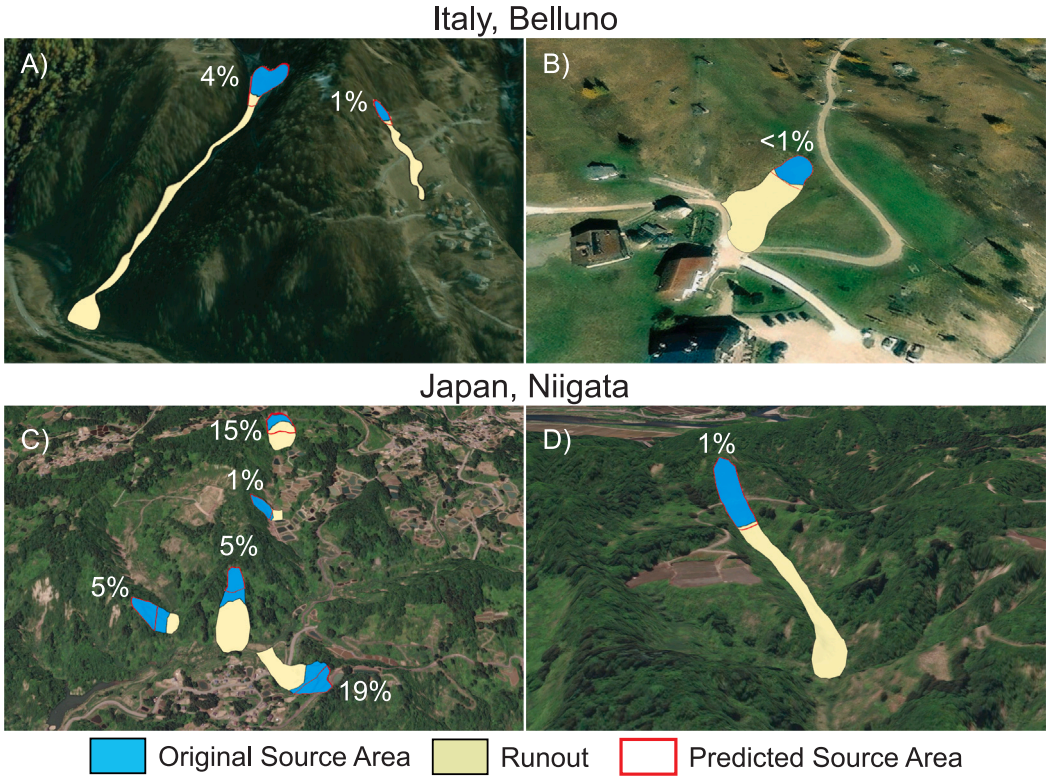


Fig. 13. Predictions for the regions where transferability is tested. (A and C) shows landslide clusters from each region, whereas (B and D) are single landslide examples. Each landslide is accompanied by the respective deviation metric (Eq. (3)), highlighting the model's accuracy. Base map source: World Imagery, credits: Esri, Maxar, Earthstar Geographics, and the GIS User Community (Esri, 2023).

In Colombia, the Mesetas region exhibited a mean scarp area of μ_S=750.0 m^2 for flow movements and μ_S=483.5 m^2 for sliding movements. These figures represent 20% (μ_T=3917.5 m^2) and 24% (μ_T=1990.6 m^2) of the mean total landslide areas for flow and sliding failure movements, respectively. The observation that sliding movements cover smaller areas than flows suggests that flow movements typically have longer runouts. This contrast is likely due to the higher mobility of flow movements, which enhances their travel distance beyond that of sliding movements.

For the historical debris flow inventory of Chile, the mean scarp μ_S=10780.8 m^2, accounting for 14% of the mean total debris flow area (μ_T=73927 m^2). Compared to other regions, the area percentage of the source is way smaller in Chile.

5.2. Temporal deployment

We also used our model to classify the multi-temporal landslide inventory of the 2008 Wenchuan co-seismic and post-seismic landslides (Fan et al., 2019). We trained the model on the proof-of-concept regions and then applied the model to the Wenchuan subsets for 2008, 2011, 2013, 2015, 2017, and 2018 (Fig. A.3). On the offset of using the entire landslide body for estimating their sizes, we find the area of the source/scarp zone to be much smaller than the runout.

We also calculated the mean and standard deviation of the area for the source (μ_S, σ_S) and the entire landslide body (μ_T, σ_T), respectively, in the Wenchuan case (Fig. A.4). The source zone of the 2008 co-seismic landslides (μ_S) accommodates 14% of the μ_T of the co-seismic

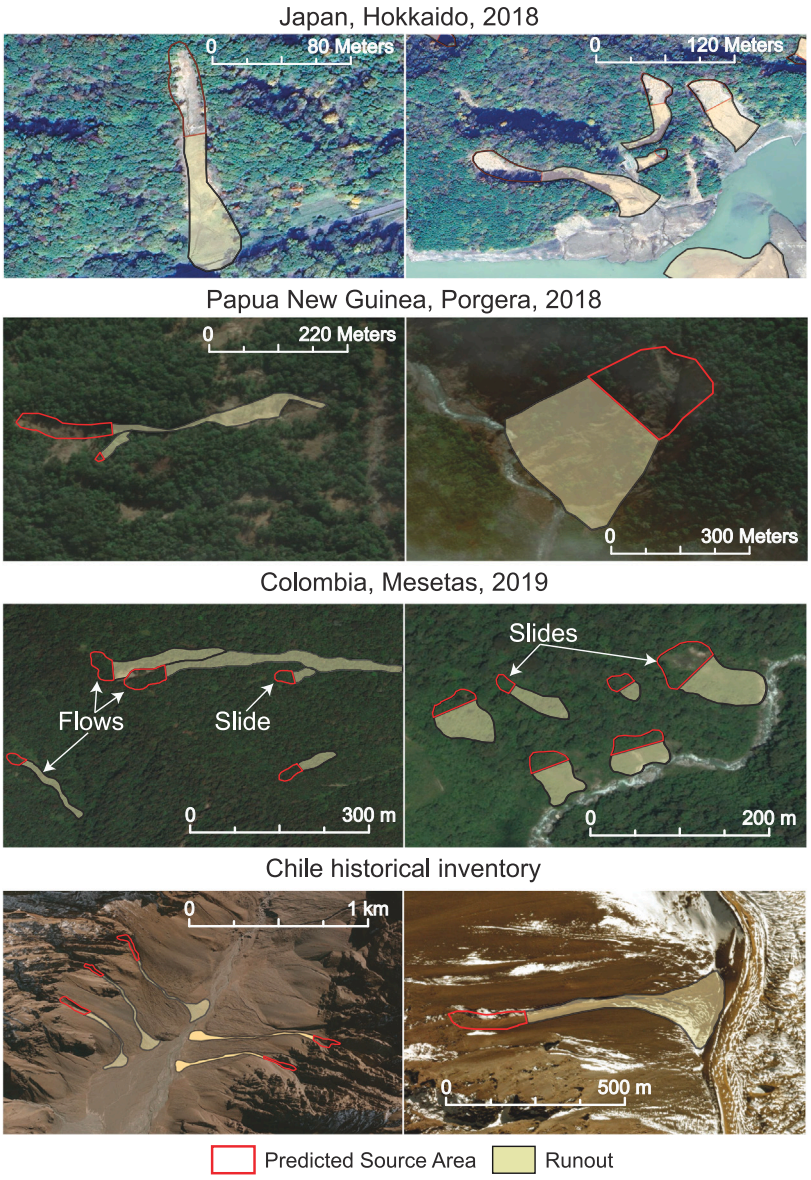


Fig. 14. Model predictions for the regions of Japan, Papua New Guinea, Colombia, and Chile. For Colombia, we also show the movement types reported in the original inventory, i.e., slides and flows. Base map source: World Imagery, credits: Esri, Maxar, Earthstar Geographics, and the GIS User Community (Esri, 2023).

landslides while the average μ_S for the inventories between 2011 and 2018 is accounted to be around 16.5% of the μ_T. The low source area percentage can be attributed to the region's significant topographic prominence. In the case of Wenchuan, this is likely due to the weak material properties and deposits of erodible debris from the 2008 coseismic landslides, combined with post-seismic rainfall (Huang and Li, 2014; Yu et al., 2014). Saturated and re-mobilized landslides (e.g., debris flows) tend to travel further, resulting in smaller scarp areas but longer runouts.

6. Discussion

This section will discuss the overall use of the newly introduced tool. Landslide kinematic separation could, above all, help improve area-based volume estimates of observed landslides. Another potential use is pairing the method with automatic landslide mapping to provide more nuanced data, i.e., landslide source and runout polygons separately instead of mere landslide polygons encompassing both kinematic stages. We also list some limitations of our tool, which may limit deployment without adjustments, such as lack of diverse samples and inherent limitations of DEMs. The current model is part of a series of tools that aim to improve landslide data for targeted use (Rana et al., 2022), especially for hazard assessment. Hence, we dedicate a small subsection to provide insights about potential future work.

6.1. Implications on area–volume scaling

Measuring landslide volumes directly on-site can be limiting because of restricted access, adverse site conditions, and time. Volume estimates based on empirical area–volume scaling have been introduced for global applications (Guzzetti et al., 2009), and further improved for regional use to decrease residual errors, for example, after major landslide triggering events (Yunus et al., 2023; Parker et al., 2011; Tebbens, 2020; Ju et al., 2023). Furthermore, extracting information about surface ruptures, particularly for landslide magnitudes, is of significant importance. This information can also provide valuable insights into orogenic growth and erosion equilibrium (Larsen and Montgomery, 2012).

We extracted the source areas of 566 coseismic landslides of the 2018 Hokkaido Eastern Iburi earthquake (M_W6.6) from Wang et al.

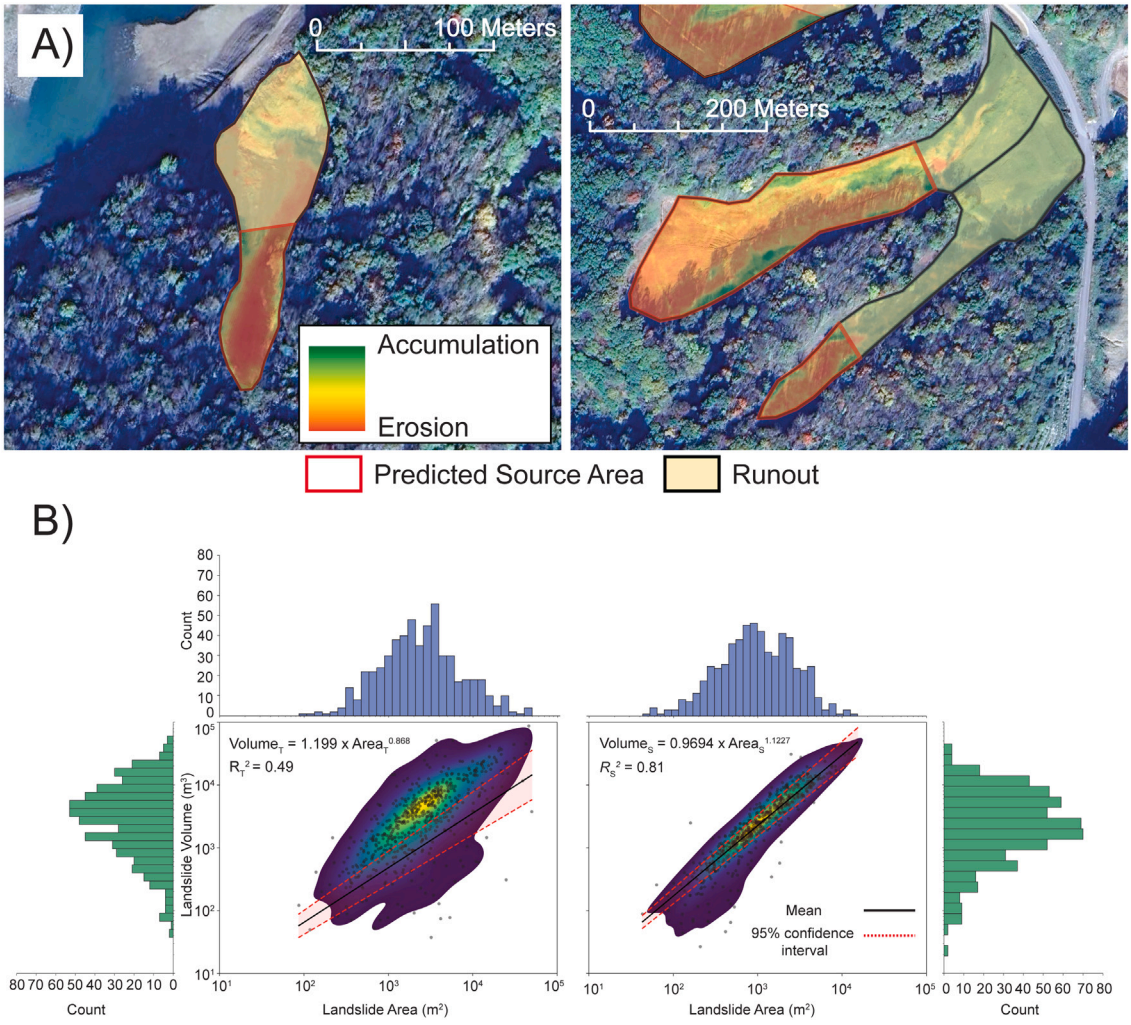


Fig. 15. A shows the predicted source areas on the Hokkaido 2018 landslides underlain with the subtracted post- and pre-DEMs, highlighting each landslide's erosion and accumulation parts. B depicts the area–volume scaling based on the total landslide area (Wang et al., 2019; Dou et al., 2020) and the source area that we mapped, respectively. Volume estimate using only scarp area achieves an $R^2 = 0.81$, which was $R^2 = 0.49$ based on total landslide polygons. The respective histograms of the area and volumes for both the total and source regions are presented as well.

(2019) and Dou et al. (2020). We derived an area–volume scaling relationship (plotted area and volume in log–log coordinates) of $V_S = 0.9694 \times \text{Area}_S^{1.12}$ ($R_S^2 = 0.81$). Our $\beta = 1.12$ is closer to an estimate extracted through post- and pre-earthquake DEMs ($\beta = 1.18$, $R^2 = 0.91$; Yunus et al., 2023) compared to the total area from the modified inventory using the total landslide planforms ($\beta = 0.86$, $R_T^2 = 0.49$), as shown in Fig. 15-B. We demonstrate that the source area provides a more accurate fitting (by 32%) for the landslides' area–volume scaling. Moreover, the size distribution of the landslides' source zones compared to their total planforms shows that the former is generally smaller than the latter (Fig. A.1).

6.2. Addition to automated-landslide-mapping pipeline

Landslide inventories with separated source areas are rare as separating landslide zones/stages is complicated and requires considerable time and resources (such as high-resolution imagery, 3D capability, and post- and pre-landslide DEMs). Our approach offers a straightforward and automated solution, depending solely on DEMs and landslide polygons to split landslide polygons. The overall deviation metric while mapping the source of landslides ranges between 7–25% in our various testing schemes (Figs. 9, 11, and 12).

In addition to mapping source areas across various geographical settings, the model demonstrates its capability with multi-temporal landslides initiated by different triggers (Fig. 10). The model is also robust to the specific landslide triggers, as it consistently performs well (7–12%) across various types of inventories, including rainfall (Fig. 10-A, B), seismic (Fig. 10-C, D), and historical inventories (Fig. 10-E, F). Furthermore, the model effectively identifies source areas regardless of variations in movement types, such as slides and flows, as demonstrated in tests conducted for Dominica and Turkey (Fig. 11).

If coupled with an automated landslide mapping tool, such as those from Bhuyan et al. (2023) and Tang et al. (2022), our model could enrich mapped landslide information and substantially support facilitating manual interventions. However, automatically mapped landslide inventories may contain all landslide movement types in one category. The underlying dynamics of each landslide failure type vary dramatically, which can decrease the accuracy and reliability of our models' predictions as we witnessed in the experiments of Dominica and Turkey in Fig. 11. In such an adaptation, we recommend an intermediate type classification (Bhuyan et al., 2024) before the kinematic classification. The development of pipelines that encompass all phases of mapping, movement type identification, and source/runout separation lays the groundwork for a system capable of producing reliable landslide mapping on regional and even continental scales. But, manual verification and validation by experts remain essential to ensure the accuracy and practical application of such automated outputs, particularly when cross-referenced with high-quality satellite imagery, contour, and slope

maps. This validation step is crucial for refining the data before its use in predictive modelling and hazard assessment.

6.3. Limitations and future work

The performance of the proposed model is directly tied to the quality of the landslide samples used for training, which depends on how accurately interpreters or mappers delineate landslides, including both source and runout areas, using methods such as geomorphological mapping, field surveys, and remote sensing image-based interpretation. If the delineation of source and runout boundaries in the training data is imprecise, the model will inevitably yield inaccuracies in the predicted landslide inventory. Furthermore, while the model demonstrates some degree of transferability, it still requires additional data to enhance its overall capacity for accurate predictions. TDA relies on nodes that depict the landslide polygons to capture properties like compactness and sinuosity. As such, the distribution of nodes while digitizing landslide boundaries becomes important. Differences in mapping scales and DEM resolutions could affect the level of detail in which landslides are seen and digitized. However, as long as key landslide features (e.g., scarps boundaries, long runouts, or debris fans) are well-represented, the model's performance remains consistent across different scales and regions. Our method relies on parameters of two different natures, i.e., topology and morphology. While topology explains the shape parameters of the landslide polygons, morphology expresses the planimetric representation of landslide properties such as travel distance, angle, height, and estimated velocity. If the landslide sample polygons are inaccurate, the parameters obtained from said samples will represent the landslides inaccurately, thereby affecting model training and prediction. Thus, it is essential to conduct quality checks on the inventories, and continuously update and expand the training dataset with diverse examples to ensure the model's robustness and accuracy.

This facet of diverse data is particularly evident in the case of landslides with larger scarps, for example, rotational failure types. Such landslides tend to suffer from inaccurate predictions. We attribute this to a bias in the model, as landslides with large scarps are under-represented in the training inventories. Consequently, the model may produce faulty predictions for landslides with large source areas. This under-representation also highlights the need for more diverse training data that includes a wider range of scarp lengths to improve the model's generalization capability and robustness. Another limitation arising from the diversity of data occurs in the case of landslides mapped with multiple source areas, for example, debris flows that enter a channel. Our model is unable to distinguish the individual source areas within a single landslide polygon that contains numerous branches. This limitation necessitates an additional step: isolating each branch into separate polygons before applying our model. Future efforts will attempt to resolve this issue via automatic "de-branching" of said landslide examples.

Currently, a perpendicular line (from the scarp) delineates the boundary between the source area and runout predictions, which does not realistically occur in nature. We would expect the separation boundary to conform to the topography or, at the very least, follow a semi-ellipsoidal shape (Cruden and Varnes, 1996). However, using a perpendicular separation line simplifies the modelling process and enhances the computational efficiency without significantly compromising the model's accuracy. Our method facilitates a clear delineation between source areas and runout zones, enabling more straightforward interpretations and applications in practical scenarios. Furthermore, while a more complex boundary may provide a marginally more accurate representation, the benefits in operational settings may not justify the additional computational cost and complexity. Future efforts could focus on refining the boundary definition using more natural, semi-ellipsoidal delineations while minimizing computational demands.

Our current prediction algorithm is designed to identify the highest elevation point from which it demarcates the boundaries of landslides. When the highest elevation cannot be accurately determined due to the coarser resolution of the underlying elevation model, the algorithm may incorrectly place the boundary. Even in this non-optimal scenario, the model calculates the correct prediction ratio (Eq. (1)). The DEM resolution plays a critical role in capturing smaller landslides. For instance, low-resolution datasets, like SRTM, mean excluding smaller landslides such as those under a pixel size, 900 m^2. For those small landslides greater than a pixel size, the model could potentially omit key features that could influence the overall accuracy. Hence, the DEM resolution could impact the model's predictive ability for landslides predominantly falling within the smaller size range (e.g., 2–4 pixels depending on the DEM resolution). At the same time, freely available DEMs ensure the reproducibility and reusability of any model. We utilized SRTM data at a 30-metre resolution for most study areas unless higher-resolution DEMs were available, i.e., ranging from 1 to 12.5 m resolution. Using DEMs of 5–30 m resolution is rather a choice of feasibility concerning processing efficiency for our large-scale assessments and predictive modelling, (at least > 18 sqkm) (e.g., in susceptibility, Catani et al., 2013; Ahmed et al., 2024). At these resolutions, elevation models could adequately capture the essential terrain features and patterns concerning most landsliding processes (Catani et al., 2013). Nevertheless, using higher-resolution DEMs should improve performance of models that rely on morphometrics up to 4-metre resolution (Smith et al., 2019).

Beyond resolution, elevation models' quality is essential to capture accurate morphometrics (Ouyang et al., 2023; Simard et al., 2024). Elevation models with precise positional accuracy, for example, LiDAR and Terrestrial Laser Scanners (TLS), outperforms SRTM at virtually any resolution (Jaboyedoff et al., 2012; Stumvoll et al., 2021; Tarolli, 2014) and should be used when available. Their use has consistently shown benefits in investigations of landslides, rock falls, debris-flows, and earth-flows, particularly in exploring detailed topographic features—including micro-topography, internal deformation structures, scarps, tension cracks, surface roughness, and displacement field computation (e.g., the Lamosano and Perarolo di Cadore landslides, Teza et al., 2008; the Radicofani landslide, Baldo et al., 2009; and large landslides in the Tatun volcano group, Chang et al., 2018). Building on the examples set by national topographic projects in Italy (e.g., TINITALY, Tarquini et al., 2007) and the United States (e.g., OpenTopography, Krishnan et al., 2011), we anticipate that more countries will follow suit, making quality and higher-resolution datasets increasingly accessible for local applications. Users of our model should consider using such DEMs – particularly those derived from LiDAR – when available. High-resolution elevation models may reveal information regarding micro-topographies related to subtle surface changes, such as scarp evolution, multiple source zones, fault displacements, or non-linear separation zones between stages of failure propagation. We recommend future work to explore these potentials to advance the model.

Regarding geology and its variability, our model's performance is closely tied to the characteristics of the landslides in the dataset, most of which are shallow. Shallow landslides, predominantly composed of superficial materials like soil, regolith, or colluvium, only partially reflect the underlying geological units. This has allowed the model to generalize well for shallow landslides across different regions, as their geotechnical properties tend to converge globally. However, we observed reduced accuracy when predicting large landslides (presumably, deep-seated landslides), as seen in Figs. 9-B and 12-B, which aligns with geological differences that introduce noise into the model. These discrepancies highlight the limitations of the current model in fully capturing geological variability, as TDA (topological data analysis) only implicitly accounts for geological factors via proxy data from landslides used in training. Introducing geology into the analysis could increase the model's performance but at the same time, would also add reliance on site-specific data.

Finally, the consideration of revealing source and runout zone morphometric variations as a function of the trigger mechanism is a topic which was not covered in this study but represents an important area for future research. Landslides triggered by rainfall and seismic events often exhibit differing properties, such as travel distances and heights, but our analysis (Fig. 3) did not reveal significant variations at first sight, similar to the findings of Rana et al. (2021). Although combining landslides from different triggers adds complexity, we did so to build a more holistic, transferable model capable of learning diverse characteristics across terrains. This choice allowed testing the model on landslides triggered by either mechanism (including landslides initiated by mixed triggers in the case of Turkey's historical inventory). However, key differences likely exist between rainfall-induced and co-seismic landslides. Future work should integrate previous findings (Rana et al. 2021; Bhuyan et al. 2024), which differentiate landslide triggers and movement types, with the kinematic zone separation model to better compare landslides based on their movement styles and kinematics to comment on the triggers (Baroň et al., 2024), focusing on location and runout differences.

6.4. Landslide hazard and runout modelling

Most contemporary approaches involve using the crown of the landslide as a point or entire landslide bodies to capture information pertaining to landsliding conditions and then train statistical models for susceptibility (Lombardo et al., 2020). Our model would allow training such susceptibility and hazard models in designated source areas of landslides. For example, susceptibility estimates relying on segments of the landslide polygon may increase prediction accuracy by 10% when compared to a model trained with total landslide polygons (Ozturk et al., 2021). Hazard tools underlying a kinematic separation step may lead to more nuanced hazard estimates, potentially leading to more targeted mitigation strategies.

In addition to the findings of the volume estimates (Fig. 15), precise volumes can be estimated that are contributed simply due to the scarp regions and not the entire polygons, thereby allowing us to revisit statistical scaling relationships and allow for better estimates at larger scales. Alas, this can also enable visiting more crucial questions that concern slope dynamics and the details regarding landscape alteration due to slope erosion.

Numerical models designed for mass flow (Mergili et al., 2017) and runout path (Wichmann, 2017; Ye et al., 2024) simulations require the scarp region to define the release area. Our method can help create said simulations and effectively validate the runout paths feasibly with empirical data, with the possibility of simulating at larger regional scales beyond slope-scale experimental models. Other related studies have also pointed out improved accuracies in physically modelling landslide runout and behaviour, led by a defined source zone as release areas (Sorbino et al., 2010; Blahut et al., 2010). Hence, the indication of release areas is crucial for enhancing the precision of these simulations and therefore, can lead to more targeted mitigation strategies and policies (e.g., Gatto et al., 2024).

7. Conclusions

Although some landslide inventories, including those used in this study, distinguish between the source and runout zones, this process of separation is time-expensive and hence, it is often absent in many databases, where landslide source and runout zones are commonly combined into a single entity. This practice can compromise predictive modelling accuracy, often resulting in overestimations of landslide occurrences. In this study, we successfully developed and validated a methodology that distinctly separates the source and scarp regions from complete landslide polygons, utilizing topological and morphometric data. Initially tested in regions with available ground truth, our approach demonstrated predictive accuracies within 15%–20% of the deviation metric. Extensive testing across diverse triggers and movement types confirmed its adaptability and effectiveness, with applications in Italy and Japan further establishing its robustness under a deviation metric of less than 25%. Notably, the application of this methodology to co-seismic landslides from the 2018 Hokkaido earthquake resulted in a 32% improvement in empirical area–volume estimates, underscoring significant advancements over traditional models that only analyse entire landslide polygons. Our method not only offers a more precise method for mapping landslide scarp dimensions but also sets a new standard for creating enhanced landslide inventories critical for predictive modelling and hazard analysis.

CRediT authorship contribution statement

Kushanav Bhuyan: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kamal Rana:** Writing – review & editing, Software, Methodology, Data curation. **Ugur Ozturk:** Writing – review & editing, Investigation, Formal analysis. **Lorenzo Nava:** Writing – review & editing. **Ascanio Rosi:** Writing – review & editing, Investigation. **Sansar Raj Meena:** Writing – review & editing. **Xuanmei Fan:** Writing – review & editing. **Mario Floris:** Writing – review & editing, Supervision. **Cees van Westen:** Writing – review & editing, Supervision. **Filippo Catani:** Writing – review & editing, Supervision, Funding acquisition.

Code availability

Code is made available at a GitHub repository in https://github.com/kushanavbhuyan/Delineating-failure-kinematics. The code is an enhancement to the existing Landsifier library by Rana et al. (2022).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors express their sincere gratitude to Dr. Tolga Gorum and Dr. Hakan Tanyas for granting access to the Turkey landslide inventory. This work was supported by the National Science Fund for Distinguished Young Scholars of China (Grant No. 42125702) and the Natural Science Foundation of Sichuan Province, China (Grant No. 2022NSFSC003); "The Geosciences for Sustainable Development" project (Budget Ministero dell'Universitá e della Ricerca-Dipartimenti di Eccellenza 2023–2027 [CUP C93C23002690001] of the Department of Geosciences, University of Padova, USA; the research focus point "Earth and Environmental Systems" of the University of Potsdam, Germany; and the Sichuan Science and Technology Program, China (No. 2024JDHJ0038, 2024ZYD0140).

Appendix. Figures

See Figs. A.1–A.4.

Data availability

The data is shared with the manuscript at our GitHub repository at https://github.com/kushanavbhuyan/Delineating-failure-kinematics.

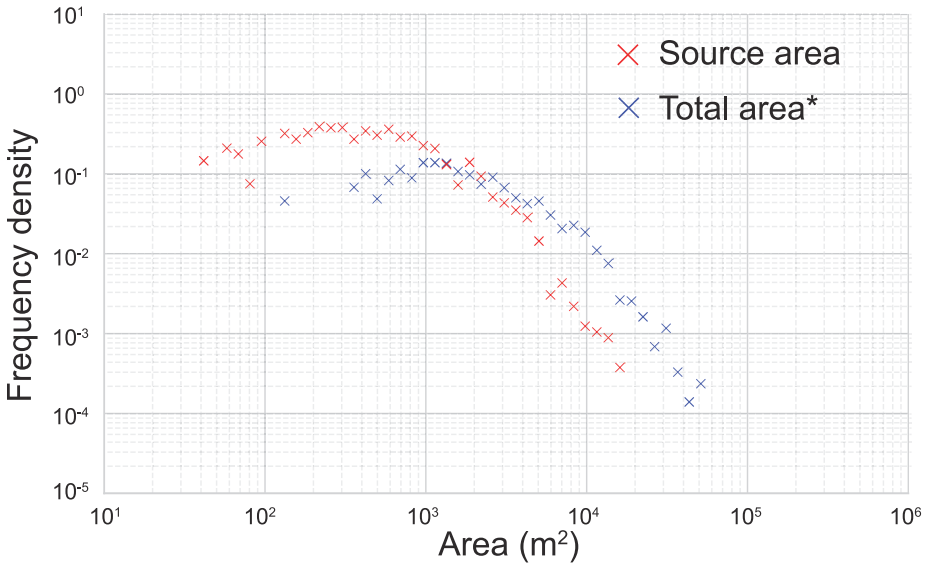


Fig. A.1. Frequency area distribution (FAD) for the landslides in Japan, Hokkaido affected by the 2018 seismic event. The FAD is determined for areas based on both the source area region (predicted by the model) and the total landslide body (indicated by *) based on the modified inventories from Wang et al. (2019) and Dou et al. (2020).

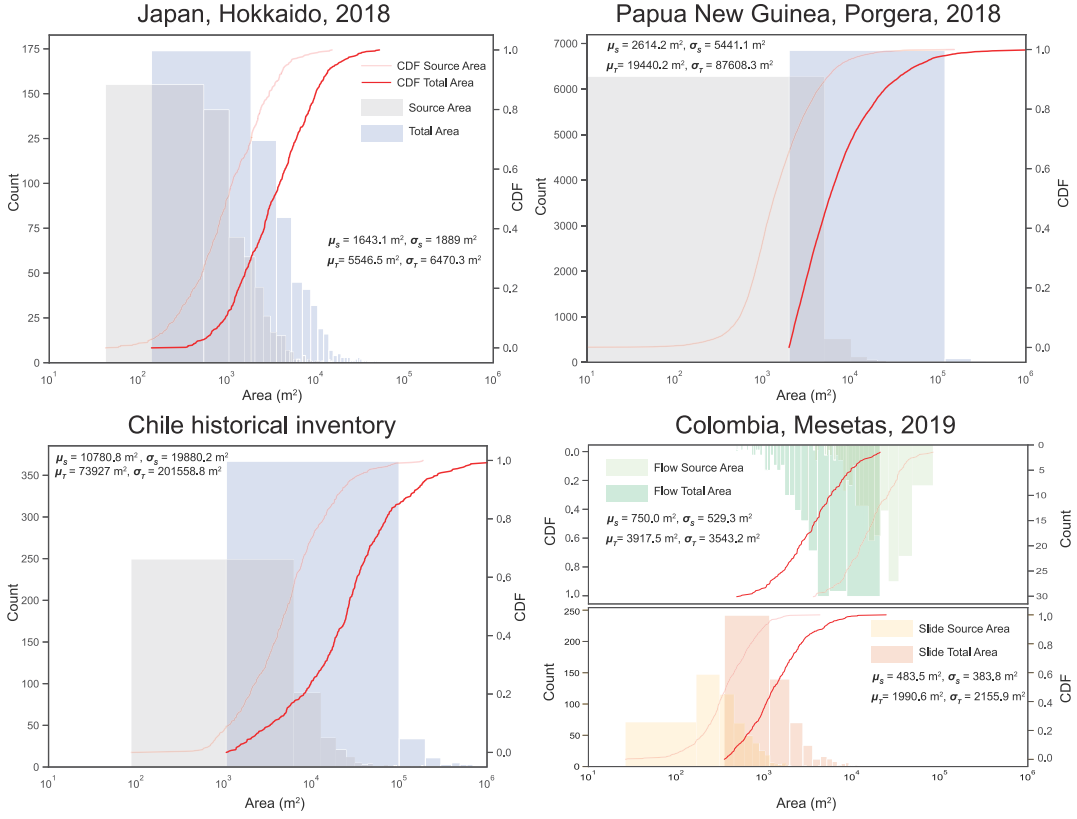


Fig. A.2. Statistics of source (μ_S, σ_S) and total landslide (μ_T, σ_T) geometric area based on the model predictions for Japan (Hokkaido), Papua New Guinea, Chile, and Colombia. For the region of Colombia, we also show the geometric area information based on their movement types (i.e., Slides and Flows).

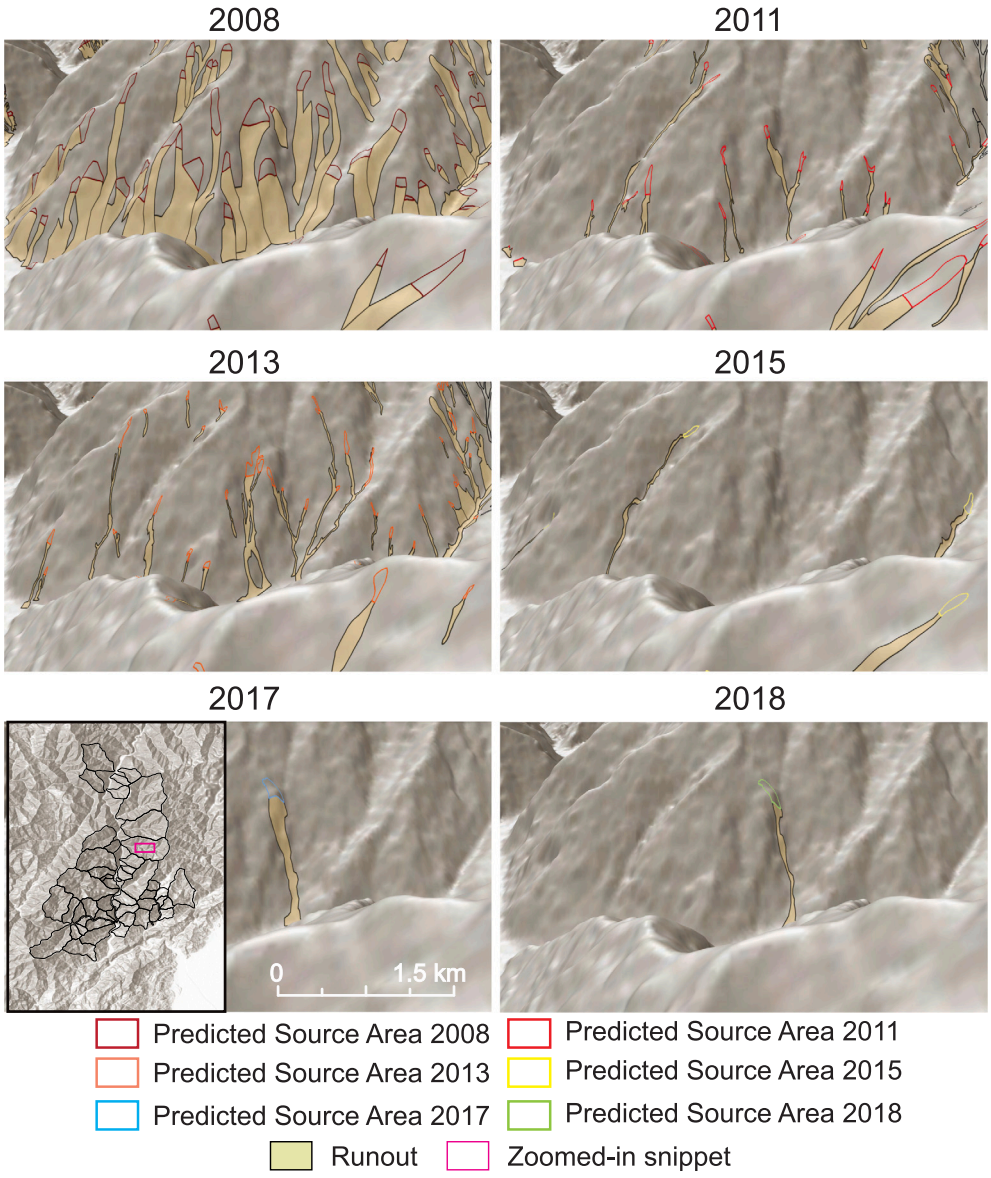


Fig. A.3. Model predictions for the 2008–2018 multi-temporal inventory of the Wenchuan 2008 earthquake and post-seismic landslides. Base map source: World Imagery, credits: Esri, Maxar, Earthstar Geographics, and the GIS User Community (Esri, 2023).

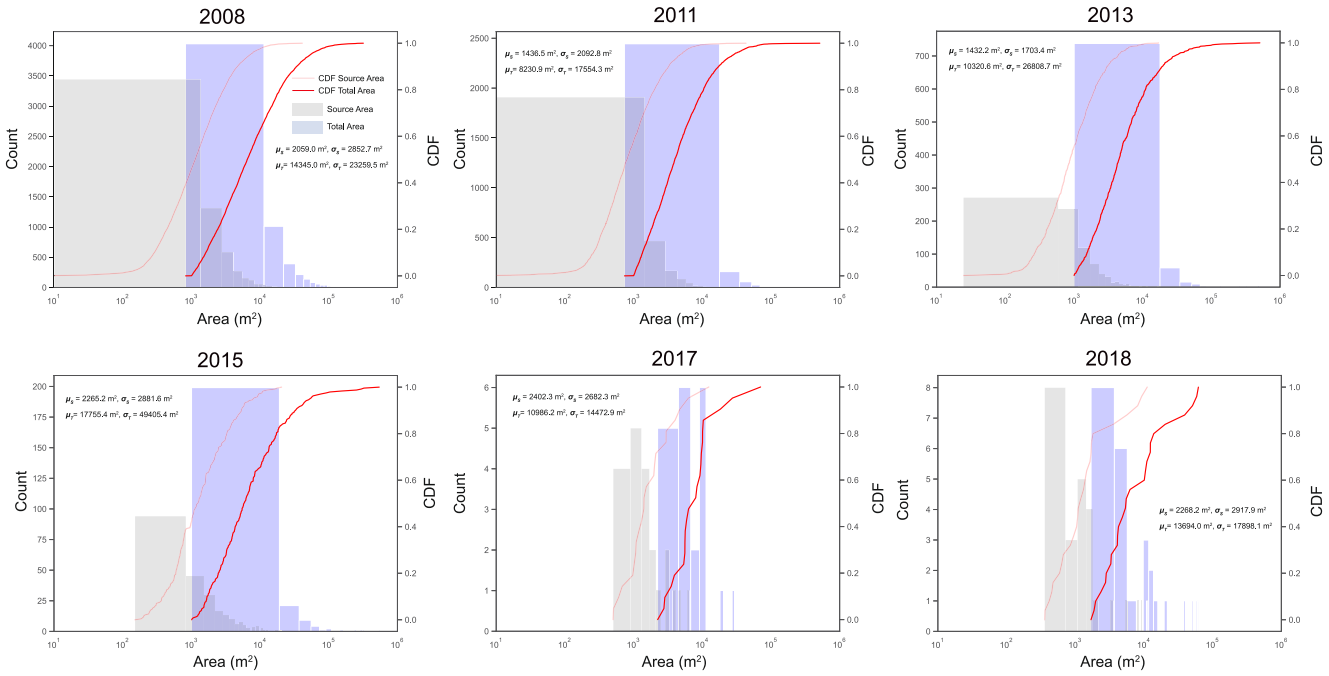


Fig. A.4. Statistics (mean and standard deviation) of source (μ_S, σ_S) and total landslide (μ_T, σ_T) geometric area based on the model predictions for the 2008–2018 multi-temporal inventory of Wenchuan 2008 earthquake and post-seismic landslides.

References

Ahmed, M., Titti, G., Trevisani, S., Borgatti, L., Francioni, M., 2024. Is higher resolution always better? Open-access DEM comparison for slope units delineation and regional landslide prediction. Nat. Hazards Earth Syst. Sci. Discuss. 2024, 1–32.

Amato, G., Palombi, L., Raimondi, V., 2021. Data–driven classification of landslide types at a national scale by using artificial neural networks. Int. J. Appl. Earth Obs. Geoinf. 104, 102549.

ARPAV, 2021. https://www.arpa.veneto.it/temi-ambientali/meteo/riferimenti/documenti/documenti-meteo/IPCCECAMBIAMENTICLIMATICIINVENETO.pdf. (Accessed 14 May 2023).

Baldo, M., Bicocchi, C., Chiocchini, U., Giordan, D., Lollino, G., 2009. LIDAR monitoring of mass wasting processes: The Radicofani landslide, Province of Siena, Central Italy. Geomorphology 105 (3–4), 193–201.

Baroň, I., Jelének, J., Klimeš, J., Dong, J.-J., Melichar, R., Šutjak, M., Chen, Y., Yang, C.-M., Zhang, E.-L., Méndez, J., et al., 2024. Source area morphometry and high depletion rate of landslides may indicate their coseismic origin. Eng. Geol. 330, 107424.

Behling, R., Roessner, S., Golovko, D., Kleinschmit, B., 2016. Derivation of long-term spatiotemporal landslide activity—A multi-sensor time series approach. Remote Sens. Environ. 186, 88–104.

Benson, C., Clay, E., Michael, F.V., Robertson, A.W., 2001. Dominica: Natural disasters and economic development in a small island state. In: Dominica: Natural Disasters and Economic Development in a Small Island State. p. 129.

Bhuyan, K., Rana, K., Ferrer, J.V., Cotton, F., Ozturk, U., Catani, F., Malik, N., 2024. Landslide topology uncovers failure movements. Nat. Commun. 15 (1), 2633.

Bhuyan, K., Tanyaş, H., Nava, L., Puliero, S., Meena, S.R., Floris, M., Van Westen, C., Catani, F., 2023. Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. Sci. Rep. 13 (1), 162.

Blahut, J., Van Westen, C.J., Sterlacchini, S., 2010. Analysis of landslide inventories for accurate prediction of debris-flow source areas. Geomorphology 119 (1–2), 36–51.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Can, T., Nefeslioglu, H.A., Gokceoglu, C., Sonmez, H., Duman, T.Y., 2005. Susceptibility assessments of shallow earthflows triggered by heavy rainfall at three catchments by logistic regression analyses. Geomorphology 72 (1–4), 250–271.

Carlsson, G., 2009. Topology and data. Bull. Amer. Math. Soc. 46 (2), 255–308.

Catani, F., Lagomarsino, D., Segoni, S., Tofani, V., 2013. Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues. Nat. Hazards Earth Syst. Sci. 13 (11), 2815–2831.

Chang, K.-J., Chan, Y.-C., Chen, R.-F., Hsieh, Y.-C., 2018. Geomorphological evolution of landslides near an active normal fault in northern Taiwan, as revealed by lidar and unmanned aircraft system data. Nat. Hazards Earth Syst. Sci. 18 (3), 709–727.

CHARIM, 2018. https://www.cdema.org/virtuallibrary/index.php/charim-hbook/why-charim. (Accessed 13 July 2024).

Crosta, G., Imposimato, S., Roddeman, D., 2003. Numerical modelling of large landslides stability and runout. Nat. Hazards Earth Syst. Sci. 3 (6), 523–538.

Cruden, D., Varnes, D., 1996. Landslide, Investigation and Mitigation. Transportation Research Board Special Report 247.

Dahal, A., Tanyaş, H., Lombardo, L., 2024a. Full seismic waveform analysis combined with transformer neural networks improves coseismic landslide prediction. Commun. Earth Environ. 5 (1), 75.

Dahal, A., Tanyas, H., van Westen, C., van der Meijde, M., Mai, P.M., Huser, R., Lombardo, L., 2024b. Space–time landslide hazard modeling via Ensemble Neural Networks. Nat. Hazards Earth Syst. Sci. 24 (3), 823–845.

Delgado, H.G., Chaparroa, J., Gamboa, C., Ruiz, G., 2021. Landslides triggered by the December 24, 2019 Mesetas (Meta, Colombia) Earthquake.

Doglioni, C., 1990. Thrust tectonics examples from the Venetian Alps. Stud. Geol. Camerti, Special Issue 117–129.

Dou, J., Yunus, A.P., Merghadi, A., Shirzadi, A., Nguyen, H., Hussain, Y., Avtar, R., Chen, Y., Pham, B.T., Yamagishi, H., 2020. Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning. Sci. Total Environ. 720, 137320.

Duman, T.Y., Çan, T., Emre, Ö., Keçer, M., Doğan, A., Ateş, Ş., Durmaz, S., 2005. Landslide inventory of northwestern Anatolia, Turkey. Eng. Geol. 77 (1–2), 99–114.

Esri, 2023. World imagery. https://services.arcgisonline.com/ArcGIS/rest/services/World_Imagery/MapServer. (Accessed 14 May 2023).

Fan, X., Scaringi, G., Domènech, G., Yang, F., Guo, X., Dai, L., He, C., Xu, Q., Huang, R., 2019. Two multi-temporal datasets that track the enhanced landsliding after the 2008 Wenchuan earthquake. Earth Syst. Sci. Data 11 (1), 35–55.

Farr, T.G., Kobrick, M., 2000. Shuttle Radar Topography Mission produces a wealth of data. EOS Trans. Am. Geophys. Union 81 (48), 583–585.

Francis, O., Fan, X., Hales, T., Hobley, D., Xu, Q., Huang, R., 2022. The fate of sediment after a large earthquake. J. Geophys. Res. Earth Surf. 127 (3), e2021JF006352.

Gallen, S.F., Clark, M.K., Godt, J.W., 2015. Coseismic landslides reveal near-surface rock strength in a high-relief, tectonically active setting. Geology 43 (1), 11–14.

Gatto, A., Martellozzo, F., Clò, S., Ciulla, L., Segoni, S., 2024. The downward spiral entangling soil sealing and hydrogeological disasters. Environ. Res. Lett..

Goetz, J., Kohrs, R., Parra Hormazábal, E., Bustos Morales, M., Araneda Riquelme, M.B., Henríquez, C., Brenning, A., 2021. Optimizing and validating the Gravitational Process Path model for regional debris-flow runout modelling. Nat. Hazards Earth Syst. Sci. 21 (8), 2543–2562.

Görüm, T., 2019. Landslide recognition and mapping in a mixed forest environment from airborne LiDAR data. Eng. Geol. 258, 105155.

Government of the Commonwealth of Dominica, 2017. Post-disaster Needs Assessment: Hurricane Maria: September 18, 2017. Government of the Commonwealth of Dominica Roseau, Dominica.

Guzzetti, F., Ardizzone, F., Cardinali, M., Rossi, M., Valigi, D., 2009. Landslide volumes and landslide mobilization rates in Umbria, central Italy. Earth Planet. Sci. Lett. 279 (3–4), 222–229.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.-T., 2012. Landslide inventory maps: New tools for an old problem. Earth-Sci. Rev. 112 (1–2), 42–66.

Heim, A., 1932. Bergsturz und menschenleben.
Hovius, N., Meunier, P., Lin, C.-W., Chen, H., Chen, Y.-G., Dadson, S., Horng, M.-J., Lines, M., 2011. Prolonged seismically induced erosion and the mass balance of a large earthquake. Earth Planet. Sci. Lett. 304 (3–4), 347–355.
Huang, R., Li, W., 2014. Post-earthquake landsliding and long-term impacts in the Wenchuan earthquake area, China. Eng. Geol. 182, 111–120.
Hungr, O., Leroueil, S., Picarelli, L., 2014. The Varnes classification of landslide types, an update. Landslides 11, 167–194.
Iadanza, C., Trigila, A., Starace, P., Dragoni, A., Biondo, T., Roccisano, M., 2021. IdroGEO: A collaborative web mapping application based on REST API services and open data on landslides and floods in Italy. ISPRS Int. J. Geo-Inf. 10 (2), 89.
Jaboyedoff, M., Oppikofer, T., Abellán, A., Derron, M.-H., Loye, A., Metzger, R., Pedrazzini, A., 2012. Use of LIDAR in landslide investigations: a review. Nat. Hazards 61, 5–28.
Ju, L.-Y., Zhang, L.-M., Xiao, T., 2023. Power laws for accurate determination of landslide volume based on high-resolution LiDAR data. Eng. Geol. 312, 106935.
Keck, J., Istanbulluoglu, E., Campforts, B., Tucker, G., Horner-Devine, A., 2023. A landslide runout model for sediment transport, landscape evolution and hazard assessment applications. EGUsphere 2023, 1–42.
Krishnan, S., Crosby, C., Nandigam, V., Phan, M., Cowart, C., Baru, C., Arrowsmith, R., 2011. OpenTopography: a services oriented architecture for community access to LIDAR topography. In: Proceedings of the 2nd International Conference on Computing for Geospatial Research & Applications. pp. 1–8.
Lai, J.-S., 2020. Separating landslide source and runout signatures with topographic attributes and data mining to increase the quality of landslide inventory. Appl. Sci. 10 (19), 6652.
Larsen, I.J., Montgomery, D.R., 2012. Landslide erosion coupled to tectonics and river incision. Nat. Geosci. 5 (7), 468–473.
Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., Huser, R., 2020. Space-time landslide predictive modelling. Earth-Sci. Rev. 209, 103318.
Lucas, A., Mangeney, A., Ampuero, J.P., 2014. Frictional velocity-weakening in landslides on Earth and on other planetary bodies. Nat. Commun. 5 (1), 3417.
Marinelli, A., Medici, C., Rosi, A., Tofani, V., Bianchini, S., Casagli, N., 2022. Shallow landslides and rockfalls velocity assessment at regional scale: a methodology based on a morphometric approach. Geosciences 12 (4), 177.
Martha, T.R., Roy, P., Jain, N., Khanna, K., Mrinalni, K., Kumar, K.V., Rao, P., 2021. Geospatial landslide inventory of India—an insight into occurrence and exposure on a national scale. Landslides 18 (6), 2125–2141.
Meena, S.R., Puliero, S., Bhuyan, K., Floris, M., Catani, F., 2022. Assessing the importance of conditioning factor selection in landslide susceptibility for the province of Belluno (region of Veneto, northeastern Italy). Nat. Hazards Earth Syst. Sci. 22 (4), 1395–1417.
Mergili, M., Fischer, J.-T., Krenn, J., Pudasaini, S.P., 2017. r. avaflow v1, an advanced open-source computational framework for the propagation and interaction of two-phase mass flows. Geosci. Model Dev. 10 (2), 553–569.
Meyrat, G., McArdell, B., Ivanova, K., Müller, C., Bartelt, P., 2022. A dilatant, two-layer debris flow model validated by flow density measurements at the Swiss illgraben test site. Landslides 19 (2), 265–276.
Munch, E., 2017. A user's guide to topological data analysis. J. Learn. Anal. 4 (2), 47–61.
Ouyang, Z., Zhou, C., Xie, J., Zhu, J., Zhang, G., Ao, M., 2023. SRTM DEM correction using ensemble machine learning algorithm. Remote Sens. 15 (16), 3946.
Ozturk, U., Pittore, M., Behling, R., Roessner, S., Andreani, L., Korup, O., 2021. How robust are landslide susceptibility estimates? Landslides 18, 681–695.
Parker, R.N., Densmore, A.L., Rosser, N.J., De Michele, M., Li, Y., Huang, R., Whadcoat, S., Petley, D.N., 2011. Mass wasting triggered by the 2008 Wenchuan earthquake is greater than orogenic growth. Nat. Geosci. 4 (7), 449–452.
Rana, K., Malik, N., Ozturk, U., 2022. Landsifier v1. 0: a Python library to estimate likely triggers of mapped landslides. Nat. Hazards Earth Syst. Sci. Discuss. 2022, 1–26.
Rana, K., Ozturk, U., Malik, N., 2021. Landslide geometry reveals its trigger. Geophys. Res. Lett. 48 (4), e2020GL090848.
Roback, K., Clark, M.K., West, A.J., Zekkos, D., Li, G., Gallen, S.F., Chamlagain, D., Godt, J.W., 2018. The size, distribution, and mobility of landslides caused by the 2015 Mw7. 8 Gorkha earthquake, Nepal. Geomorphology 301, 121–138.
Rouse, C., 1990. The mechanics of small tropical flowslides in Dominica, West Indies. Eng. Geol. 29 (3), 227–239.
Rouse, W., Reading, A., Walsh, R., 1986. Volcanic soil properties in Dominica, West Indies. Eng. Geol. 23 (1), 1–28.
Sassa, K., 2005. Landslide disasters triggered by the 2004 Mid-Niigata Prefecture earthquake in Japan. Landslides 2 (2), 135–142.
Sekiguchi, T., Sato, H., 2006. Feature and distribution of landslides induced by the mid niigata prefecture earthquake in 2004, Japan. J. Jpn. Landslide Soc. 43 (3), 142–154.
Simard, M., Denbina, M., Marshak, C., Neumann, M., 2024. A global evaluation of radar-derived digital elevation models: SRTM, NASADEM, and GLO-30. J. Geophys. Res. Biogeosci. 129 (11), e2023JG007672.
Smith, T., Rheinwalt, A., Bookhagen, B., 2019. Determining the optimal grid resolution for topographic analysis on an airborne lidar dataset. Earth Surf. Dyn. 7 (2), 475–489.
Sorbino, G., Sica, C., Cascini, L., 2010. Susceptibility analysis of shallow landslides source areas using physically based models. Nat. Hazards 53, 313–332.
Stumvoll, M., Schmaltz, E., Glade, T., 2021. Dynamic characterization of a slow-moving landslide system–assessing the challenges of small process scales utilizing multi-temporal TLS data. Geomorphology 389, 107803.
Tang, X., Tu, Z., Wang, Y., Liu, M., Li, D., Fan, X., 2022. Automatic detection of coseismic landslides using a new transformer method. Remote Sens. 14 (12), 2884.
Tanyaş, H., Kirschbaum, D., Görüm, T., van Westen, C.J., Lombardo, L., 2021. New insight into post-seismic landslide evolution processes in the tropics. Front. Earth Sci. 9, 700546.
Tanyaş, H., van Westen, C.J., Allstadt, K.E., Anna Nowicki Jessee, M., Görüm, T., Jibson, R.W., Godt, J.W., Sato, H.P., Schmitt, R.G., Marc, O., et al., 2017. Presentation and analysis of a worldwide database of earthquake-induced landslide inventories. J. Geophys. Res. Earth Surf. 122 (10), 1991–2015.
Tarolli, P., 2014. High-resolution topography for understanding earth surface processes: Opportunities and challenges. Geomorphology 216, 295–312.
Tarquini, S., Isola, I., Favalli, M., Mazzarini, F., Bisson, M., Pareschi, M.T., Boschi, E., 2007. TINITALY/01: a new triangular irregular network of Italy. Ann. Geophys..
Tauzin, G., Lupo, U., Tunstall, L., Pérez, J.B., Caorsi, M., Medina-Mardones, A.M., Dassatti, A., Hess, K., 2021. giotto-tda:: A topological data analysis toolkit for machine learning and data exploration. J. Mach. Learn. Res. 22 (39), 1–6.
Taylor, F.E., Malamud, B.D., Witt, A., Guzzetti, F., 2018. Landslide shape, ellipticity and length-to-width ratios. Earth Surf. Process. Landf. 43 (15), 3164–3189.
Tebbens, S., 2020. Landslide scaling: a review. Earth Space Sci. 7 (1), e2019EA000662.
Teza, G., Pesci, A., Genevois, R., Galgaro, A., 2008. Characterization of landslide ground surface kinematics from terrestrial laser scanning and strain field computation. Geomorphology 97 (3–4), 424–437.
Trigila, A., Iadanza, C., Guerrieri, L., Hervás, J., 2007. The IFFI project (Italian landslide inventory): Methodology and results. In: Guidelines for Mapping Areas at Risk of Landslides in Europe, vol. 23, ISPRA Rome, Italy, p. 15.
Van Westen, C., 2016. National Scale Landslide Susceptibility Assessment for Dominica. University of Twente.
van Westen, C., Zhang, J., 2018. Landslides and Floods Triggered by Hurricane Maria (18 September, 2017) in Dominica, Digital or Visual Products, UNITAR-UNOSAT. Wiley, London, UK.
Wang, F., Fan, X., Yunus, A.P., Siva Subramanian, S., Alonso-Rodriguez, A., Dai, L., Xu, Q., Huang, R., 2019. Coseismic landslides triggered by the 2018 Hokkaido, Japan (M w 6.6), earthquake: spatial distribution, controlling factors, and possible failure mechanism. Landslides 16, 1551–1566.
Wichmann, V., 2017. The Gravitational Process Path (GPP) model (v1. 0)–a GIS-based simulation framework for gravitational processes. Geosci. Model Dev. 10 (9), 3309–3327.
Ye, T., Jiang, Q., Zhang, C., Ma, Z., Li, C., 2024. High-speed long-runout landslide scraping and entrainment effects: A case study on shuicheng landslide. Eng. Geol. 341, 107722.
Yu, B., Wu, Y., Chu, S., 2014. Preliminary study of the effect of earthquakes on the rainfall threshold of debris flows. Eng. Geol. 182, 130–135.
Yunus, A.P., Xinyu, C., Catani, F., Subramaniam, S.S., Fan, X., Jie, D., Sajinkumar, K., Gupta, A., Avtar, R., 2023. Earthquake-induced soil landslides: volume estimates and uncertainties with the existing scaling exponents. Sci. Rep. 13 (1), 8151.