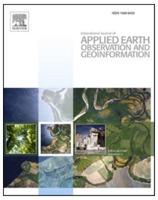
International Journal of Applied Earth Observation and Geoinformation 140 (2025) 104608

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



Machine learning and clustering for supporting the identification of active landslides at national scale

Camilla Medici [a,*], Alessandro Novellino [b], Claire Dashwood [b], Silvia Bianchini [a]

[a] Earth Sciences Department, University of Firenze, Via La Pira, 4 50121 Firenze Italy
[b] British Geological Survey, Environmental Science Centre, Keyworth, Nottingham NG12 5GG, UK

ARTICLE INFO

Keywords:
EGMS
Machine learning
Clustering
Landslide
Slope units
Great Britain

ABSTRACT

Landslides are one of the major geohazards causing significant economic damage and loss of life, with impacts expected to increase due to climate change. Landslide inventory maps (LIMs) are essential for risk management and reduction, but they usually remain an overlooked issue, especially over very large areas, i.e. at a regional or national level. Nowadays, extensive interferometric satellite radar data with wide area coverage are profitably available, but their potential can be not fully exploited due to the challenge of managing them. In this context, we used space-borne advanced Interferometric Synthetic Aperture Radar (InSAR) data at a national scale, in order to create a useful database of active slope instability movement areas to rely on where the landslide inventory map is missing or largely incomplete. Specifically, we provide insights into a new approach, proposing a national-scale method that combines Machine Learning (ML) and clustering tools, which are crucial to manage a huge amount of data. The proposed methodology has been applied to Great Britain. The use of a ML algorithm, specifically Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) for noise filtering, has allowed the InSAR dataset to be reduced from approximately 6.5 million points to about 3.8 million points per component. Thus, implementing ML along with Slope Units for geomorphological reliability, and tools for identifying and classifying active deformation areas yields an InSAR landslide inventory map. Through this process, 336,557 Slope Units have been classified; of these, 5% show discrepancies between landslide inventory and InSAR data. Identifying these areas, along with those classified as landslides by both datasets, is crucial for risk management as it highlights areas that require closer inspections.

1. Introduction

Landslides, one of the most catastrophic geohazards, cause significant economic damage and loss of life every year. The Emergency Events Database (Delforge et al., 2023), https://public.emdat.be/) reports that landslides have resulted in around 73,000 fatalities and $24 billion in damages worldwide from 1900 to 2023. Therefore, there is a critical need to enhance our understanding of landslide risks and to develop strategies to reduce their human and socio-economic impacts.

Earth Observation (EO) techniques, particularly Multi-Temporal Interferometry Synthetic Aperture Radar (MTInSAR), are now well established and adopted approaches for the analysis and monitoring of slow-moving landslides (Novellino et al., 2024). MTInSAR has proven its effectiveness in landslide studies by enabling analysis over large areas but also detailed investigations, and monitoring of inaccessible areas reducing time and costs. Moreover, the availability of extensive historical datasets allows for the analysis of long-term trends in instability phenomena. Particularly in this context, an important turning point has been the large amount of free data acquired by the Sentinel-1 (S1) satellite constellation with worldwide coverage, millimetre accuracy and a short revisiting time, combined with the advancements in computational capabilities and processing techniques (Crosetto et al., 2016; Showstack, 2014). InSAR data has thus been widely used to date for deriving landslide hazard and risk information (Bianchini et al., 2017; Intrieri et al., 2018; Raspini et al., 2018; Solari et al., 2020). InSAR data applications for landslides analysis cover a wide range of uses. For example, they have been extensively employed for landslide mapping, particularly to update landslide inventories by refining information on the perimeter, state of activity, and intensity (Cascini et al., 2010; Cigna et al., 2013; Righini et al., 2012; Strozzi et al., 2013). Additionally, these data are often integrated with optical imagery (Farina et al., 2006) or in situ instruments (Bovenga et al., 2012; Cascini et al., 2010; Colesanti

---

* Corresponding author.
E-mail address: camilla.medici@unifi.it (C. Medici).

https://doi.org/10.1016/j.jag.2025.104608
Received 25 July 2024; Received in revised form 17 March 2025; Accepted 15 May 2025

et al., 2003; Tofani et al., 2013) to enhance the monitoring of landslide phenomena.

In this context, it is noteworthy that the starting point for any landslide risk assessment is a Landslide Inventory Map (LIM), a comprehensive database recording information on past landslides, including their location, size, type, date of occurrence and state of activity. Creating and keeping up-to-date LIMs is a time-consuming and costly task which often relies on ground campaigns and/or aerial photo interpretation (Corominas et al., 2013; Guzzetti et al., 2012). In literature, recent studies have used InSAR data to update or analyse landslide inventories over large areas (Confuorto et al., 2023; Rosi et al., 2018) as far as we can nowadays rely on the unprecedented amount of ground deformation data provided by the European Ground Motion Service (EGMS – https://egms.land.copernicus.eu/) within the Copernicus Land Monitoring Service (CLMS).

In this context, the adoption of semi-automatic and automatic techniques has become essential due to the extraordinary amount of data, especially on a national scale. Therefore, recent works have relied on Machine Learning (ML, Prakash et al., 2020; Tehrani et al., 2022), Deep Learning (DL, Bhuyan et al., 2023; Nava et al., 2022) or spatial clustering (Barra et al., 2017; Festa et al., 2022; Han et al., 2023) approaches for automatically analyse landslides. Throughout the approach developed in this work, we have used ML techniques to semi-automatically filter national scale datasets. This approach significantly reduced the need for manual intervention and streamlined data management while enabling the identification of slopes that can be associated with landslides. Additionally, we took into account the spatio-temporal characteristics of the EGMS datasets along with the geomorphometric characteristics of the slopes over the area of interest. We have selected Slope Units (SUs) as basic mapping units (e.g., Bryce et al., 2025) since they are defined by geomorphological criteria and bounded by drainage lines (Alvioli et al., 2016). They provide several advantages over gridded units, such as better capturing terrain geometry, improved incorporation of geospatial landslide-occurrence data in different formats and are widely employed for landslide analysis (e.g., Chang et al., 2023; Huang et al., 2019; Lombardo et al., 2020).

The resulting map from the procedure proposed in this study is a useful product providing information about the presence and absence of landslides over the whole Great Britain (GB) territory according to both the existing landslide inventory and InSAR data. We chose GB given the availability of a single and consistent National Landslide Database (NLD), managed by a single organisation, the British Geological Survey (BGS). The existence of this type of inventory reduces the issues associated with the presence of multiple databases at different scales and compiled with different methods.

This work is providing insights into a new facet: a first application at national scale of ML and clustering tools for supporting active landslides identification.

2. Study area

GB extends over an area of ~230,000 km^2 and information derived from the 50 m NEXTMap® DTM (https://catalogue.ceda.ac.uk/uuid/8f6e1598372c058f07b0aeac2442366d) shows that most of the landmass is dominated by gentle topography and steepness of slopes in mountain regions generally does not exceed ~ 15–20° (Cigna et al., 2014). Based on the topography, the SAR visibility analysis at national scale, informing about InSAR feasibility for mapping landslides (Del Soldato et al., 2021), shows that only less than 20 km^2 of GB surface is affected by shadow and layover (Novellino et al., 2017) so making InSAR information valuable over most of the territory.

In terms of landsliding, Great Britain is a low-risk environment with small scale failures and low fatality rates (Pennington et al., 2015), in the order of 25 fatalities over the last 50 years, relatively low in comparison with other European countries (Haque et al., 2016). Economic impacts of landslides are difficult to quantify, Gibson et al. (2013) suggested an annual cost of £10 million whilst more recent estimates suggest the impact on the rail network alone is almost £40 million a year, as revealed by the New Civil Engineer (NCE, https://www.newcivilengineer.com/).

Spatial analysis of landslide events has led to group GB territory in different hierarchical landslide domains (https://www.bgs.ac.uk/geology-projects/shallow-geohazards/landslide-domains/), which represent areas of similar physiographic, meteorological, climatic and geological characteristics that shaped the style of movement (Dashwood et al., 2017). Such domains provide preliminary information on where InSAR can be effective for mapping landslides in terms of style of motion and velocity. For example, areas with compound/complex landslides in the Cambrian Mountains and Pennines Mountains are particularly suitable compared to predominantly rock falls areas in the Northwest Highlands and Grampian Mountains (Fig. 1).

3. Materials and methods

The proposed methodology relies on a complete multi-scale approach starting from a country-wide mosaicked mean deformation velocity map and ending up with a classified active motion areas database. Results are then presented and discussed through site-specific, detailed scale analysis.

The flow chart of the methodological approach, depicted in Fig. 2, is broken down into multiple parts: the gathering of EGMS data, the ML component where HDBSCAN has been used to filter the original EGMS data, the combination of the filtered InSAR data with the ADA inside the SUs, the spatio-temporal analysis and the combination with the NLD.

The procedure required the following datasets: the BGS NLD (https://www.bgs.ac.uk/geologyprojects/landslides/national-landslide-database/, Section 3.1), the land cover map (https://www.ceh.ac.uk/data/ukceh-land-cover-maps, Section 3.2), the Digital Terrain Model (DTM) and DTM-derived products (Section 3.2), and the EGMS InSAR data (Section 2.3).

3.1. GB national landslide database

The BGS' NLD (https://www.bgs.ac.uk/geology-projects/landslides/national-landslide-database/) is the most extensive source of information on landslides in Great Britain. The database currently holds over 18,000 records and it is regularly updated as new landslide events occur. The inventory pulls together information from different sources such as maps data, journal articles, research reports and site investigations (Foster et al., 2012). The main source of information for modern events in the NLD comes from news reports, social media, and field works and it is, inevitably, biased towards the most populated areas (Pennington et al., 2015) with EO-techniques only deployed following major storm events.

As can be observed in Fig. 3, population distribution has an impact on landslide density with location of landslides on slopes around major urban centre. On one hand, Scotland, suffers fewer landslides (13 %) than may be anticipated from its topography and geology, and on the other hand, England has the largest percentage (72 %) despite its morphology. This can be attributed to the surface area of these territories and their population density. The NLD suffers from other two limitations: it is point based, with the point usually corresponding to the source area, and the state of activity is unknown in most cases. Aware of these limitations, the proposed approach does not rely on the pre-existing landslide database as its starting point. Instead, the NLD is employed only at the end of the procedure as a reference for comparison.

Considering the applicability of InSAR according to its visibility (Del Soldato et al., 2021), we filter the NLD data, retaining only slow movements recorded between 2018 and 2021.

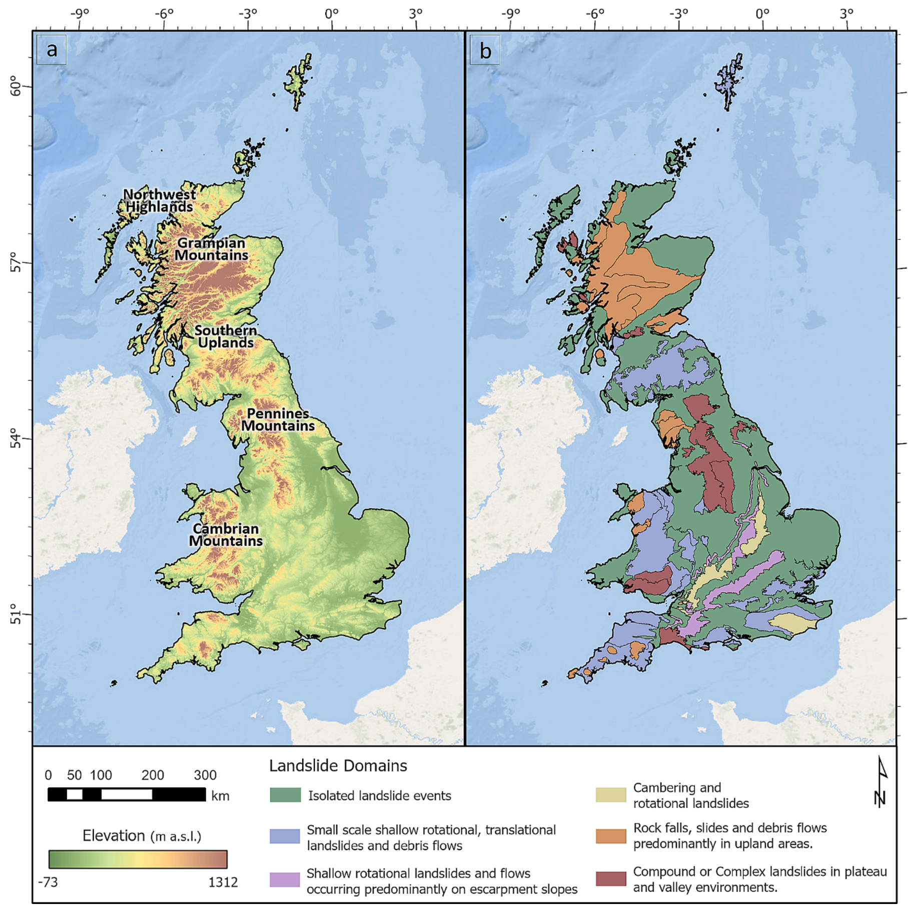


Fig. 1. A) GB elevation with the main mountain chains; b) landslide domains map (https://www.bgs.ac.uk/geology-projects/shallow-geohazards/landslide-domains/).

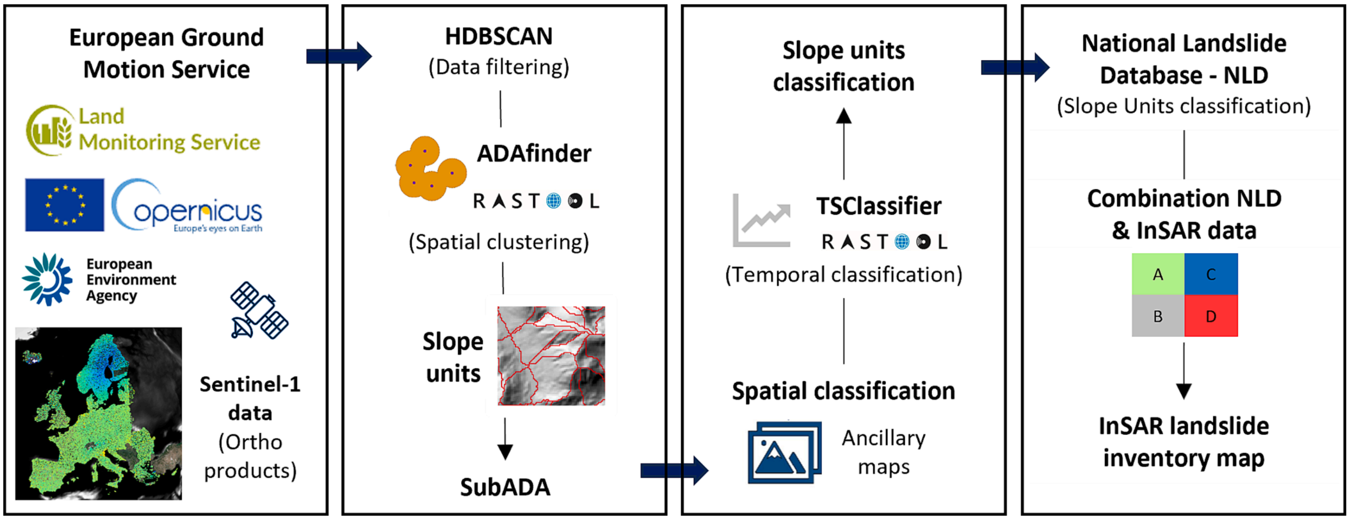


Fig. 2. Flow chart of the methodological approach.

3.2. Ancillary data

The chosen territorial units for the present study are the Slope Units (SUs) produced with a GRASS GIS script, the same used by Bryce et al., 2025 for assessing landslide susceptibility in Scotland. While pixels are often used as the basic unit in the scientific literature (Reichenbach et al., 2018), this approach entails numerous approximations and, most importantly, does not consider the topographic and morphological characteristics that should instead underpin any analysis related to geohazards. Indeed, SUs, being bounded by drainage lines, are characterised by morphometric and environmental homogeneity, as described in Alvioli et al. (2016), and thus can be more realistic units when we need to map landslides. In the process of creating the SUs, a systematic approach was incorporated to exclude flat areas based on terrain

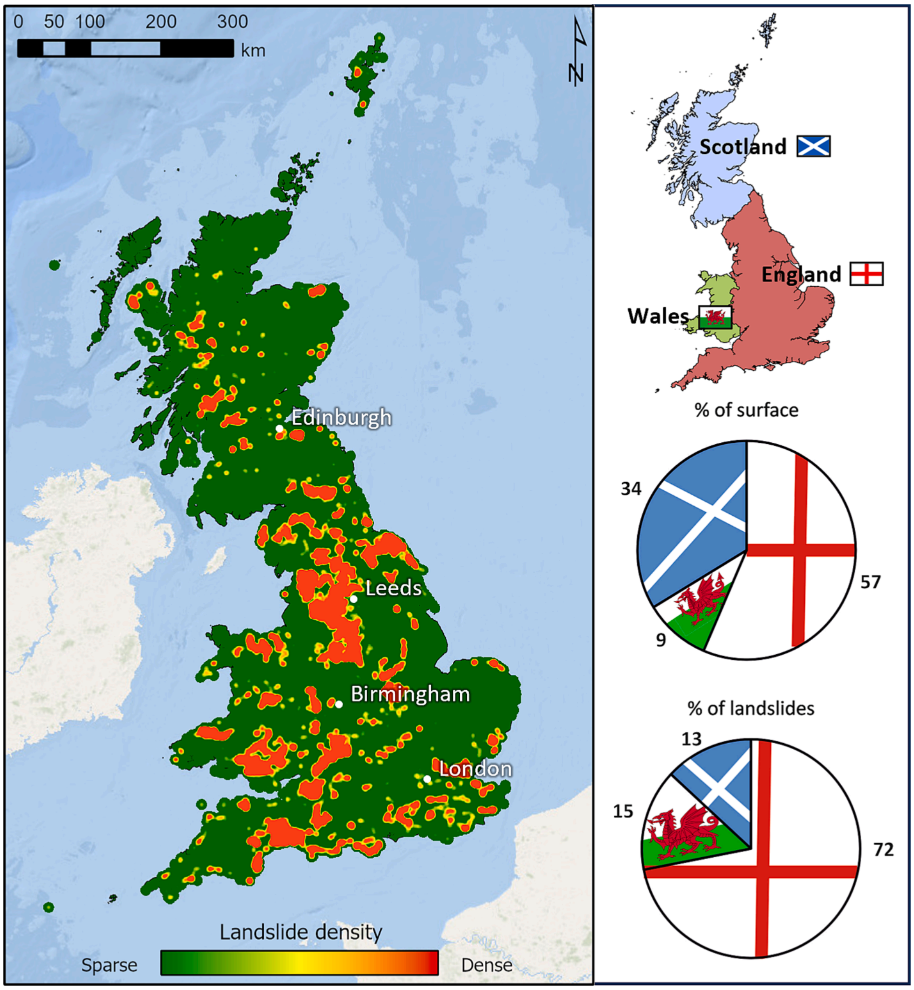


Fig. 3. GB landslide density and pie charts of the percentage of surface and landslides among England, Scotland and Wales.

classification using TPI (Topographic Position Index). Essentially, if the SU primarily falls within a class associated with flat topographies, it is not generated in those areas initially.

A total of 336,557 SUs have been derived by the University of Twente over Great Britain using the 50 m NEXTMap® DTM. The same DTM has been used to extract the slope and the eastness, namely the degree to which a particular area is oriented towards the East. The values range from −1 to 1, where positive values indicate that the surface faces eastward, while negative values denote a western exposure. In particular, these morphometric parameters have been used alongside the InSAR data to classify active deformation areas.

Finally, the 2021 land cover map at 10 m resolution, provided by the UK Centre for Ecology & Hydrology, has been used as part of the classification of the active deformation areas (https://www.ceh.ac.uk/data/ukceh-land-cover-maps).

3.3. Satellite InSAR data

The InSAR data comes from the European Ground Motion Service (EGMS) (https://egms.land.copernicus.eu/), a Copernicus Land Monitoring Service (CLSM) programme carried out under the auspices of the European Environment Agency (EEA) (Costantini et al., 2021; Crosetto et al., 2021, 2020). The data are derived from the multi-temporal interferometric analysis of high-resolution Sentinel-1 radar images. Specifically, Ortho L3 data have been used in this work. The selected data includes time series and average velocities of the vertical and horizontal (East-West) components, resampled to a 100 m grid and have been downloaded through the EGMStream (https://cpc.unifi.it/EGMStream_v1.zip), a desktop application developed by Festa and Del Soldato, (2023) that facilitates EGMS data management. It is indeed a user-friendly interface that allows downloading data from the EGMS Explorer and automatically converting them from csv format into geospatial databases, allowing also to crop EGMS data on the area of interest.

The EGMS data collected throughout GB consists of approximately 6.5 million points for each velocity component, i.e. horizontal and vertical measurements (Fig. 4). Despite Level 3 not corresponding to the original resolution of the InSAR datasets, we used such lower-resolution input as we are testing a new methodology at large scale. Indeed, using the original resolution would need a further level of analysis, movements along the satellite Line of Sight vs. the steepest slope direction, that we plan to undertake in future work.

3.4. Methodology

The proposed methodology leverages a hybrid approach that integrates automatic techniques, with expert judgment. Thus, the procedure starts from the EGMS data and comprises a series of sequential data analysis steps which employ ML algorithm and clustering tool to reduce computation times and manage large quantities of data as well as the expert judgment necessary for the analysis and employment of the results obtained from automatic techniques.

To address the challenge of InSAR data volume and make it exploitable, the initial step of the data analysis consists of applying a ML

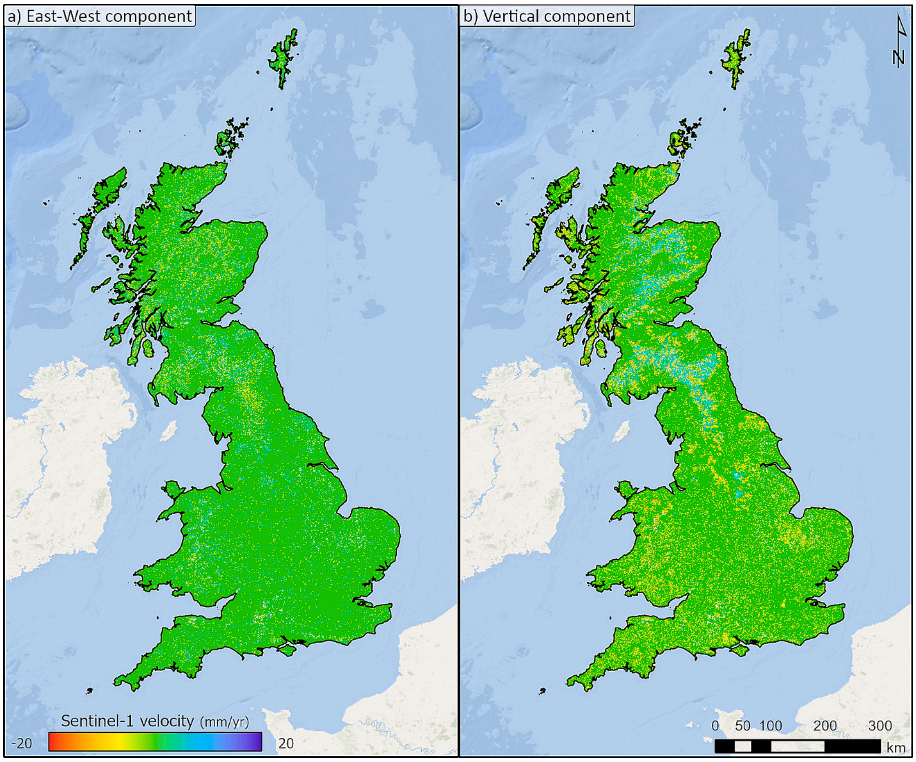


Fig. 4. S1 velocity of the a) horizontal (East-West) and b) vertical component over GB.

algorithm, the HDBSCAN, (McInnes et al., 2017) with the purpose of identifying and removing points classified as noise or outliers within the EGMS dataset. HDBSCAN, which stands for Hierarchical Density-Based Spatial Clustering of Applications with Noise, is a spatial clustering algorithm that operates based on the concept of density reachability, which quantifies how close a point is to a cluster (Rygus et al., 2023). It is based on the DBSCAN algorithm, but it performs a more robust clustering through a condensed tree representation of the data, a hierarchical approach. In GB, the algorithm has been applied within each SU with at least three measurement points. The SU limits have been integrated because they represent a spatial boundary with a morphological meaning. The clustering groups measurement points close in space and with similar velocity patterns. The purpose of applying this ML algorithm is to filter out outliers, i.e., points that cannot be considered reliable, thereby reducing the input data for subsequent steps. This objective is achieved by defining the minimum cluster size and minimum samples parameters, whose values are determined by iteratively running the algorithm until the best combination is obtained based on the relative validity function.

The data resulting from HDBSCAN have been used as an input for the subsequent phase, where unstable areas have been identified. This is grounded in the concept that a cluster of points is inherently more significant, reliable, and representative of a phenomenon than a single point (Barra et al., 2017; Montalti et al., 2019; Tomás et al., 2019). For this purpose, we have used a tool developed as part of a European project named *EGMS RASTOOL* (https://rastool.cttc.es/), which provides a set of tools for an easy exploitation of EGMS data including ADAFinder. The latter allows for the semi-automatic detection of the Active Deformation Areas (ADAs) from MTInSAR-derived displacement maps, significantly reducing the time required for the analysis of extensive datasets (Barra et al., 2017). The so-called ADAs are unstable areas characterised by a minimum number of measurement points within a specific clustering radius and with velocity values greater than a certain threshold.

However, ADAs delimit areas regardless of the geomorphological characteristics, so ADAs have then been intersected with SUs and thus became possibly representative of potential slope movements. At this stage the unstable areas, intersected and grouped by each SU, could be related to different causes; therefore, considering the purpose of the work, a classification is necessary to filter only those areas potentially related to landslide phenomena. The classification of the unstable areas has been carried out based on slope angle, land cover map and consistency between slope eastness and horizontal velocity as to consider only downslope movements. Moreover, for each unstable area associated with landslide phenomena, the mean trend of the displacement time series has been assessed by using the TS classifier tool (https://rastool.cttc.es/), developed within the *EGMS RASTOOL* project. It is a user-friendly tool which by leveraging minimal input parameters allows to classify the mean time series trend as accelerating, decelerating or constant. Therefore, this tool enabled the addition of an important temporal characterisation for the SUs affected by landslide active deformation areas.

Subsequently, each SU has been classified considering a SU as landslide-prone if it encompasses at least one landslide point of the NLD. Conversely, regarding the InSAR data, the classification of the SUs is based on what is termed ‘ADA density’, representing the percentage of the SU area occupied by active deformation due to landslides. Through a tuning against well-known landslides, we determined the following two criteria to identify landslide SUs:

- with an ADA extension > 30 % of the SU;
- with an ADA extension between 10 % and 30 % of the SU, we need 5 mm/year for the average velocity or 10 mm/year for the maximum velocity of ADA at least.

In the end, the information derived from the NLD and the InSAR data has been combined to produce a map wherein each SU depicts the presence or absence of landslides according to both methodologies.

Because of the national purpose of the study, which imposes some limitations, the assessment of the proposed methodology and its accuracy is intrinsically restricted. Furthermore, the absence of ground-based data has precluded the validation of the obtained results. An analysis of well-known landslide sites has been done to address this problem and evaluate the approach.

4. Results

The first phase of the procedure consists in the application of the HDBSCAN algorithm to both horizontal and vertical velocity datasets. To achieve the best result, multiple combinations of the algorithm's required parameters have been run, specifically the minimum cluster size and the minimum samples, selecting the best combination based on the relative validity function, which considers the cohesion within clusters and the separation between clusters (Rygus et al., 2023). Among the tested values, the combination that has yielded the highest relative validity score is a minimum cluster size of 3 and a minimum samples value of 1. In this study, HDBSCAN has been applied over the GB to identify points that do not adhere to the clustering criteria. The removal of such points, recognised as noise, has constituted the primary objective of this initial phase of data analysis, leading to a reduction in the initial datasets by almost half, from approximately 6.5 million points to about 3.8 million points per component.

The subsequent identification of unstable areas through the clustering tool, ADAFinder, required a rapid analysis of the pre-filtered EGMS data; this, combined with values reported in the literature, enabled the selection of the few input parameters needed by the tool. Specifically, a velocity threshold of 2.5 mm/yr, corresponding to twice the standard deviation of the velocity of the whole InSAR population (Fig. 5a, Fig. 5b), a minimum number of measurement points equal to 3, and a clustering radius of 100 m have been chosen. The latter value is due to the use of ortho EGMS data resampled on a 100 m grid. Therefore, by simply selecting these few input parameters, the tool provides, for both velocity datasets, the unstable areas corresponding to the set criteria (Fig. 5c).

After intersecting the identified ADAs with the SUs to attribute a geomorphological significance to the unstable areas, the resulting 'sub-ADAs' have been filtered to identify those associated with landslide phenomena. Therefore, we have isolated unstable areas with slope angle greater than 5°, which are more likely to be associated with landslide. Then the consistency between slope eastness and velocity has been taken into account, this pertains to a basic concept associated with InSAR. Specifically, in the quest for landslide detection, and thus downslope movement, it is imperative that the eastness and horizontal velocity values representative of the inspected area share the same sign. Moreover, the last parameter involved in the classification is the land cover map, which has been used to filter out peatbogs. Peatbogs extend over large areas of Great Britain (approximately 9,490 km^2, https://www.ceh.ac.uk/data/ukceh-land-cover-maps) and, being particularly prone to vertical movements, have been excluded (Alshammari et al., 2020; Bradley et al., 2021).

After the unstable areas related to landslide have been identified, they have been temporally classified based on the timeseries. The mean trend in the displacement time series for each landslide active deformation area has been identified by the TSclassifier tool by using a velocity threshold of 2.5 mm/year to identify active points, a temporal window of 120 days, and acceleration and deceleration thresholds of 1.5 mm/year and 0.5 mm/year, respectively. Therefore, as shown in Fig. 6, the areas are classified into accelerating trends, decelerating trends, or not significant change.

Finally, each SU has been classified as landslide or non-landslide using two distinct criteria. In the first case, the classification is based on the NLD (Fig. 7a), where SUs with at least one recorded landslide are labelled as landslide SUs and the other ones as no-landslide. In the second case, the classification is performed using active landslide areas identified through the InSAR methodology developed in this study, following the criteria previously explained (Fig. 7b).

Once the SUs have been classified according to these criteria, the two sets of information have been integrated according to the matrix depicted in Fig. 8. This combination has resulted in an InSAR landslide inventory map (Fig. 8), which delineates the potential presence or absence of landslides within each SU according to the NLD and the InSAR data.

Particular attention should be given to areas where the NLD data and InSAR data are inconsistent, which might indicate the presence of unmapped landslides that are moving such as those in the North-West of England, Isle of Mull (Scotland), and southern part of Wales.

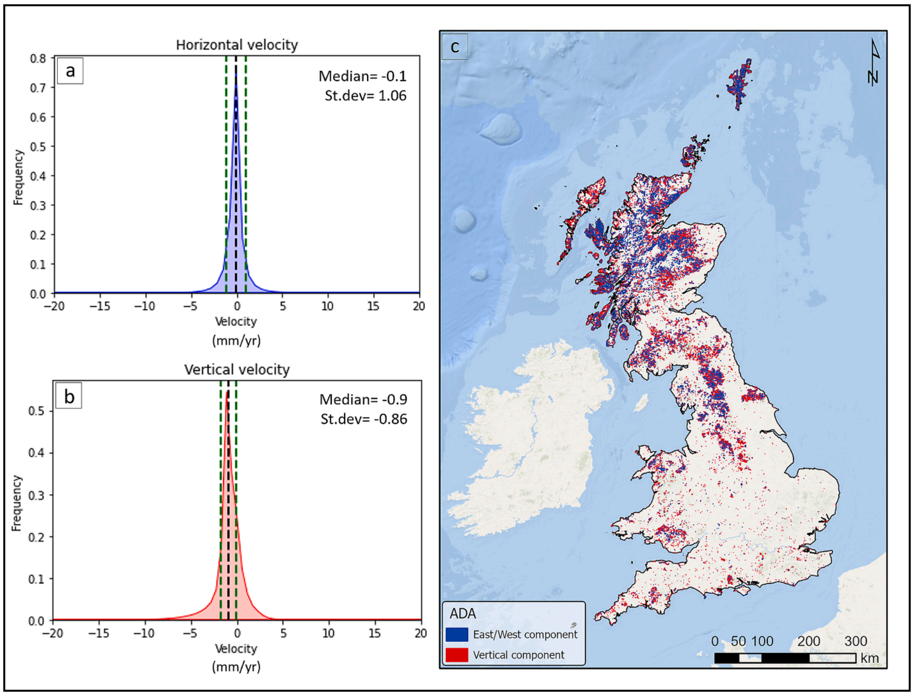


Fig. 5. a) Horizontal and b) vertical velocity distribution; c) ADAs for both components.

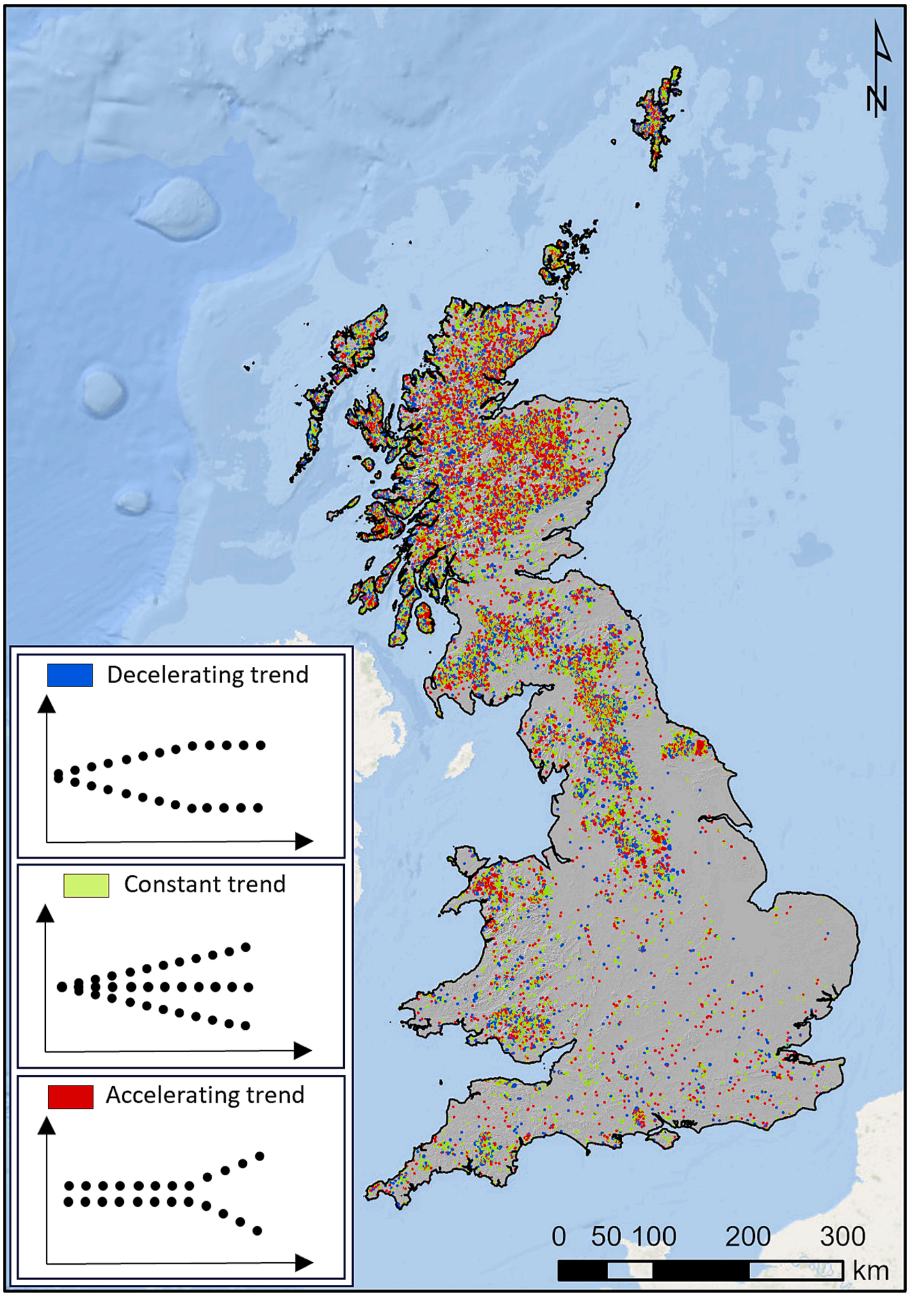


Fig. 6. ADAs classified into accelerating trends, decelerating trends, or not significant change.

In order to assess the performance of the applied methodology, given the national scope of the study and the lack of ground-based data for a deeper validation, we have considered well-known landslide areas, i.e., Blaina (Wales), Snake Pass (Derbyshire, England), and New Tredegar (Wales) (Fig. 9). This analysis has revealed that all three areas fall into category D of our matrix, indicating that both methods classify the area as a landslide zone.

The Isle of Wight, a southern island of the United Kingdom, and particularly its southern coastline with the town of Ventnor, is a well-known site for its slope instability (O'Connor et al., 2021). The landslide known as the Ventnor Undercliff Landslide Complex is recognised as the largest urban area affected by landsliding in Great Britain and has therefore been the subject of investigation for several decades (Chandler, 1984; Hourston et al., 2024; Moore et al., 2010) (Fig. 10). This complex is characterised by a large, deep-seated landslide, with displacement rates that are typically slow but have nonetheless caused significant damage. For these reasons, the Ventnor area has been chosen as the specific site to illustrate the various steps of the procedure proposed in this study (Fig. 11).

In Fig. 11a, the EGMS data for the Ventnor area are presented, specifically the vertical velocity component. Following this, the HDBSCAN algorithm has been applied to these data to remove outliers (black dots in Fig. 11b). The filtered data, for both geometries, has been then used as input for identifying active deformation areas. In this area, three ADAs have been identified from the East-West velocity data and two from the vertical velocity data (Fig. 11c). These ADAs have then been intersected with the SUs to provide geomorphological significance getting the subADAs (Fig. 11d). Subsequently, subADAs have been temporally classified in accelerating, decelerating or constant trend (Fig. 11e) and filtered to focus on those associated with landslide phenomena (Fig. 11f). Thereafter, SUs have been classified as landslide or non-landslide based on the two criteria: NLD (Fig. 11g) and ADA density

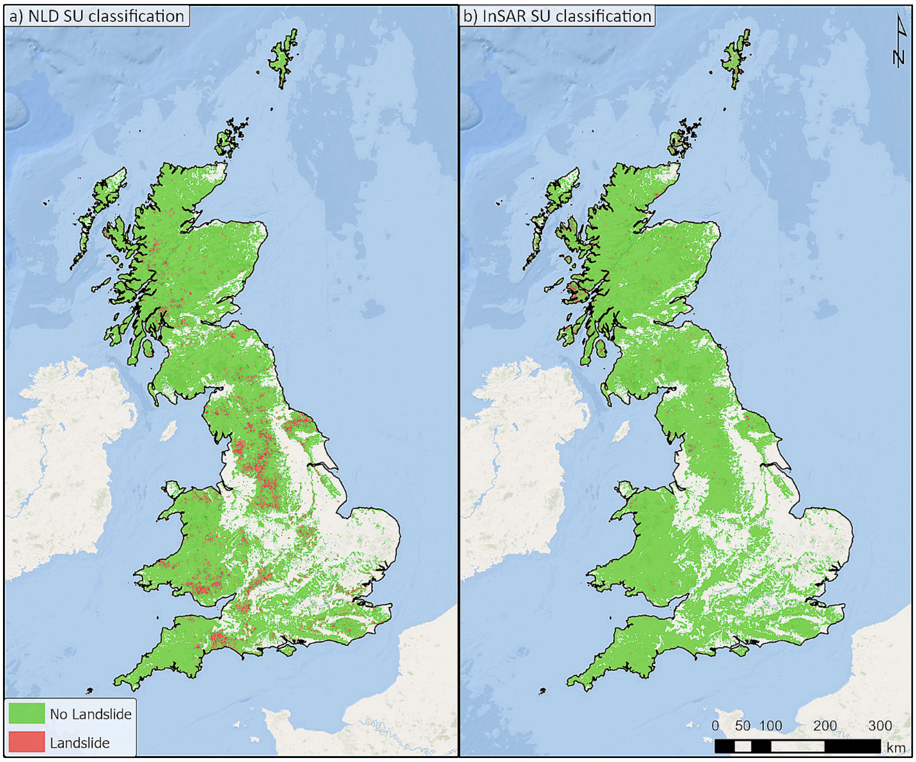


Fig. 7. a) SUs classification according to the National Landslide Database, where green indicates SUs without recorded landslides and red indicates SUs with at least one landslide. b) SUs classification based on landslides active areas identified using the InSAR methodology developed in this study.

(Fig. 11*h*). Ultimately, the matrix has been used to combine these two sets of data – ADA density and NLD – to create the final map (Fig. 11i). The resulting map in this area reveals the presence of all four classes: SUs where both methods are consistent (indicating either the presence or absence of a landslide) and SUs where the information is inconsistent, requiring further detailed analysis.

5. Discussion

The integration of MTInSAR, ML and clustering techniques significantly contributes to landslide detection. In this context, this study has shown that by leveraging the large volume of high-resolution data provided by the S1 satellite constellation, it is possible to identify areas characterised by the presence of movements associated with landslides on a national scale. The proposed approach allows for the efficient processing and analysis of massive datasets, which is crucial for applications that operate on country-wide basis.

Within this framework, most of the InSAR literature focuses on refining the boundaries of existing landslides (Hussain et al., 2023) or updating their state of activity (Ciuffi et al., 2024). In contrast, our approach represents a significant step forward, as it enables the detection of unstable slopes where slow movements (≤cm/yr) related to landslides are occurring, without relying on pre-existing inventories. This is valuable even in cases where high-resolution optical data fail to identify any sign of instability (Van Wyk De Vries et al., 2024).

Although the current methodology does not delineate landslide perimeters, it represents an important screening tool for identifying unstable slopes. Recent studies employ InSAR for mapping landslides within SUs (He et al., 2023), and our approach builds on this by introducing key innovations. These include filtering the InSAR data to remove isolated points before further analysis and considering the consistency between movement direction and slope aspect. Therefore, the proposed approach lays the foundation for subsequent detailed analyses. Indeed, the methodology is particularly helpful in identifying moving areas hotspots where resources need to be directed for assessing local scale conditions of the slope and eventually deploying landslide risk monitoring and/or mitigation activities. The proposed semi-automated approach facilitates the work of landslide specialists by being a practical solution for large-scale geohazard management.

Moreover, as mentioned, we have worked by SUs, a pragmatic approach to gather terrain deformation information that can be representative of slope conditions, even in situations where full InSAR coverage due to land cover and topography is absent. With this regard, we are also aware that the use of the velocity component involves an approximation due to the resampling data on a 100 m grid of the EGMS data. This analysis might also include large scale processes such as glacio-eustatic rebound, particularly evident in the northern part of GB even if with a limited magnitude (in the order of few mm per yr), as recorded by GNSS stations (Stockamp et al., 2015). However, this pattern cannot be excluded where its contribution is greater (e.g., Scandinavia).

Furthermore, the analysis of displacement trends by using the TSclassifier tool adds a valuable temporal dimension to the study. The classification of areas as accelerating, decelerating or no significant change trends makes it possible to prioritise interventions and monitor changes over time. This temporal characterisation, while insightful since it provides additional information for each unstable area, enabling a temporal characterization of the phenomenon, which can be very helpful for planning necessary actions, it represents a preliminary result that will need a deeper study.

Nevertheless, the proposed approach is not without its drawbacks. The reliance on InSAR data implies the limitations of this technique and, in particular, it means that the approach is effective for detecting only slow-moving landslides. The use of InSAR data also entails another limitation, as it causes the analysis to focus only on landslides which have been active during the InSAR revisiting period. Despite this limitation, the location of currently active landslides is of importance when assessing the potential exposure of transport routes, infrastructure and population centres.

While this approximation is suitable for a nationwide study, future

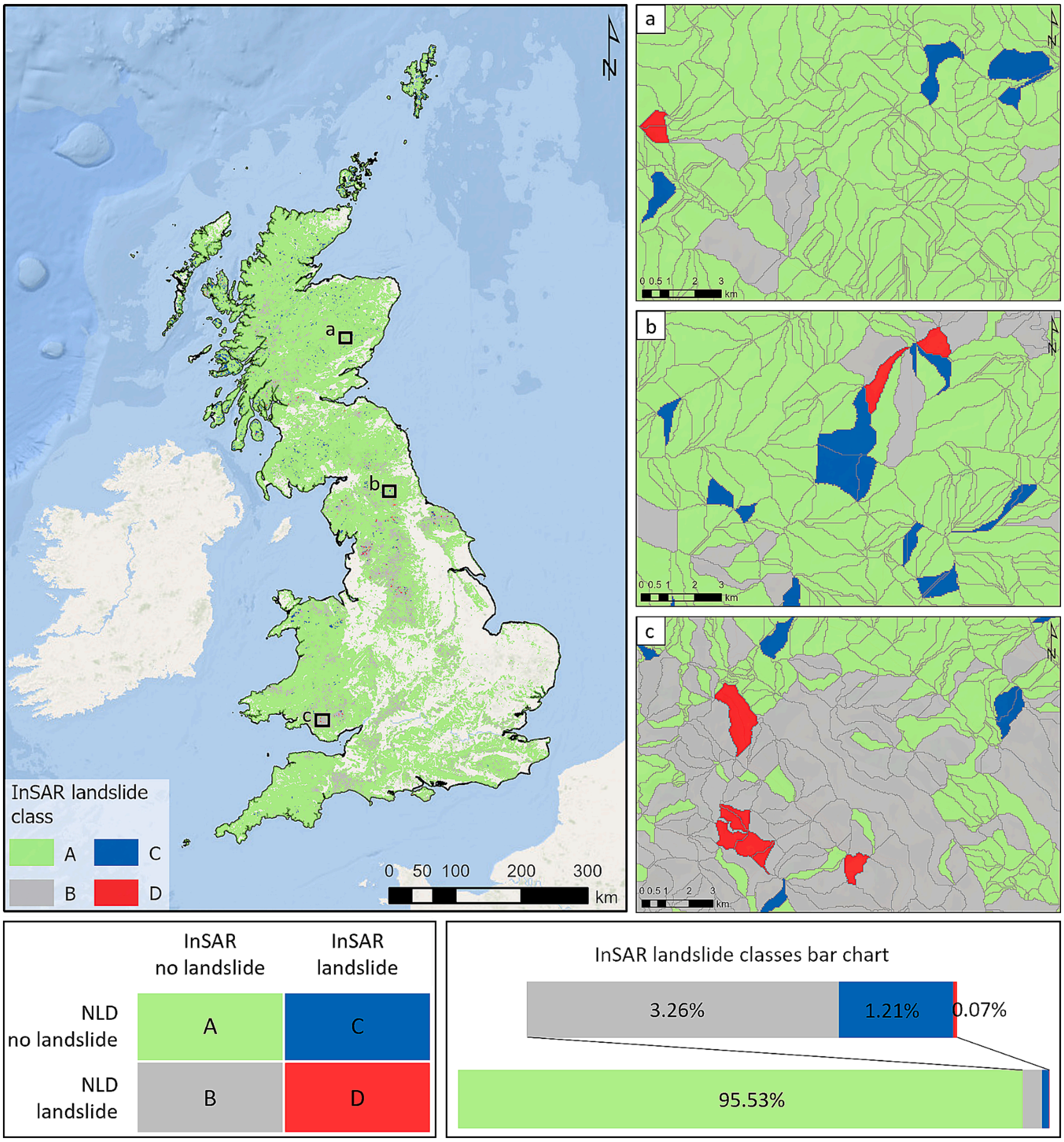


Fig. 8. InSAR landslide inventory map with associated matrix and the InSAR landslide classes bar chart. Additionally, three zooms of the map from Scotland (a), England (b) and Wales (c).

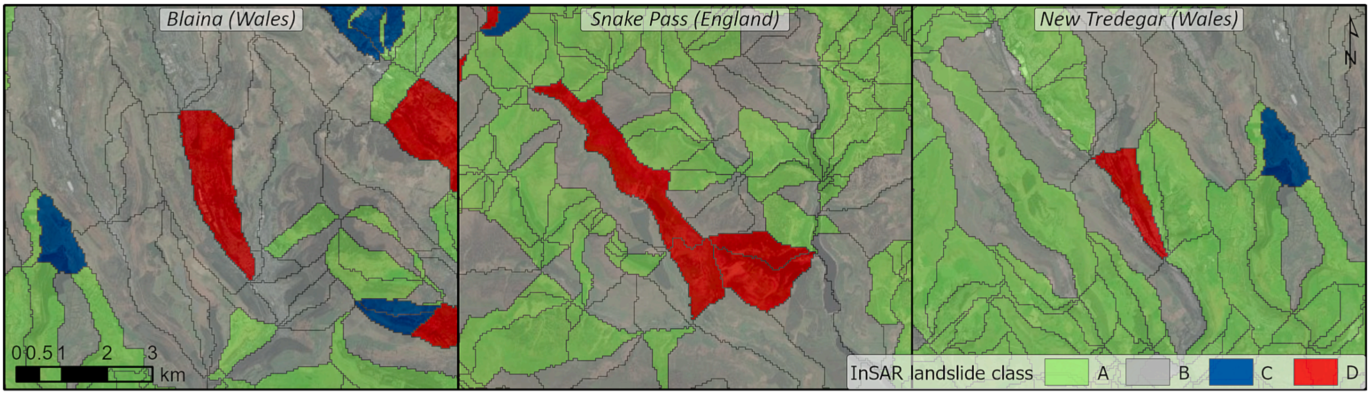


Fig. 9. Zoom of the InSAR landslide inventory map in Blaina (Wales), Snake Pass (Derbyshire, England), and New Tredegar (Wales).

developments will include the analysis of velocities along the satellite Line of Sight (LoS) to improve the accuracy of the results.

Additionally, the national scope of this study and the lack of ground-based data for large scale validation impose certain limitations in assessing the methodology and its accuracy. To address this issue and evaluate the approach, the analysis of results over well-known landslide

Fig. 10. Ventnor, Isle of Wight, case study; a) Toe of the Undercliff; b) Property damage; c) Ventnor Graben (http://geoscenic.bgs.ac.uk/).

sites has been conducted. This provides a preliminary validation but highlights the need for future ground-based surveys to fully verify and refine the methodology. However, given the time and financial constraints inherent to national-scale studies, a comprehensive validation across the entire country remains unfeasible.

The integration of ancillary data, like the land cover map, the DTM and DTM-derived products, into the methodology entails a geomorphological significance for the methodology, a fundamental aspect for landslide activity. To the same end, the SUs have been used as spatial unit in place of the commonly used pixel.

6. Conclusions

This study demonstrates the feasibility and effectiveness of analysing InSAR data with ML techniques for the detection of areas with associated landslide movement on a national scale. The approach provides a replicable model that can be applied to any area and any scale under the condition of data availability and is particularly useful for large scale analyses where ground truth data are inevitably scarce. The integration of EGMS data with a landslide inventory offers a comprehensive view of slow-movement landslide mapping and their state of activity, facilitating better decision-making in risk management. By leveraging the strengths of both remote sensing and semi-automatic techniques, it is possible to create detailed and timely maps of landslide-prone areas, thereby improving the ability to mitigate the impacts of this significant geohazard. Through the described methodology, a total of 336,557 SUs have been classified. Of these, the majority have been identified as non-landslide areas by both databases. Discrepancies between the landslide inventory and the InSAR data have been observed in 5 % of the units, and a small percentage has been classified as landslide areas by both criteria. Identifying areas classified as landslide-prone by both datasets or where the datasets are inconsistent, is crucial for risk management as it highlights regions requiring more detailed inspections. However aware of the weaknesses of this approach, further efforts will be required in the future to increase the robustness of the methodology. An example of future development could be the employment of geomorphometric parameters (e.g. topological data analysis, Bhuyan et al., 2023) to enable the delineation of landslide perimeters rather than merely identifying unstable slopes, also providing a more comprehensive understanding of landslide dynamics. In particular, an advanced step will involve filtering the data through the HDBSCAN algorithm based on the time series rather than using just the average velocity, which cannot fully represent the time series displacement, especially with seasonal movements. Furthermore, the temporal analysis performed in this study represents a preliminary result, and recognising the criticality of this aspect, future perspectives consist of improving the temporal analysis. This will include the ability to detect seasonality in time series data, make more nuanced distinctions within accelerating or decelerating trends, and identify the time interval when a trend change occurred. The next step will also be to consider displacement from full resolution data to improve the accuracy of the results by reducing the spatial approximation. The on-the-ground validation of results is also a crucial aspect that will be taken into consideration for a quantitative assessment of our approach, as field surveys and local observations can provide critical data for verifying and calibrating results and also for landslide risk management and mitigation.

CRediT authorship contribution statement

Camilla Medici: Writing – original draft, Software, Methodology, Conceptualization. **Alessandro Novellino:** Writing – review & editing, Methodology, Conceptualization. **Claire Dashwood:** Writing – review & editing, Validation. **Silvia Bianchini:** Writing – review & editing, Methodology, Conceptualization.

Funding

The research was supported by the ALMEO project, part of the BGS ‘Geoscience to tackle Global Environmental Challenges’ grant (NE/X006255/1).

Declaration of competing interest

The authors declare that they have no known competing financial

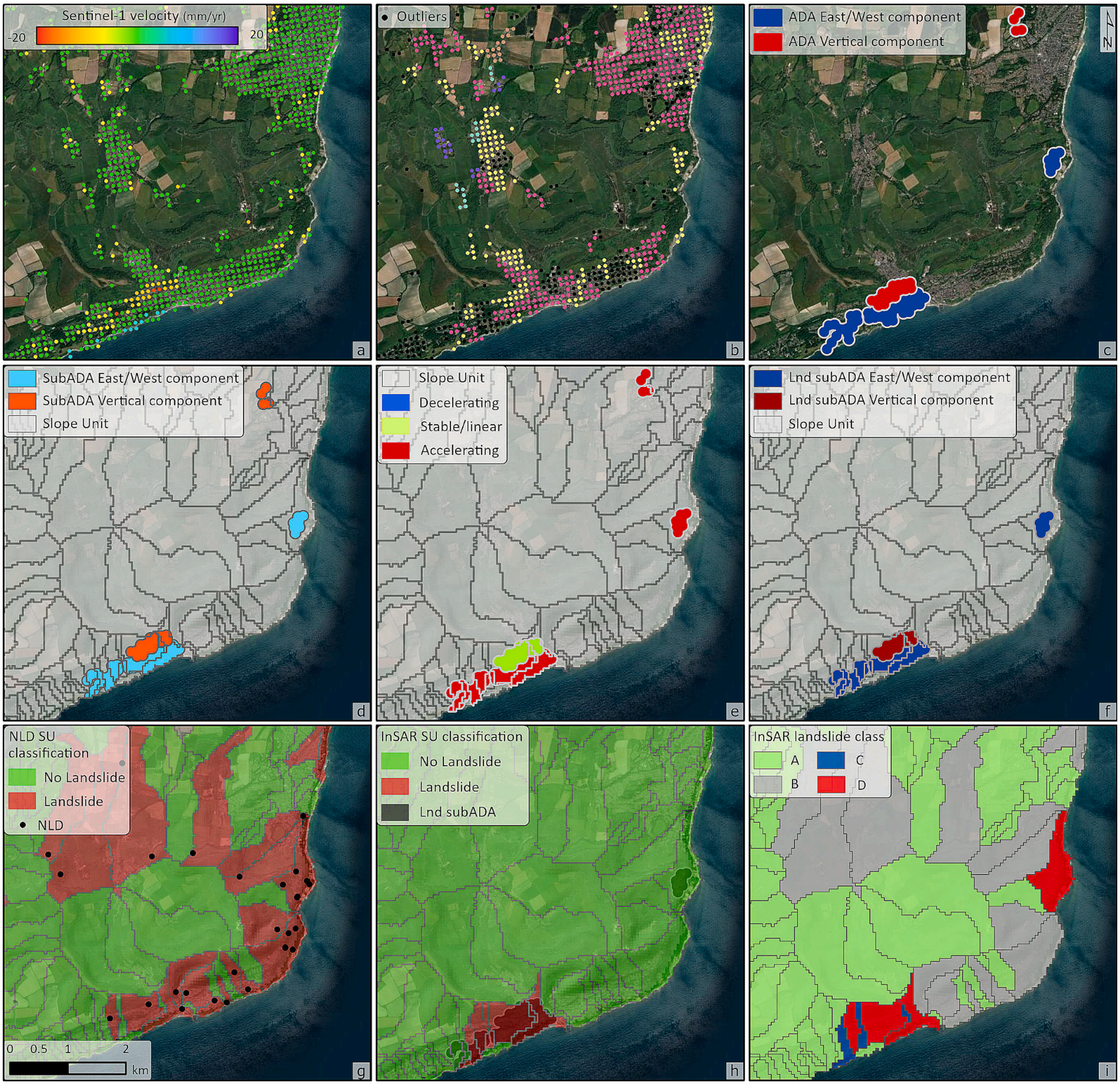


Fig. 11. Step-by-step procedure in the Ventnor area. a) S1 vertical velocity from EGMS data; b) HDBSCAN result; c) ADAs for both velocity components; d) subADAs from the intersection between ADAs and SUs; e) subADAs temporal classification; f) filtering of the subADAs to get only those related to landslide; g) SUs classification according to the NLD; h) SUs classification according to the ADA density; i) InSAR landslide inventory map.

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the UK Centre for Ecology & Hydrology for providing the land cover map (Order Number: NB7M549R). The authors also acknowledge the crucial contribution of the European project EGMS RASTOOL*: European ground motion risk assessment tool* co-financed by the EU-Union Civil Protection Mechanism under grant agreement n° 101048474 (https://rastool.cttc.es/).

Data availability

Data will be made available on request.

References

Alshammari, L., Boyd, D.S., Sowter, A., Marshall, C., Andersen, R., Gilbert, P., Marsh, S., Large, D.J., 2020. Use of surface motion characteristics determined by InSAR to assess peatland condition. JGR Biogeosciences 125, e2018JG004953. https://doi.org/10.1029/2018JG004953.

Alvioli, M., Marchesini, I., Reichenbach, P., Rossi, M., Ardizzone, F., Fiorucci, F., Guzzetti, F., 2016. Automatic delineation of geomorphological slope units with r.slopeunits v1.0 and their optimization for landslide susceptibility modeling. Geoscientific Model Development 9, 3975–3991. https://doi.org/10.5194/gmd-9-3975-2016.

Barra, A., Solari, L., Béjar-Pizarro, M., Monserrat, O., Bianchini, S., Herrera, G., Crosetto, M., Sarro, R., González-Alonso, E., Mateos, R., Ligüerzana, S., López, C., Moretti, S., 2017. A methodology to detect and update active deformation areas based on sentinel-1 SAR images. Remote Sensing 9, 1002. https://doi.org/10.3390/rs9101002.

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. Scientific Reports 13, 162.

Bianchini, S., Solari, L., Casagli, N., 2017. A GIS-based procedure for landslide intensity evaluation and specific risk analysis supported by persistent scatterers interferometry (PSI). Remote Sensing 9, 1093.

Bovenga, F., Wasowski, J., Nitti, D.O., Nutricato, R., Chiaradia, M.T., 2012. Using COSMO/SkyMed X-band and ENVISAT C-band SAR interferometry for landslides analysis. Remote Sensing of Environment 119, 272–285.

Bradley, A.V., Andersen, R., Marshall, C., Sowter, A., Large, D.J., 2021. Identification of typical eco-hydrological behaviours using InSAR allows landscape-scale mapping of peatland condition. Earth Surface Dynamics Discussions 2021, 1–28.

Bryce, E., Castro-Camilo, D., Dashwood, C., Tanyas, H., Ciurean, R., Novellino, A., Lombardo, L., 2025. An updated landslide susceptibility model and a log-Gaussian Cox process extension for Scotland. Landslides 22 (2), 517–535.

Cascini, L., Fornaro, G., Peduto, D., 2010. Advanced low-and full-resolution DInSAR map generation for slow-moving landslide analysis at different scales. Engineering Geology 112, 29–42.

Chandler, M.P., 1984. The coastal landslides forming the Undercliff of the Isle of Wight (PhD Thesis). University of London.

Chang, Z., Catani, F., Huang, F., Liu, G., Meena, S.R., Huang, J., Zhou, C., 2023. Landslide susceptibility prediction using slope unit-based machine learning models considering the heterogeneity of conditioning factors. Journal of Rock Mechanics and Geotechnical Engineering 15, 1127–1143.

Cigna, F., Bateson, L.B., Jordan, C.J., Dashwood, C., 2014. Simulating SAR geometric distortions and predicting Persistent Scatterer densities for ERS-1/2 and ENVISAT C-band SAR and InSAR applications: Nationwide feasibility assessment to monitor the landmass of Great Britain with SAR imagery. Remote Sensing of Environment 152, 441–466. https://doi.org/10.1016/j.rse.2014.06.025.

Cigna, F., Bianchini, S., Casagli, N., 2013. How to assess landslide activity and intensity with Persistent Scatterer Interferometry (PSI): the PSI-based matrix approach. Landslides 10, 267–283. https://doi.org/10.1007/s10346-012-0335-7.

Ciuffi, P., Bayer, B., Berti, M., Franceschini, S., Simoni, A., 2024. InSAR stacking to detect active landslides and investigate their relation to rainfalls in the Northern Apennines of Italy. Geomorphology 457, 109242.

Colesanti, C., Ferretti, A., Novali, F., Prati, C., Rocca, F., 2003. SAR monitoring of progressive and seasonal ground deformation using the permanent scatterers technique. IEEE Transactions on Geoscience and Remote Sensing 41, 1685–1701.

Confuorto, P., Casagli, N., Casu, F., De Luca, C., Del Soldato, M., Festa, D., Lanari, R., Manzo, M., Onorato, G., Raspini, F., 2023. Sentinel-1 P-SBAS data for the update of the state of activity of national landslide inventory maps. Landslides 20, 1083–1097. https://doi.org/10.1007/s10346-022-02024-0.

Corominas, J., Van Westen, C., Frattini, P., Cascini, L., Malet, J.-P., Fotopoulou, S., Catani, F., Van Den Eeckhaut, M., Mavrouli, O., Agliardi, F., Pitilakis, K., Winter, M. G., Pastor, M., Ferlisi, S., Tofani, V., Hervás, J., Smith, J.T., 2013. Recommendations for the quantitative analysis of landslide risk. Bulletin of Engineering Geology and the Environment. https://doi.org/10.1007/s10064-013-0538-8.

Costantini, M., Minati, F., Trillo, F., Ferretti, A., Novali, F., Passera, E., Dehls, J., Larsen, Y., Marinkovic, P., Eineder, M., Brcic, R., Siegmund, R., Kotzerke, P., Probeck, M., Kenyeres, A., Proietti, S., Solari, L., Andersen, H.S., 2021. European Ground Motion Service (EGMS), in: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Presented at the IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Brussels, Belgium, pp. 3293–3296. https://doi.org/10.1109/IGARSS47720.2021.9553562.

Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., Crippa, B., 2016. Persistent Scatterer Interferometry: A review. ISPRS Journal of Photogrammetry and Remote Sensing 115, 78–89. https://doi.org/10.1016/j.isprsjprs.2015.10.011.

Crosetto, M., Solari, L., Balasis-Levinsen, J., Bateson, L., Casagli, N., Frei, M., Oyen, A., Moldestad, D.A., Mróz, M., 2021. Deformation Monitoring AT European Scale: The Copernicus Ground Motion Service. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLIII-B3-2021, 141–146. https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-141-2021.

Crosetto, M., Solari, L., Mróz, M., Balasis-Levinsen, J., Casagli, N., Frei, M., Oyen, A., Moldestad, D.A., Bateson, L., Guerrieri, L., Comerci, V., Andersen, H.S., 2020. The evolution of wide-area DInSAR: From regional and national services to the European ground motion service. Remote Sensing 12, 2043. https://doi.org/10.3390/rs12122043.

Dashwood, C., Pennington, C., Bee, E., Freeborough, K., Dijkstra, T., 2017. Creation of a National Landslide Domain Map to Aid Susceptibility Mapping in Great Britain. In: Mikos, M., Tiwari, B., Yin, Y., Sassa, K. (Eds.), Advancing Culture of Living with Landslides. Springer International Publishing, Cham, pp. 973–979. https://doi.org/10.1007/978-3-319-53498-5_110.

Del Soldato, M., Solari, L., Novellino, A., Monserrat, O., Raspini, F., 2021. A new set of tools for the generation of InSAR visibility maps over wide areas. Geosciences 11, 229. https://doi.org/10.3390/geosciences11060229.

Delforge, D., Wathelet, V., Below, R., Sofia, C.L., Tonnelier, M., Loenhout, J.V., Speybroeck, N., 2023. EM-DAT: the Emergency Events Database. https://doi.org/10.21203/rs.3.rs-3807553/v1.

Farina, P., Colombo, D., Fumagalli, A., Marks, F., Moretti, S., 2006. Permanent Scatterers for landslide investigations: Outcomes from the ESA-SLAM project. Engineering Geology 88, 200–217.

Festa, D., Bonano, M., Casagli, N., Confuorto, P., De Luca, C., Del Soldato, M., Lanari, R., Lu, P., Manunta, M., Manzo, M., 2022. Nation-wide mapping and classification of ground deformation phenomena through the spatial clustering of P-SBAS InSAR measurements: Italy case study. ISPRS Journal of Photogrammetry and Remote Sensing 189, 1–22.

Festa, D., Del Soldato, M., 2023. EGMStream, a desktop app for EGMS data downstream. Remote Sensing 15, 2581. https://doi.org/10.3390/rs15102581.

Foster, C., Pennington, C.V.L., Culshaw, M.G., Lawrie, K., 2012. The national landslide database of Great Britain: Development, evolution and applications. Environment and Earth Science 66, 941–953. https://doi.org/10.1007/s12665-011-1304-5.

Gibson, A.D., Culshaw, M.G., Dashwood, C., Pennington, C.V.L., 2013. Landslide management in the UK—the problem of managing hazards in a 'low-risk' environment. Landslides 10, 599–610. https://doi.org/10.1007/s10346-012-0346-4.

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-Science Reviews 112, 42–66. https://doi.org/10.1016/j.earscirev.2012.02.001.

Han, J., Guo, X., Jiao, R., Nan, Y., Yang, H., Ni, X., Zhao, D., Wang, S., Ma, X., Yan, C., 2023. An automatic method for delimiting deformation area in insar based on hnsw-dbscan clustering algorithm. Remote Sensing 15, 4287.

Haque, U., Blum, P., Da Silva, P.F., Andersen, P., Pilz, J., Chalov, S.R., Malet, J.-P., Auflič, M.J., Andres, N., Poyiadji, E., Lamas, P.C., Zhang, W., Peshevski, I., Pétursson, H.G., Kurt, T., Dobrev, N., García-Davalillo, J.C., Halkia, M., Ferri, S., Gaprindashvili, G., Engström, J., Keellings, D., 2016. Fatal landslides in Europe. Landslides 13, 1545–1554. https://doi.org/10.1007/s10346-016-0689-3.

He, K., Tanyas, H., Chang, L., Hu, X., Luo, G., Lombardo, L., 2023. Modelling InSAR-derived hillslope velocities with multivariate statistics: A first attempt to generate interpretable predictions. Remote Sensing of Environment 289, 113518.

Hourston, H.I., Gonzalez Alvarez, I., Bateson, L., Hussain, E., Novellino, A., 2024. Automated INSAR time-series analysis tool for geological interpretations in near-real time.

Huang, J., Xie, M., Farooq, A., Williams, E.J., 2019. DInSAR technique for slow-moving landslide monitoring based on slope units. Survey Review 51, 70–77. https://doi.org/10.1080/00396265.2017.1380947.

Hussain, S., Pan, B., Afzal, Z., Ali, M., Zhang, X., Shi, X., Ali, M., 2023. Landslide detection and inventory updating using the time-series InSAR approach along the Karakoram Highway. Northern Pakistan. Scientific Reports 13, 7485.

Intrieri, E., Raspini, F., Fumagalli, A., Lu, P., Del Conte, S., Farina, P., Allievi, J., Ferretti, A., Casagli, N., 2018. The Maoxian landslide as seen from space: detecting precursors of failure with Sentinel-1 data. Landslides 15, 123–133. https://doi.org/10.1007/s10346-017-0915-7.

Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., Huser, R., 2020. Space-Time Landslide Predictive Modelling. Earth-Science Reviews 209, 103318.

McInnes, L., Healy, J., Astels, S., 2017. hdbscan: Hierarchical density based clustering. JOSS 2, 205. https://doi.org/10.21105/joss.00205.

Montalti, R., Solari, L., Bianchini, S., Del Soldato, M., Raspini, F., Casagli, N., 2019. A Sentinel-1-based clustering analysis for geo-hazards mitigation at regional scale: A case study in Central Italy. Geomatics, Natural Hazards and Risk 10, 2257–2275. https://doi.org/10.1080/19475705.2019.1690058.

Moore, R., Carey, J.M., McInnes, R.G., 2010. Landslide behaviour and climate change: Predictable consequences for the Ventnor Undercliff, Isle of Wight. QJEGH 43, 447–460. https://doi.org/10.1144/1470-9236/08-086.

Nava, L., Bhuyan, K., Meena, S.R., Monserrat, O., Catani, F., 2022. Rapid mapping of landslides on SAR data by attention U-Net. Remote Sensing 14, 1449.

Novellino, A., Cigna, F., Brahmi, M., Sowter, A., Bateson, L., Marsh, S., 2017. Assessing the feasibility of a national InSAR ground deformation map of great Britain with sentinel-1. Geosciences 7, 19. https://doi.org/10.3390/geosciences7020019.

Novellino, A., Pennington, C., Leeming, K., Taylor, S., Alvarez, I.G., McAllister, E., Arnhardt, C., Winson, A., 2024. Mapping landslides from space: A review. Landslides 21, 1041–1052. https://doi.org/10.1007/s10346-024-02215-x.

O'Connor, W., Mider, G., Lawrence, J.A., Agar, S., Mason, P.J., Ghail, R., Scoular, J., 2021. An Investigation into ground movement on the ventnor landslide complex, UK using persistent scatterer interferometry. Remote Sensing 13, 3711. https://doi.org/10.3390/rs13183711.

Pennington, C., Freeborough, K., Dashwood, C., Dijkstra, T., Lawrie, K., 2015. The national landslide database of Great Britain: Acquisition, communication and the role of social media. Geomorphology 249, 44–51. https://doi.org/10.1016/j.geomorph.2015.03.013.

Prakash, N., Manconi, A., Loew, S., 2020. Mapping landslides on EO data: Performance of deep learning models vs. traditional machine learning models. Remote Sensing 12, 346.

Raspini, F., Bianchini, S., Ciampalini, A., Del Soldato, M., Solari, L., Novali, F., Del Conte, S., Rucci, A., Ferretti, A., Casagli, N., 2018. Continuous, semi-automatic monitoring of ground deformation using Sentinel-1 satellites. Scientific Reports 8, 7253. https://doi.org/10.1038/s41598-018-25369-w.

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., Guzzetti, F., 2018. A review of statistically-based landslide susceptibility models. Earth-Science Reviews 180, 60–91.

Righini, G., Pancioli, V., Casagli, N., 2012. Updating landslide inventory maps using Persistent Scatterer Interferometry (PSI). International Journal of Remote Sensing 33, 2068–2096. https://doi.org/10.1080/01431161.2011.605087.

Rosi, A., Tofani, V., Tanteri, L., Tacconi Stefanelli, C., Agostini, A., Catani, F., Casagli, N., 2018. The new landslide inventory of Tuscany (Italy) updated with PS-InSAR: Geomorphological features and landslide distribution. Landslides 15, 5–19. https://doi.org/10.1007/s10346-017-0861-4.

Rygus, M., Novellino, A., Hussain, E., Syafiudin, F., Andreas, H., Meisina, C., 2023. A clustering approach for the analysis of InSAR time series: Application to the bandung basin (Indonesia). Remote Sensing 15, 3776. https://doi.org/10.3390/rs15153776.

Showstack, R., 2014. Sentinel satellites initiate new era in earth observation. EoS Transactions 95, 239–240. https://doi.org/10.1002/2014EO260003.

Solari, L., Del Soldato, M., Raspini, F., Barra, A., Bianchini, S., Confuorto, P., Casagli, N., Crosetto, M., 2020. Review of satellite interferometry for landslide detection in Italy. Remote Sensing 12, 1351. https://doi.org/10.3390/rs12081351.

Strozzi, T., Ambrosi, C., Raetzo, H., 2013. Interpretation of aerial photographs and satellite SAR interferometry for the inventory of landslides. Remote Sensing 5, 2554–2570.

Tehrani, F.S., Calvello, M., Liu, Z., Zhang, L., Lacasse, S., 2022. Machine learning and landslide studies: Recent advances and applications. Natural Hazards 114, 1197–1245. https://doi.org/10.1007/s11069-022-05423-7.

Tofani, V., Raspini, F., Catani, F., Casagli, N., 2013. Persistent Scatterer Interferometry (PSI) technique for landslide characterization and monitoring. Remote Sensing 5, 1045–1065.

Tomás, R., Pagán, J.I., Navarro, J.A., Cano, M., Pastor, J.L., Riquelme, A., Cuevas-González, M., Crosetto, M., Barra, A., Monserrat, O., Lopez-Sanchez, J.M., Ramón, A., Ivorra, S., Del Soldato, M., Solari, L., Bianchini, S., Raspini, F., Novali, F., Ferretti, A., Costantini, M., Trillo, F., Herrera, G., Casagli, N., 2019. Semi-Automatic identification and pre-screening of geological–geotechnical deformational processes using persistent scatterer interferometry datasets. Remote Sensing 11, 1675. https://doi.org/10.3390/rs11141675.

Van Wyk De Vries, M., Arrell, K., Basyal, G.K., Densmore, A.L., Dunant, A., Harvey, E.L., Jimee, G.K., Kincey, M.E., Li, S., Singh Pujara, D., Shrestha, R., Rosser, N.J., Dadson, S.J., 2024. Detection of slow-moving landslides through automated monitoring of surface deformation using Sentinel-2 satellite imagery. Earth Surf Processes Landf 49, 1397–1410. https://doi.org/10.1002/esp.5775.