Landslides (2026) 23:969–988
DOI 10.1007/s10346-025-02671-z
Received: 9 April 2025
Accepted: 27 November 2025
Published online: 7 January 2026

Matteo Berti · Marco Pizziolo · Michele Scaroni · Mauro Generali · Stefano Olivucci · Giampiero Gozza · Paolo Formicola · Vincenzo Critelli · Marco Mulas · Melissa Tondo · Francesco Lelli · Cecilia Fabbiani · Francesco Ronchetti · Giuseppe Ciccarese · Nicola Dal Seno · Elena Ioriatti · Rodolfo Rani · Alessandro Zuccarini · Tommaso Simonelli · Alessandro Corsini

Testing NDVI and U-Net for automated mapping of multiple-occurrence regional landslide events using satellite and aerial multispectral data (Casola Valsenio, Emilia-Romagna, Northern Apennines, Italy)

Abstract In May 2023, the Emilia-Romagna Region (northern Italy) experienced a severe multiple occurrence regional landslide event, resulting in more than 80,000 debris slides and flows as well as some hundreds of rock-block slides. One of the more severely affected areas was the Municipality of Casola Valsenio, with more than 5200 landslides in 84 km². In this study, we assess the performances of two established methods for automated landslide mapping (NDVI change and U-Net) applied to satellite and aerial multispectral data of resolutions ranging from 10 to 0.2 m (the latter resampled to 1 and 2 m). In doing so, we used a basic configuration (i.e., no additional data layer in the analysis) to simulate their application in the aftermath of a MORLE scenario, when time and training data are limited. The results show that, even with limited training, U-Net outperforms NDVI in terms of accuracy only in case of 2 m resolution aerial data (achieving an F1-score higher than 0.6), while with 10 m resolution satellite data, the NDVI performs better (F1 up to 0.5) than U-Net (F1 lower than 0.4). The enhanced performance of U-Net with higher resolution data is ascribed to its proficiency in detecting landslides not only by changes in vegetation but also by analyzing the landslide's physical shape. However, the research ultimately evidence that in the case study area, automated mapping using NDVI change and U-Net in basic configuration still is affected by too many false positives and misses, so it cannot eliminate the need for manual verification, especially for precise identification of affected buildings and roads.

Keywords Automated mapping · NDVI · U-Net

Introduction

In May 2023, the Emilia-Romagna Region (northern Italy) has experienced one of the most severe multiple occurrence regional landslide event (Crozier 2005) (Fig. 1). The inventory of the effects on slopes of the May 2023 MORLE in Emilia-Romagna has been recently published by Berti et al. (2025). It includes approximately 81,000 landslides, classified as debris slides (DS), debris flows (DF), rock-block slides (RS) (in the coarse-grained bedrock units), and earth slides (ES), earth flows (EF) (in the fine-grained bedrock units), which have been manually mapped by geologists on the basis of high-resolution multispectral aerial images. The event has been caused by the subsequent occurrence of two major rainfall events: May 1–3 2023, which delivered approximately 200 mm of rain in 48 h, and May 16–17, which resulted in rainfall reaching 200–250 mm in 48 h. The estimated return period of the two combined events exceeds 500 years (Brath et al. 2023).

One of the areas more severely affected by the May 2023 event is the Municipality of Casola Valsenio. In 84 km², a total of 5507 landslides have been triggered, damaging roads and endangering buildings (Berti et al. 2025). With reference to this area, that is representative of the geological setting of the sector of the Emilia-Romagna region affected by the disaster, and on the basis of 4-bands (RGB + NIR) satellite and aerial imageries at different resolution, this study aims to (i) test the performance of automated landslides mapping techniques based either on NDVI changes (Normalized Difference Vegetation Index; Tucker 1979) or U-Net convolutional neural network (Ronneberger et al. 2015); (ii) assess the extent to which false positives and misses resulting from automated mapping can affect the inventory of buildings and roads impacted by landslides.

Background and rationale

The use of remote sensing data is crucial for landslides inventories after major hydrogeological disasters (Guzzetti et al. 2012; Behling et al. 2014; Tanoli et al. 2017; Gorelick et al. 2017; Ghorbanzadeh et al. 2020; Ferrario and Livio 2023). In Europe, the Copernicus Emergency Management Service (CEMS) makes use mid- to high-resolution pre- and post-event satellite imagery (such as Sentinel-2 MSI at 10 m resolution in visible and NIR and Pléiades-1A/B at 0.5 m resolution) to provide rapid mapping of landslides and flooded areas using both visual and automatic methods (Joubert-Boitat et al. 2020). Similarly, in the USA, services of Rapid Response Landslide Inventory are provided by the USGS (Martinez et al. 2021) using visual interpretation of mid- to high-resolution pre- and post-event satellite imagery including Sentinel-2, WorldView (0.3–0.5-m resolution), and Planet (2.7–4.0-m). ALADIM-VHR is a service created by CNRS-EOST in Strasbourg, France, designed to map landslides triggered by significant events. It utilizes high-resolution or

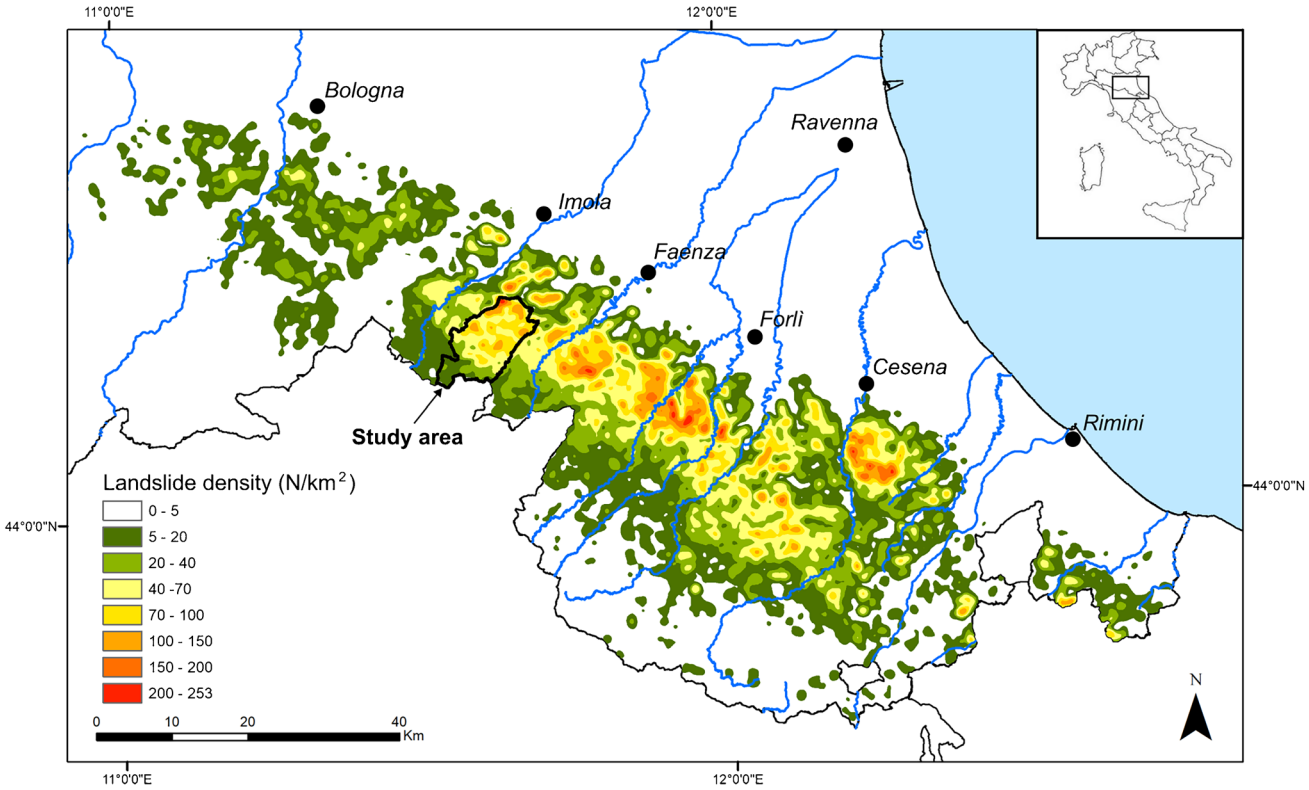


Fig. 1 Overview of the May 2023 disaster in the Emilia-Romagna region (from Berti et al. 2025, modified). Colors indicate the landslide density (number of landslide per km^2) determined through manual mapping

very-high-resolution satellite imagery (typically Pléiades, Spot6/7, Geo-Eye and Planets) and employs change detection and machine learning techniques for automated mapping (Stumppf and Kerle 2010). Recently, Amatya et al. (2021) introduced a novel landslide mapping system named SALaD (NASA's Semi-Automatic Landslide Detection), which integrates object-based image analysis (OBIA) and machine learning (ML) into open-source Python modules. This system is designed to quickly map landslides across large areas after extreme events.

The increasing availability of free services and open-access tools for automated landslide detection highlights the vibrant nature of this research field, driven by recent advancements in artificial intelligence (Prakash et al. 2020, 2021; Tehrani et al. 2022; Meena et al. 2022; Chen et al. 2023; Khalili et al. 2024). Deep learning (DL) algorithms, particularly those based on convolutional neural networks, have emerged as the predominant methods due to their outstanding pattern recognition and image segmentation capabilities (Su et al. 2021; Minaee et al. 2022). Over the past decade, a substantial body of research has been published, featuring numerous variations of the U-Net architecture originally proposed by Ronneberger et al. (2015). Notable adaptations include ResU-Net (Ghorbanzadeh et al. 2022), Attention U-Net (Nava et al. 2022), Dual-encoder U-Net (Lu et al. 2023), and U-Net with a ResNet Backbone (Chandra et al. 2023). Other recent studies have focused on enhancing traditional methods for automated mapping based on change detection indexes (Scheip and Wegmann 2021; Milledge et al. 2022; Notti et al. 2023), pixel-based approaches (Nichol and Wong 2005; Mondini et al. 2017; Lu et al. 2019), or object-oriented techniques (Holbling et al. 2017; Comert et al. 2019; Amatya et al. 2021), or integrating these methods with DL techniques (Shahabi et al. 2022).

Despite these advancements, substantial challenges persist. In a recent literature review, He et al. (2024) noted that the application of DL models in landslide detection is "still in its infancy." They pointed out that both the design of DL architectures and the enhancement of model performance and generalization capabilities require more in-depth research. Novellino et al. (2024) also emphasized the need for further studies to improve the operational effectiveness of these tools, including the establishment of standardized training protocols for DL models. These challenges are particularly crucial in the urgent context of landslide mapping. Automated methods are generally trained and tested using databases of landslides from past rainfall or seismic events, often with a training-to-test ratio of 70:30 or more. However, few studies explore the application of these methods with very limited training data available, which is a common situation in real-time emergency scenarios, which demand quick and precise mapping. Furthermore, while advanced deep learning (DL) methods benefit from high-resolution imagery—since greater spatial resolution allows for finer detail detection (Bhuyan et al. 2023)—this advantage comes with considerable practical limitations, including increased computational load, higher memory demands, and a heightened risk of overfitting, particularly when the training dataset is small (Wang et al. 2024). Although these challenges are manageable in a research setting, they can become substantial obstacles in practical

emergency applications. Indeed, the issues identified by Williams et al. (2018) regarding the difficulties of rapid landslide mapping under data and time constraints are still pertinent.

In this study, we assess two established methods for automated landslide mapping (NDVI change and U-Net) by simulating their application in a MORLE scenario. Our aim is to determine whether automated methods can truly reduce mapping efforts, even when time and calibration data are limited.

Methods

Landslides and damages dataset used for training and testing

The Casola Valsenio municipality extends for 84 km² in the eastern flank of the northern Apennines, in Italy, and it ranges in elevation from 770 to 100 m above sea level (Fig. 2). Hillslopes are predominantly made of turbidite rock masses of the Marnoso-Arenacea Formation (FMA). The structural setting is that of a monocline that results in steep 30–50° slopes with FMA dipping uphill and gentle 15–20° slopes with FMA dipping downhill.

Prior to the May 2023 MORLE, the Italian National Landslides Inventory (IFFI; Trigila et al. 2008) included 730 landslides in Casola Valsenio, the majority being dormant rock-block slides (Fig. 2a). In the May 2023, the municipality has been affected by a total of 5507 landslides, almost all of which were first-time failures (Berti et al. 2025) (Fig. 2b). Debris flows (DF) and debris slides (DS) have widely affected steep forested slopes, leading to clearance of vegetation in the source and track areas and deposited bare debris along or at the bottom of slopes. At the same time, rock-block slides (RS) have affected gentle slopes, determining clearly recognizable detachment scarps and displaced rock blocks and boulders. Altogether, in Casola Valsenio, the May 2023 MORLE landslides inventory of Berti et al (2025) includes 4224 DS (1.73 km²), 1076 DF (1.81 km²), and 111 RS (0.63 km²). Based on the size classification suggested by McColl and Cook (2024), debris slides are generally small with an average area of about 400 m², whereas debris flows and rock slides are considered medium-sized, averaging about 1700 m² and 5700 m², respectively. These landslides affected 64 houses and 185 roads, considering structures inside or within 20 m of the landslides. These elements were manually identified using aerial imagery with a high-resolution of 0.2 m per pixel, taken 1 week after the event. The data will be used to evaluate the accuracy of automated methods in detecting damage, which requires precise identification of landslide boundaries.

Satellite and aerial multispectral datasets

Table 1 provides details of the original satellite and aerial multispectral datasets (4-bands, i.e., RGB + NIR) used in this study as reference for pre- and post- May 2023 MORLE conditions.

Satellite datasets: (i) pre-event dataset S1 includes 10 m resolution Sentinel-2 images from May and June 2022; images just before the event (April 2023) were unsuitable due to persistent cloud cover, and images of March 2023 have been discarded as early spring is marked by rapid vegetation growth and intense agricultural activities. (ii) Pre-event dataset S2 includes a single 10 m resolution Sentinel-2 image captured in May 2022, exactly 1 year before the May 2023 event. (iii) Post-event dataset S3, includes the first available 10 m resolution Sentinel-2 images characterized by low cloud cover, captured on 2023-05-23 (1 week after the main event). (iv) Post-event dataset S4 includes 8 Sentinel-2 images with low cloud coverage captured during June 2023.

Aerial datasets: (i) Pre-event dataset A1 includes tiles of the 0.2 m resolution 2020 AGEA coverage (Agency for Agricultural Payments), surveyed between April and July 2020. It generally reflects the vegetation state during the event period, although there might be local variations owing to different agricultural practices. (ii) Post-event dataset C1 includes tiles of the 0.2 m resolution 2023 CGR coverage, commissioned ad hoc by the Emilia-Romagna Region in the aftermath of the 2023 MORLE; it is the same dataset used for manual landslides mapping by Berti et al (2025), and Casola Valsenio has been surveyed on 2023-05-23.

Figure 3 highlights the relevance of image resolutions (both in RGB and NIR spectrums) for detecting landslides in our scenario. It is noteworthy that the smallest landslides, those under approximately 400 m², can be identified in the high-resolution CGR images (Fig. 3a, b), whereas they are represented by a mere 4 × 4 pixels or less in the Sentinel-2 images (Fig. 3c, d).

Automated landslides mapping using NDVI changes

NDVI changes have been used to automatically map landslides that caused disruption of vegetation (Fiorucci et al. 2019; Scheip and Wegmann 2021). Ideally, pre- and post-event images should refer to a similar season on the year (so to have comparable vegetation states), and the post-event images should be from soon after the event (so to better capture changes due to landslides). This should allow to minimize the risk of mistakenly attributing to landslides vegetation changes that are actually caused by other factors, such as agricultural activities, deforestation, or fires (Mondini et al. 2011; Yang et al. 2019).

We implemented the NDVI changes method with various combinations of pre- and post-event imagery as outlined in Table 2 (NDVI-1/2/3). For NDVI-1, we computed pre-event and post-event NDVI using datasets S1 and S4, respectively. This involved averaging NDVI values from multiple images over approximately 2 months to reduce noise and anomalies. Conversely, NDVI-2 was calculated using datasets S2 and S3 to reflect the vegetation state from individual Sentinel-2 images, enabling the detection of localized vegetation changes. Lastly, NDVI-3 was conducted using high-resolution aerial images from datasets S3 and C1, with a resolution of 0.2 m. This, potentially, allows for the identification of small-scale vegetation changes not observable in images with 10 m resolution, but it also introduces additional "noise," primarily due to shadows and lighting conditions (Burgess et al. 1995; Yang et al. 2022). To address this, we resampled the original 0.2 m images to 0.5, 1, 2, 5, and 10 m pixel size using bilinear interpolation, which preserves the spectral integrity of the original spectral bands. Ultimately, a 1 m raster resolution was found to offer the best balance between accurate landslide delineation and limitation of noise. All the Sentinel-2 images used in the NDVI analysis (NDVI-1 and NDVI-2) were manually inspected and confirmed to be cloud-free in the area of interest. Therefore, no additional cloud masking (e.g., using s2cloudless or Sentinel QA bands) was applied.

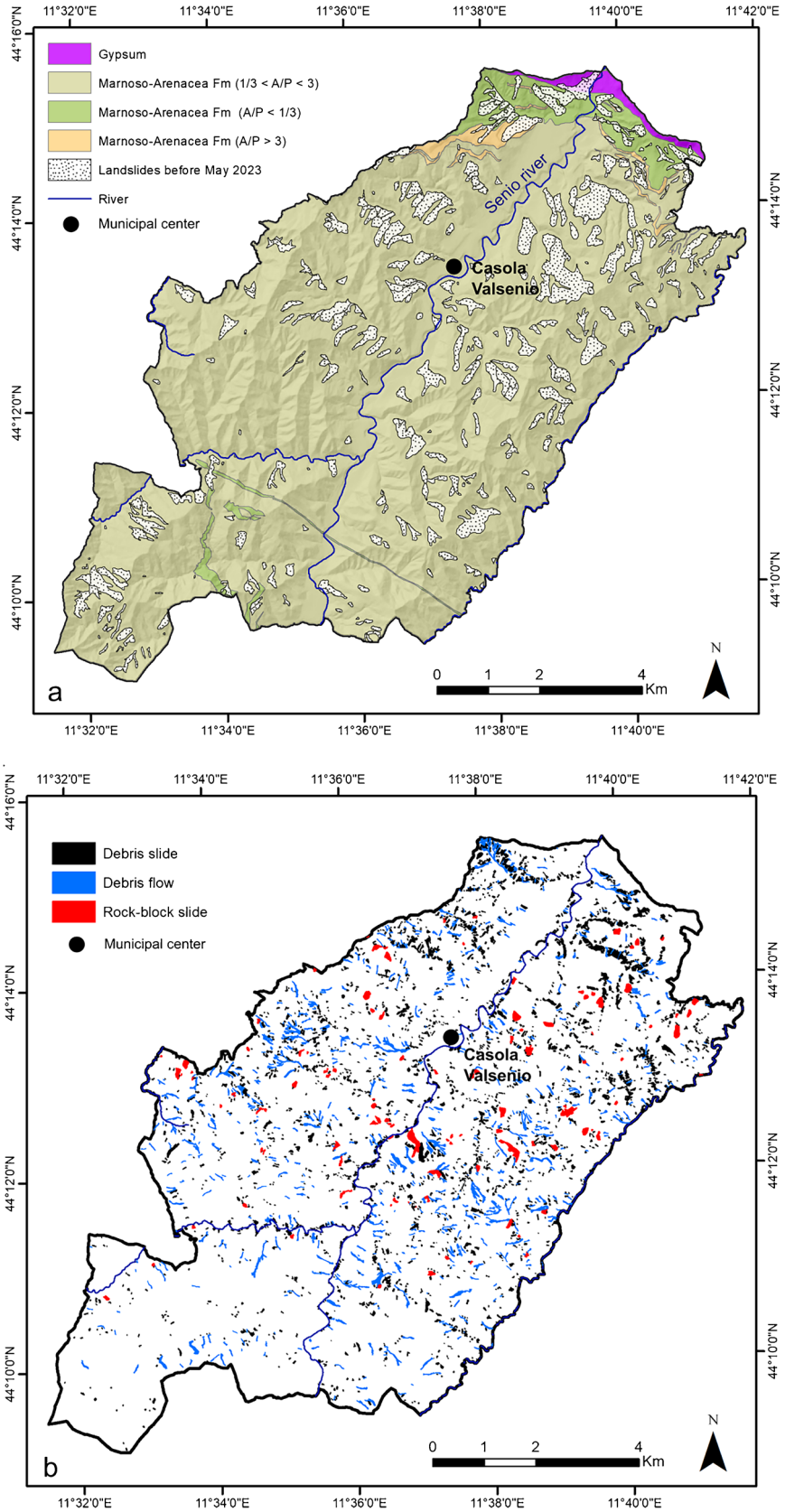


Fig. 2 **a** Schematic geological map of the municipality of Casola Valsenio (Province of Ravenna, Italy). The Marnoso-Arenacea Formation is categorized into three distinct members based on the prevailing Areneties/Pelites (A/P) ratio. **b**) Landslides triggered by the May 2024 MORLE manually identified using high-resolution aerial imagery (0.2 m resolution) (from Berti et al. 2025)

Table 1 Satellite and aerial imagery used for landslide mapping through manual and automated methods

Dataset	Period	Image source	Image type	Spectral bands	Pixel (m)	Date	Number of images
S1	Pre-event	Sentinel-2	Satellite	RGB + NIR	10	5 May–30 June 2022	9
S2	Pre-event	Sentinel-2	Satellite	RGB + NIR	10	13 May 2022	1
S3	Post-event	Sentinel-2	Satellite	RGB + NIR	10	23 May 2023	1
S4	Post-event	Sentinel-2	Satellite	RGB + NIR	10	23 May–30 July 2023	8
A1	Pre-event	AGEA	Aerial	RGB + NIR	0.2	April–July 2020	1
C1	Post-event	CGR	Aerial	RGB + NIR	0.2	23 May 2023	1

Automated landslides mapping using U-Net

U-Net is a convolutional neural network (CNN) originally developed for biomedical image segmentation (Ronneberger et al. 2015). The architecture of U-Net features a contracting (downsampling) path that captures the context and details of the input image, and an expansive (upsampling) path designed for precise localization of features. A key feature of U-Net is the incorporation of "skip connections," which allows the network to bypass deeper layers in certain instances, enabling the transfer of rich, local information (such as textures or specific shapes) directly across the network without loss of detail. This feature is particularly useful for tasks requiring high precision in image reconstruction, such as the segmentation of complex structures (Litjens et al. 2017). As in any machine learning model, the effectiveness of U-Net heavily depends on the quality and suitability of the training dataset that accurately represents the variety of conditions under which the target features can occur (Minaee et al. 2022).

U-Net has been extensively used for landslide mapping (Soares et al. 2020; Qi et al. 2020; Meena et al. 2022; Nava et al. 2022; Bhuyan et al. 2023; Lu et al. 2023; Chandra et al. 2023; Shafapourtehrany, et al. 2023; Chen et al. 2023). In these studies, the model was trained using image tiles of varying sizes, sampled through different strategies, including random and regular grid sampling, and commonly enhanced with data augmentation techniques such as random rotations and flips (Ronneberger et al. 2015). Currently, there are no established general rules or standardized procedures regarding the optimal number of tiles, their size, and other specific parameters.

For our analysis, we initially employed a standard U-Net architecture integrated into an open-source Python script (Iakubovskii 2019). This script was used to conduct several numerical experiments and to validate the outcomes against the built-in functionality in MATLAB. Ultimately, we opted to perform the final analysis using MATLAB, as it offers enhanced facilities for processing and visualizing geospatial data. The U-Net architecture itself comprises a moderately deep encoder with four levels, each containing two convolutional layers, followed by a ReLU activation function, and max-pooling layers interspersed to decrease spatial dimensions. The network employs 32 filters and utilizes the Adam solver for optimization with an adjustable initial learning rate ranging from 0.0001 to 0.01.

The U-Net analyses were carried out using both Sentinel-2 (case UNET-1, dataset S3; Table 2) and high-resolution CGR imagery (case UNET-2, dateset C1; Table 2). To manage the challenges of memory allocation and computational time, CGR images were resampled to a 2 m resolution. Deep learning models like U-Net, in fact, require significant memory for storing input images, handling augmented data and training (Wang et al. 2024). For instance, a 1 km² image at 2 m resolution requires 4 MB of memory and forms a 500 × 500 pixel array, whereas the same area at the original 0.2 m resolution requires 400 MB for a 5000 × 5000 pixel array. A computer with 12 GB of VRAM can process multiple 2 m resolution images but struggles with a single full-resolution image. Furthermore, training time escalates non-linearly with image size; training a tile for 100 epochs takes about 1 min at 2 m resolution but could extend to 5 days at 0.2 m on a GeForce RTX 3060 GPU. The decision to use lower resolution was dictated by the limits of computational power and hardware resources typically available in civil protection agencies or technical support structures, where advanced computing infrastructure may not always be accessible. A commonly adopted alternative in such cases is to preserve high resolution by splitting the imagery into smaller tiles. However, in our application, this approach proved problematic. Tiles smaller than 1 km² often contained too few landslides for effective training, and a significant number of polygons became fragmented at tile boundaries, resulting in unrealistic shapes that negatively affected model performance. For these reasons, reducing image resolution emerged as the only viable solution for running the model on standard hardware.

NDVI and U-Net training and testing approaches

Automated landslide mapping requires ground-truth datasets for training and testing models. Ideally, ground-truths should be distributed across the study area and should encompass diverse conditions related to terrain types, vegetation, and landslide types and shapes. Conventionally, ground-truth dataset is randomly divided into training and testing segments, typically using 80–20% or 70–30% proportions (Chung and Fabbri 2005; Nava et al. 2022; Chandra et al. 2023). However, such approach implies to have substantially completed the landslides inventory before being able to train and test the automated mapping algorithms. This undermines the main added value of any automated algorithm, namely its capacity to provide reliable landslides maps with limited time and effort. Consequently, we have developed an unconventional training approach to explore the possibility of training NDVI and U-Net using minimal ground-truth datasets.

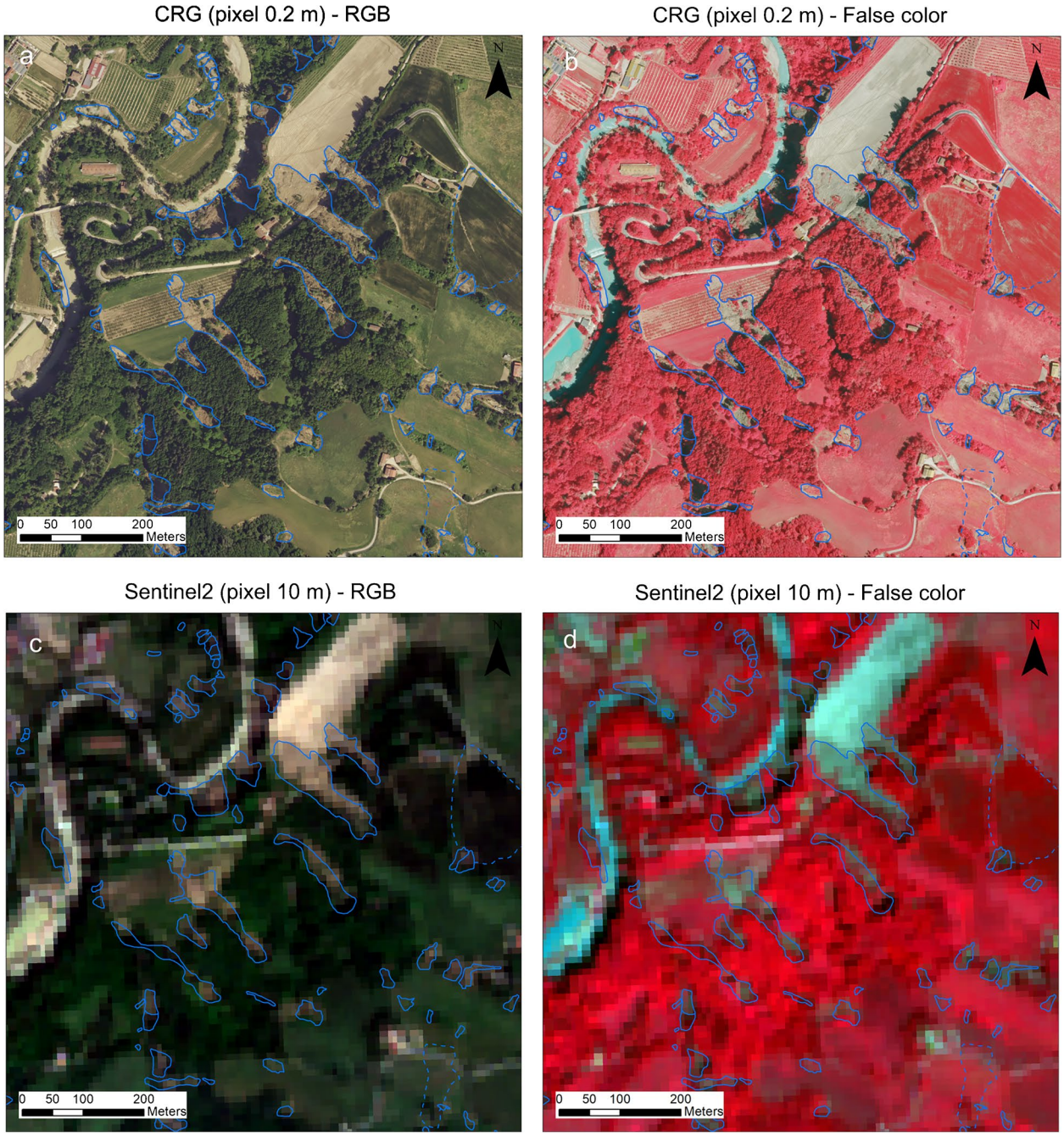


Fig. 3 Examples of imagery from a small section of the study area illustrating various products. **a** CGR aerial image with 0.2 m resolution in RGB; **b** CGR aerial image with 0.2 m resolution in NIR false colors; **c** Sentinel-2 satellite image with 10 m resolution in RGB; **d** Sentinel-2 satellite image with 10 m resolution in NIR false colors

At first, the study area was segmented into 86 grid tiles of 1 × 1 km size (Fig. 4). Each tile corresponds to 100 × 100 pixels in the 10 m Sentinel-2 imagery, 500 × 500 pixels in the CGR imagery resampled at 2 m, and 1000 × 1000 pixels in the CGR imagery resampled at 1 m. Generally, a single grid tile contains over 50 landslides, providing a reliable overview of landslide pattern. Then, we proceeded to train the NDVI and U-Net models using a limited number of tiles, specifically 1–2-4–6-8–10-12 tiles. Considering a total of 86 tiles, our training dataset ranged from 1% (1 tile) to 14% (12 tiles) of the entire area. Testing was then carried out on the remaining tiles, leading to a variable training-to-testing ratio from 1–99% to 16–84%.

Our approach incorporated two different strategies for tile selection: a random layout and a clustered layout (Fig. 4). In the random layout, training tiles are randomly selected across the study area. This approach aims to capture a wide range of conditions, thereby enhancing

Table 2 Summary of the cases considered in the automated landslide mapping. Dataset codes refer to Table 1

Method	Case	Pre-event dataset	Post-event dataset	Pixel (m)	Description
NDVI	NDVI-1	S1	S4	10	Low resolution Sentinel-2 images (10 m). Pre- and post-event NDVI maps created by averaging a series of images
	NDVI-2	S2	S3	10	Low resolution Sentinel-2 images (10 m). Pre- and post-event NDVI maps obtained by a single image
	NDVI-3	A1	C1	1	High-resolution aerial images (0.2 m) resampled at 1 m
U-Net	UNET-1	\	S3	10	Low resolution post-event Sentinel-2 image (10 m)
	UNET-2	\	C1	2	High-resolution post-event aerial image (0.2 m) resampled at 2 m

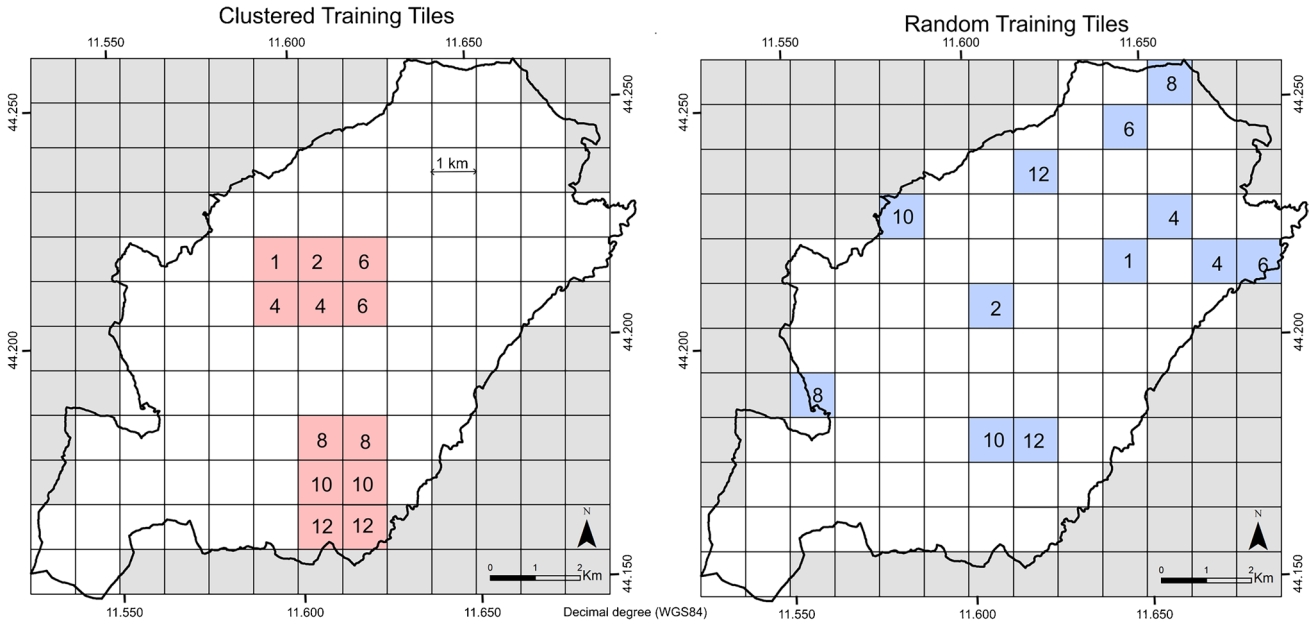


Fig. 4 Location of tiles utilized for training the automated mapping models in both clustered and randomly selected training scenarios. The numbers show the sequential inclusion of tiles in the analysis

the statistical robustness and generalizability of the results. In contrast, the clustered layout involves grouping training tiles that are geographically close to each other. This approach is designed to simulate scenarios where post-event imagery is only available for a limited area and manual mapping is possible only for specific local conditions. These selection strategies led to the creation of 14 unique training scenarios: seven in the clustered layout and seven in the random layout (Fig. 4).

In the NDVI method, training consists of identifying the NDVI change threshold (ΔNDVI) that best distinguishes landslide pixels from non-landslide pixels. All NDVI models (NDVI-1 to NDVI-3) rely on simple NDVI differencing between pre- and post-event images. To determine the optimal threshold, a range of ΔNDVI values (from −2 to + 2) was tested on the training dataset. For each value, the corresponding F1-score was calculated (Tharwat 2018), as the F1-score is well-suited for highly unbalanced datasets like ours, in which landslide pixels typically account for less than 10–15% of each tile. The threshold that yielded the highest F1-score on the training tiles was selected as the optimal value. This adaptive threshold was then applied to the testing tiles, and model performance was again evaluated using the F1-score.

In the U-Net method, training is achieved through a process where batches of data are continuously fed into the network. The network makes predictions, and the weights are adjusted based on the calculated loss from these predictions. This cycle of prediction, adjustment, and feedback is repeated across multiple epochs, or complete passes through the training dataset, until the network achieves satisfactory performance levels. Once trained, the U-Net model was applied to test tiles to assess its effectiveness in real-world scenarios. The accuracy and reliability of these predictions were evaluated using the F1-score.

To further assess the generalizability of the U-Net model and improve the statistical robustness of the performance metrics, we also implemented a k-fold cross-validation strategy. The 86 tiles

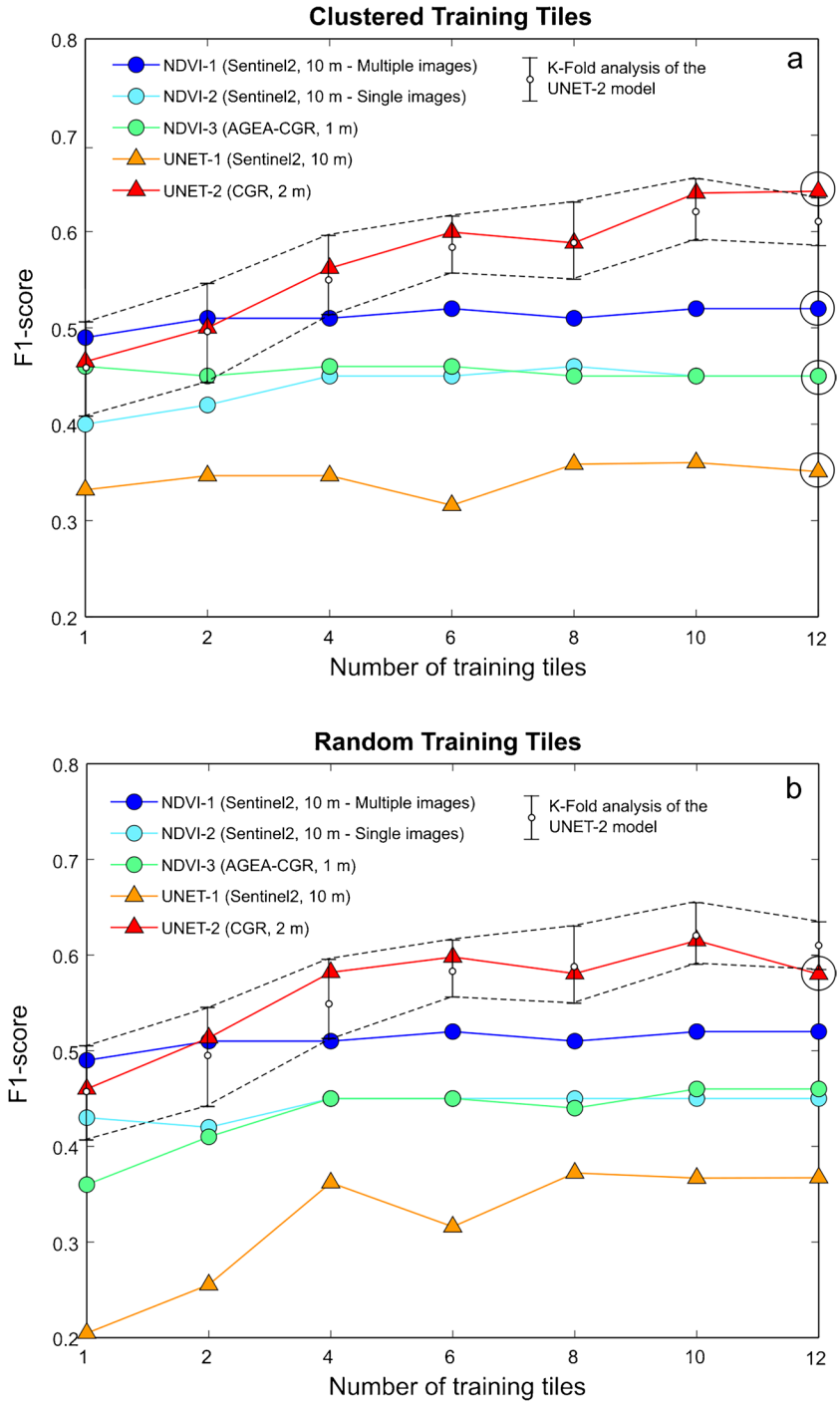


Fig. 5 F1-score values for different automated mapping methods as a function of the number of training tiles. **a** Clustered distribution of the training tiles; **b** Random distribution of the training tiles. The circled symbols mark the cases depicted in Fig. 6

were partitioned into k folds, with k set to 7, 8, 10, 15, 22, 44, and 86, corresponding to 12, 10, 8, 6, 4, 2, and 1 tiles per fold, respectively. For each value of k, the models were trained on $k - 1$ tile groups randomly selected and tested on the remaining group. This process was repeated k times, yielding k F1-scores per configuration. The average F1-score ± standard deviation across folds was then compared to the values obtained using the random and clustered training layouts.

Automated identification of affected buildings and roads

The prompt identification of damages caused by landslide is crucial in the aftermath of a MORLE. For instance, in the May 2023 event, a rapid assessment of damages and related costs was required to coordinate responses and to gather economic resources.

Table 3 Effectiveness of automated mapping models across scenarios outlined in Table 2. F1-score values represent the highest score achieved on the validation dataset. "Threshold" denotes the optimal NDVI change threshold identified

Case	Training tiles distribution	Parameter	Number of training tiles						
			1	2	4	6	8	10	12
NDVI-1	Clustered	F1-score	0.49	0.51	0.51	0.52	0.51	0.52	0.52
		Threshold	−0.1	−0.12	−0.12	−0.12	−0.12	−0.14	−0.14
NDVI-2		F1-score	0.4	0.42	0.45	0.45	0.46	0.45	0.45
		Threshold	−0.08	−0.1	−0.14	−0.16	−0.16	−0.16	−0.16
NDVI-3		F1-score	0.46	0.45	0.46	0.46	0.45	0.45	0.45
		Threshold	−0.34	−0.32	−0.35	−0.34	−0.34	−0.35	−0.35
NDVI-1	Random	F1-score	0.49	0.51	0.51	0.52	0.51	0.52	0.52
		Threshold	−0.1	−0.12	−0.16	−0.16	−0.16	−0.14	−0.14
NDVI-2		F1-score	0.43	0.42	0.45	0.45	0.45	0.45	0.45
		Threshold	−0.2	−0.22	−0.18	−0.18	−0.17	−0.17	−0.17
NDVI-3		F1-score	0.36	0.41	0.45	0.45	0.44	0.46	0.46
		Threshold	−0.47	−0.43	−0.37	−0.37	−0.41	−0.37	−0.37
UNET-1	Clustered	F1-score	0.33	0.35	0.35	0.32	0.36	0.36	0.35
UNET-2		F1-score	0.47	0.50	0.56	0.60	0.59	0.64	0.64
UNET-1	Random	F1-score	0.20	0.26	0.36	0.32	0.37	0.37	0.37
UNET-2		F1-score	0.46	0.51	0.58	0.60	0.58	0.62	0.58

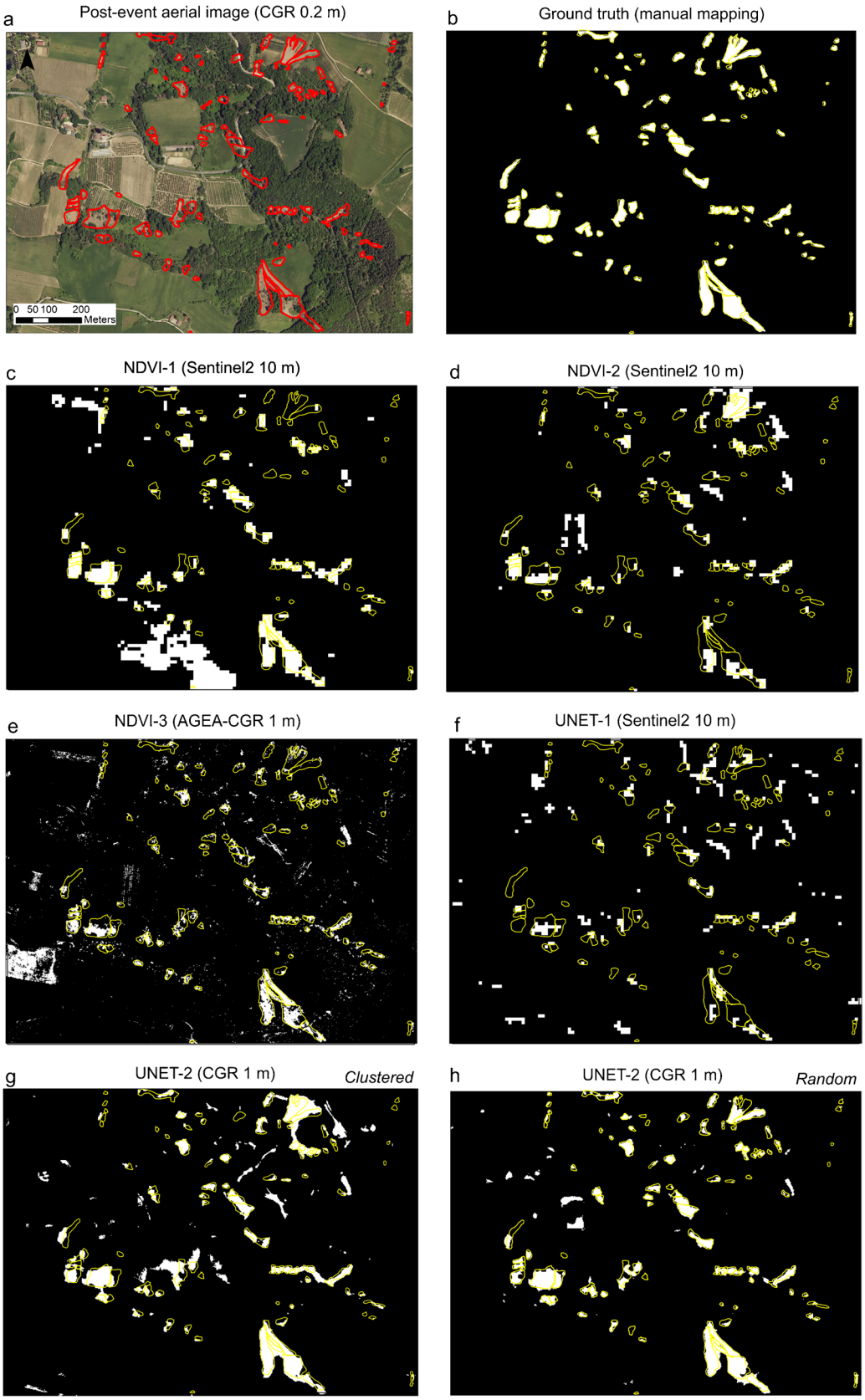
To test the automated identification of buildings and roads impacted by landslides, the maps generated by NDVI and U-Net were overlaid to the map of vulnerable elements. Such dataset (available in vector format from the Emilia-Romagna region's topographic database, updated in 2024, just before the disaster) includes buildings, cultural heritage sites, bridges, hydraulic structures, tunnels, roads, and railways. Mirroring the criteria used for manual damage identification from high-resolution aerial images (see Sect. "Landslides and damages dataset used for training and testing"), we extracted buildings and roads, and any of these elements intersected by, or within a 20 m buffer, an automatically mapped landslide, is considered possibly damaged. The effectiveness of the automated identification is assessed using a confusion matrix using the manual detections as a ground-truth, thus computing correctly identified potential damages (true positives TP), missed potential damages (false negatives, FN), and erroneous detections (false positives, FP). Since only a small fraction of buildings and roads was affected by the disaster, true negatives (TN) were not considered, and the precision of the models was measured as $P = TP/(TP + FP)$.

Results

Automated landslides mapping

The performances, in terms of F1-score, of the different automatic mapping techniques with reference to the different combinations of clustered and random training tiles show that (Fig. 5, Tab.3)

- The U-Net method outperforms the NDVI approach when using high-resolution images (UNET-2, $F1 = 0.64$), with F1-scores up to 30% higher than those achieved by the best NDVI map (NDVI-1, $F1 = 0.52$).
- U-Net's effectiveness significantly drops with lower-resolution Sentinel-2 images (UNET-1, $F1 = 0.35$), bringing its performance below that of the NDVI approach.
- NDVI's performance is more influenced by the choice of pre- and post-event images rather than image resolution. The best NDVI results are obtained using Sentinel-2 images where pre- and post-event scenes are averaged over a span of about 2 months (NDVI-1). In contrast, employing a single pre- and post-event image (NDVI-2 and NDVI-3) leads to considerably lower NDVI scores ($F1 = 0.45$ in both cases). Unexpectedly, reducing the pixel size from 10 to 1 m does not lead to an improvement in the F1-scores.
- The optimal NDVI threshold varies with image resolution, showing values around −0.14 to −0.12 for Sentinel-2 images (10 m resolution) and between −0.40 and −0.34 for high-resolution images downsampled to 1 m (Table 3).
- The arrangement of training tiles, whether clustered or randomly placed, seems to have a negligible impact on the accuracy. Notable differences in performance are only seen when U-Net is trained with a very limited number of tiles (less than 4). These findings are supported by the *k*-fold cross-validation applied to the UNET-2 model, which yields consistent F1-scores across varying numbers of training tiles (Fig. 5a, b).


a
Post-event aerial image (CGR 0.2 m)
0 50 100 200
Meters
b
Ground truth (manual mapping)
c
NDVI-1 (Sentinel2 10 m)
d
NDVI-2 (Sentinel2 10 m)
e
NDVI-3 (AGEA-CGR 1 m)
f
UNET-1 (Sentinel2 10 m)
g
UNET-2 (CGR 1 m)
Clustered
h
UNET-2 (CGR 1 m)
Random

◂**Fig. 6** Sample results of automated methods for landslide mapping in a selected area of Casola Valsenio. **a**, **b** Baseline data established through manual mapping; **c** landslide map generated using the NDVI change method with an average from multiple Sentinel-2 images; **d** similar to **c** but with NDVI change calculated using a single image; **e** landslide map derived from NDVI change in high-resolution AGEA-CGR images; **f** landslide map produced using UNET on Sentinel-2 images; **g**, **h** landslide maps created with UNET applied to CGR images, showing outcomes for both clustered (**g**) and randomly distributed (**h**) training tiles. Yellow polygons in **b**–**h** indicate the ground-truth landslide outlines

- A relatively small number of tiles, covering about 2–3% of the total study area, is sufficient to achieve the highest score for each method.

Figure 6 offers a visual comparison of the prediction maps generated by the different models over a selected portion of the study area, chosen for its diverse landscape features, including agricultural fields, forested slopes, and landslides of varying shapes and sizes. The maps were produced using 12 training tiles (indicated by circled symbols in Fig. 5). As shown, the NDVI method applied to Sentinel-2 imagery (Fig. 6c, d) successfully detects most landslides, though its accuracy is limited by the 10 m spatial resolution. Increasing the resolution to 1 m with AGEA-CGR imagery (Fig. 6e) enhances the mapping detail but introduces a noticeable salt-and-pepper effect that reduces statistical performance. In contrast, the UNET model yields accurate results only when using high-resolution imagery (compare Fig. 6f with Fig. 6g, h). Its performance appears only slightly affected by whether the training layout is clustered (Fig. 6g) or random (Fig. 6h).

The limited performances of U-Net with Sentinel-2 images (Fig. 6f) is largely due to the limited resolution of these images, which hinders the detection of smaller landslides. Landslides that are less than 400–500 m^2, roughly equivalent to 2×2 pixels, are hard to spot out even with the naked eye. The training of U-Net, which relies on pattern recognition and contextual understanding similar to human cognitive processes, encounters limitations under these conditions. On the contrary, U-Net shows its effectiveness with Sentinel-2 images for larger landslides that cover a more substantial number of pixels (Chen et al. 2023).

A closer analysis of the results from two methods reveals differences not fully captured by the F1-scores. Figure 7 presents the confusion matrix and the distribution of misclassified pixels in the two best-performing predictions (NDVI-1 and UNET-2), categorized by land use type. For each category, the bars indicate the percentage of false positives (FP) and false negatives (FN) within that class, relative to the total number of misclassified pixels. Figure 8 provides visual examples of these classification errors.

The analysis shows that NDVI-1 generates a considerable number of false positives (FP) in arable lands, forested areas, and along riverbanks. In arable lands, FP often arise when post-event imagery captures plowed fields or a reduction in vegetation cover compared to pre-event images, leading the algorithm to mistakenly identify these areas as landslides (see sample areas E1 in Fig. 8e). In forested areas, false positives typically occur where trees have been cleared (E2 in Fig. 8e), or around actual landslides, due to the coarse spatial resolution of the Sentinel-2 imagery (E3 in Fig. 8e). Additional false positives are observed along rivers, where flood-induced vegetation removal confuses the algorithm's classification (E4 in Fig. 8e). False negatives, though less frequent, typically occur when landslides develop on agricultural lands that were already bare prior to the event or, in forested slopes, where the landslide debris spreads through the trees without significantly disturbing the vegetation. Although the canopy remains largely intact, these areas were included in the landslide polygons during manual mapping.

UNET-2 encounters similar challenges but achieves higher accuracy, with a lower rate of misclassified pixels (compare the confusion matrices in Fig. 7). This improved performance is due to the ability of convolutional neural networks to consider not only changes in pixel values but also the shape and spatial patterns of the features to be identified (Ronneberger et al. 2015). For instance, some of the plowed fields marked as E1 in Fig. 8e are correctly classified as non-landslide areas by UNET (Fig. 8f). However, false positives still occur in cases where cleared forests or plowed terrain exhibit shapes that resemble landslides (E5 in Fig. 8f). Another notable source of false positives includes backyards and rooftops (urbanized areas in Fig. 7), which UNET occasionally misclassifies as landslides due to subtle color variations between pre- and post-event images, often caused by differences in illumination or viewing angle. False negatives are generally limited, except on forested slopes. In these areas, the issue arises from manual mappings that often include both the detachment and runout zones within the landslide polygons, even when vegetation is still partially present (E6 in Fig. 8f). UNET struggles to replicate this level of expert interpretation, which relies on contextual understanding beyond spectral or spatial cues.

Regarding the ability to identify the different types of landslides, both methods perform better with DS and DF, but struggle with RS. As expected, DS and DF are easier to spot as they stripped away vegetation, in contrast to RS, in which substantial parts of the sliding blocks may preserve vegetation. In such cases (see an example in Fig. 9), both NDVI and U-Net methods detect the detachment area and the fragmented debris at the toe as distinct landslide events, but they fail to identify the real full extent of the landslide.

Figure 10 provides an additional comparison between the two methods. It displays the 1×1 km tiles colored according to their F1-scores, computed from high-resolution imagery using 12 clustered training tiles and the two best-performing models NDVI-1 and UNET-2. The results indicate that UNET-2 reaches F1-scores higher than 0.5 in 75 of 86 tiles (approximately 87%), while NDVI-1 reaches similar scores in only 44 of 86 tiles (about 51%). In U-Net, the tiles with the lowest F1-scores have simultaneously low TP and high FP. This is often observed in tiles with a sparse presence of landslides, where the few positive pixels do not align well with the actual occurrences. Conversely, in the NDVI model, the lowest F1-scores are primarily due to high FP. These inaccuracies tend to be distributed randomly across the tiles and are in many cases associated with agricultural activities. As expected, the training tiles (marked with a thick black border in Fig. 10) exhibit higher F1-scores, particularly for U-Net, which is designed to identify complex data patterns. This effect is more subdued for NDVI, as this method mainly tracks changes in vegetation and is less sensitive to localized conditions.

Finally, Fig. 11 compares the landslide index (% ratio of landslide area to tile area) obtained by manual mapping (observed) and by NDVI-1 and UNET-2 for all of the 86 tiles of Casola Valsenio (predicted). U-Net demonstrates a higher R^2 and lower RMSE

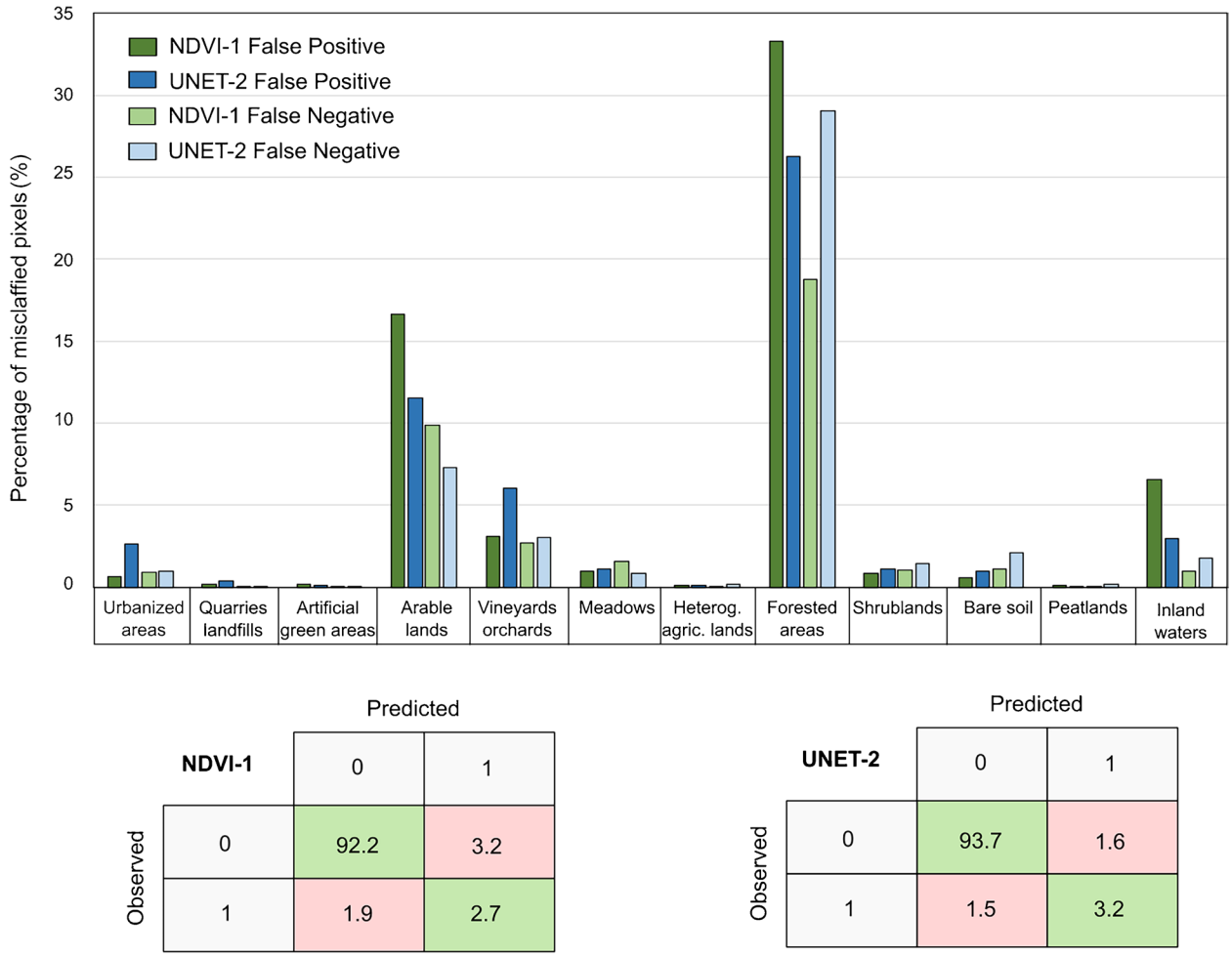


Fig. 7 Histogram showing the distribution of misclassified pixels for the two best-performing models, NDVI-1 and UNET-2. Each bar represents the percentage of false positives or false negatives associated with different land use categories, relative to the total number of misclassified pixels. Below the histogram, the confusion matrices for both models are displayed (0 = no landslide; 1 = landslide). Values in the confusion matrices are expressed as percentages of the total number of pixels

compared to NDVI, although both methods successfully estimate the proportion of landslide area over a broad range of index values, approximately from 0 to 20%.

Automated identification of affected buildings and roads

The automated identification of affected buildings and roads is somehow more problematic than the mapping of landslides. Figure 12 compares the counts of houses and roads affected by landslides determined by manual mapping or by NDVI - U-Net. Manual mapping identified 64 houses and 185 roads impacted by landslides. U-Net, utilizing high-resolution imagery (UNET-2) and trained on 12 tiles, was able to detect a significant portion of the affected roads and buildings (65% and 72%, respectively). However, it also generated a significant number of FP, particularly with buildings, often misclassifying roof tops as landslide areas. On the other hand, the NDVI using Sentinel-2 (NDVI-1) had fewer issues with false positives, but its accuracy in identifying affected buildings drops to 49%. For both methods, the precision (*P*) is below 0.52 for both roads and buildings.

Discussion

In this study, we evaluated two automated methods for rapid mapping to aid in the detection of landslides in emergency scenarios. Our evaluation took place in a selected area within the Emilia-Romagna Region which was affected by a MORLE in May 2023. The maps generated through automated processes were analyzed against a manually created map by expert geologists. Overall, on the basis of the F1 score, results of automated landslides mapping are in most cases satisfactory. However, these results come with some expectable limitations.

A first limitation is that we deliberately chose not to use additional data layers to train U-Net or integrate NDVI changes in a multicriteria framework. Our goal was to test the predictive capability achievable using only multispectral images, assuming that other datasets might not be immediately available after a MORLE. Nonetheless, incorporating additional layers can enhance accuracy. As shown in Fig. 12, adding a fifth input band representing NDVI change improves the detection of damaged buildings and roads. The UNET-A and UNET-B models extend the baseline UNET-2 (which uses four high-resolution

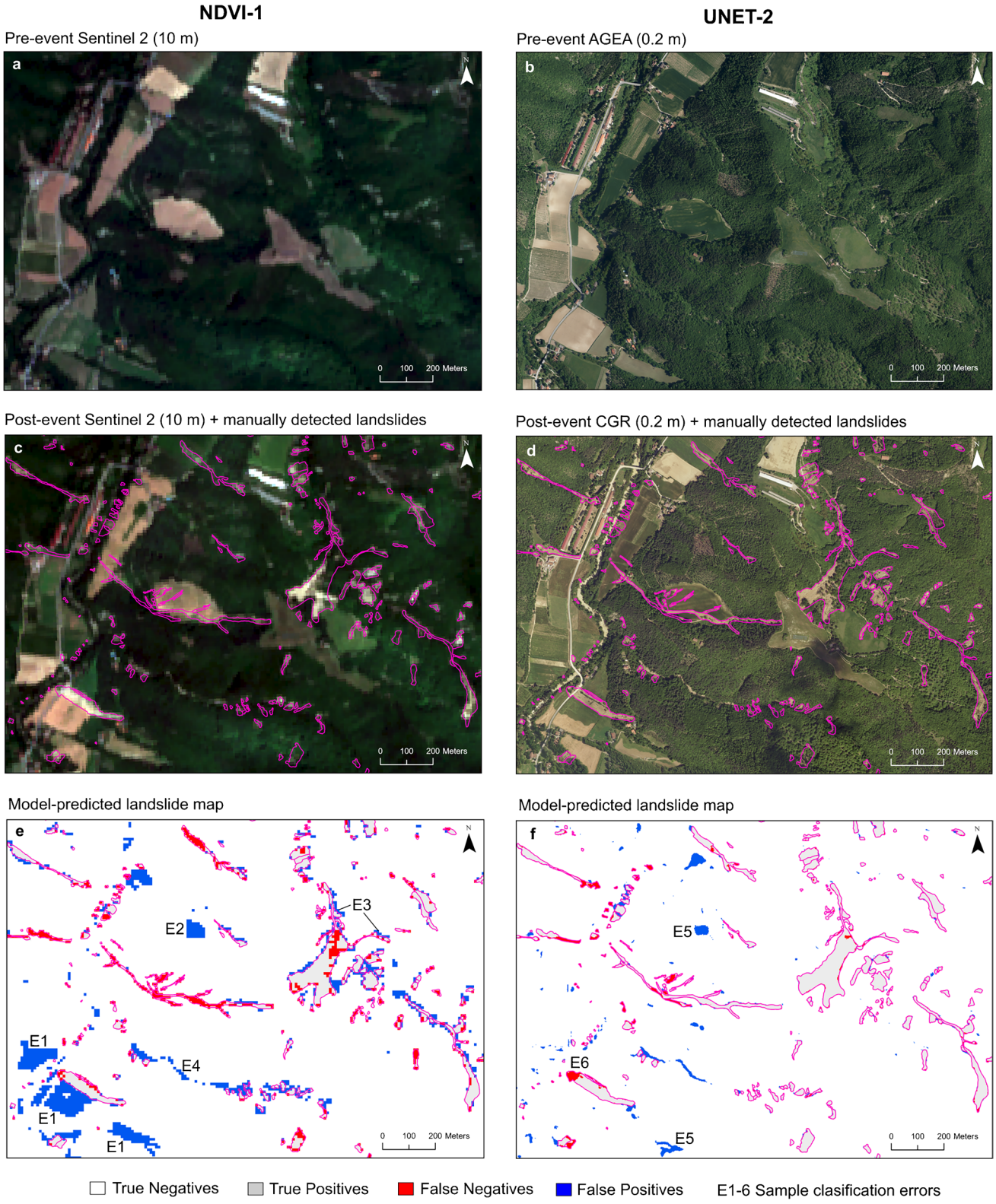


Fig. 8 Examples of classification errors for the two best-performing models, NDVI-1 and UNET-2. **a**, **b** Pre-event images; **c**, **d** post-event images with manually mapped landslides; **e**, **f** classification error maps produced by the two methods. Labels E1–E6 indicate sample areas of misclassification discussed in the text

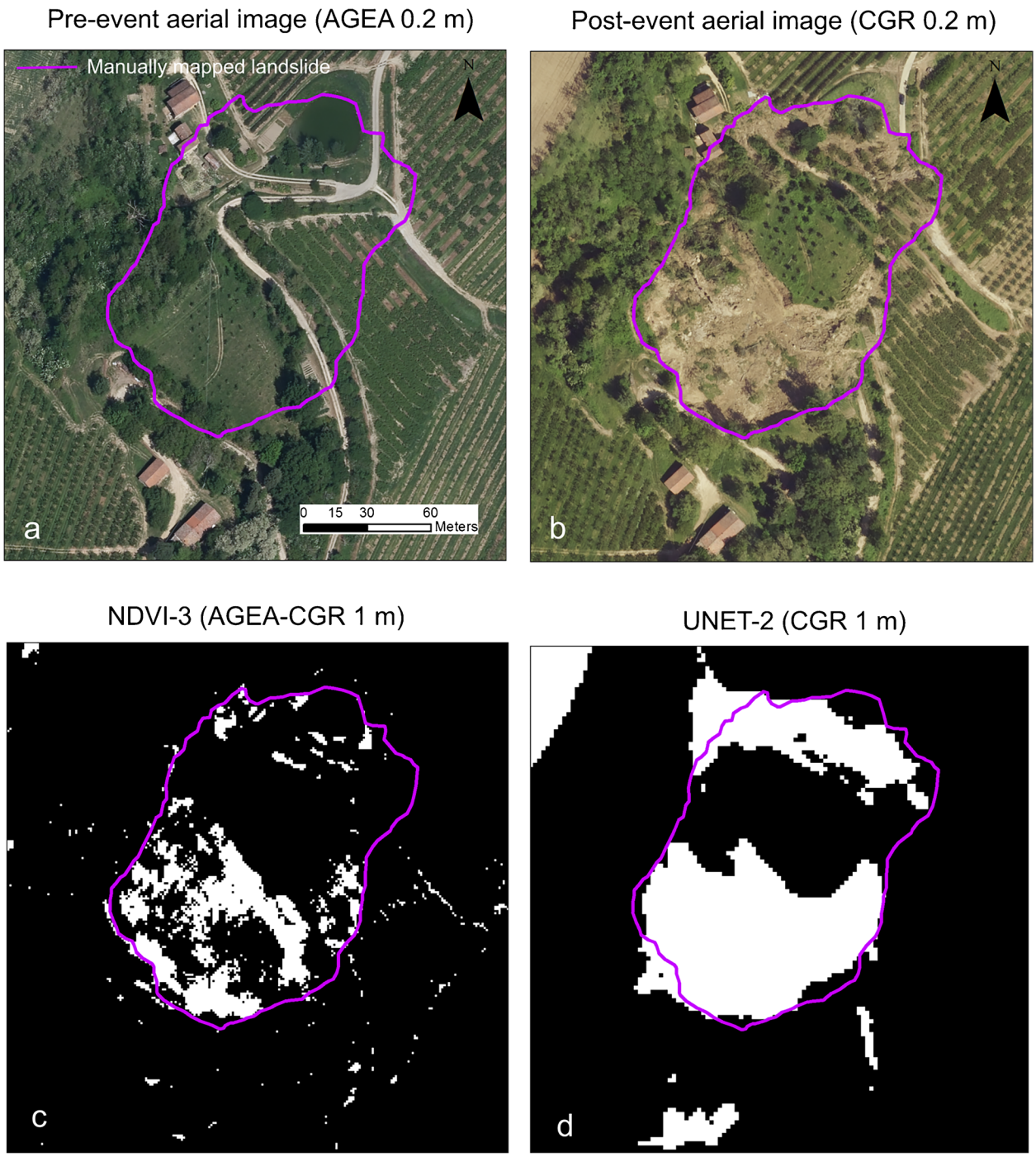


Fig. 9 Challenges in automatically mapping rock-block slides. **a**, **b** High-resolution images before and after a typical rock-block slide in the study area; **c**, **d** automated landslide maps generated from high-resolution imagery using the NDVI change method (**c**) and the UNET algorithm (**d**). Purple polygon indicates the manually mapped landslide (ground truth)

AGEA-CGR bands) by including an NDVI change layer derived from high-resolution orthophotos (UNET-A) or Sentinel-2 imagery (UNET-B). This leads to a notable increase in precision: for buildings, from 0.25 (UNET-2) to 0.46 (UNET-A) and 0.55 (UNET-B); and for roads, from 0.46 (UNET-2) to 0.69 (UNET-A) and 0.65 (UNET-B). The NDVI data help the network distinguish landslides, bare soil, and similarly colored areas such as rooftops, reducing false positives.

Several studies have underscored the effectiveness of a multi-layer approach in enhancing the mapping capabilities of convolutional neural networks, as these models are inherently adept at uncovering complex relationships among numerous parameters (Soares et al. 2020; Su et al. 2021; Bhuyan et al. 2023). In principle, incorporating topographic features such as slope, aspect, and curvature should improve segmentation performance by helping detect the irregular morphologies of

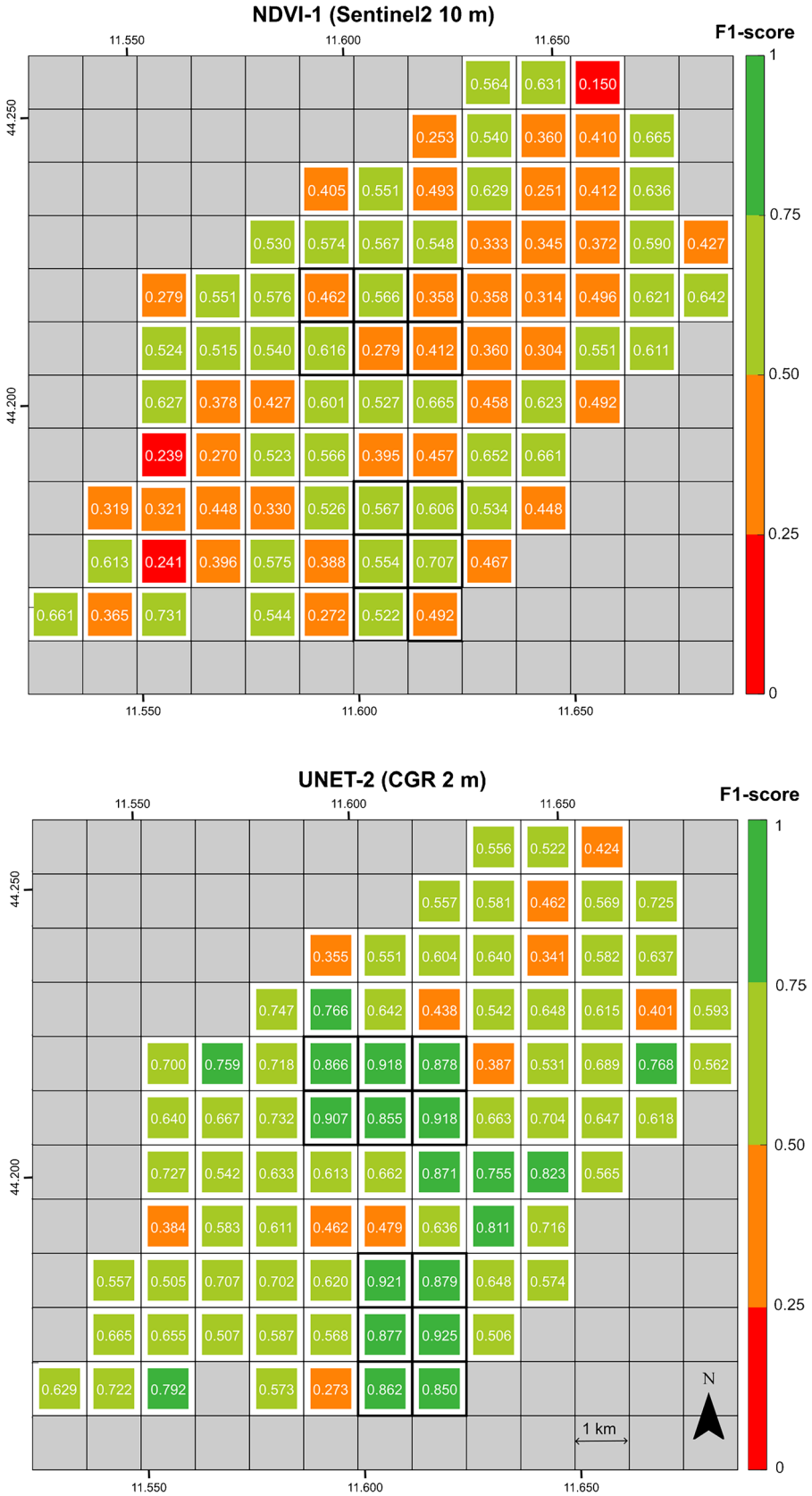


Fig. 10 Performance comparison of the landslide map produced by NDVI change (upper) versus UNET (lower) across Casola Valsenio. Each 1 km^2 tile is color-coded based on the F1-score, as indicated by the color bar. The colors indicate four accuracy levels, spanning from very high (F1 > 0.75) to very low (F1 < 0.25). Tiles utilized for training are distinctly marked with a thick black border

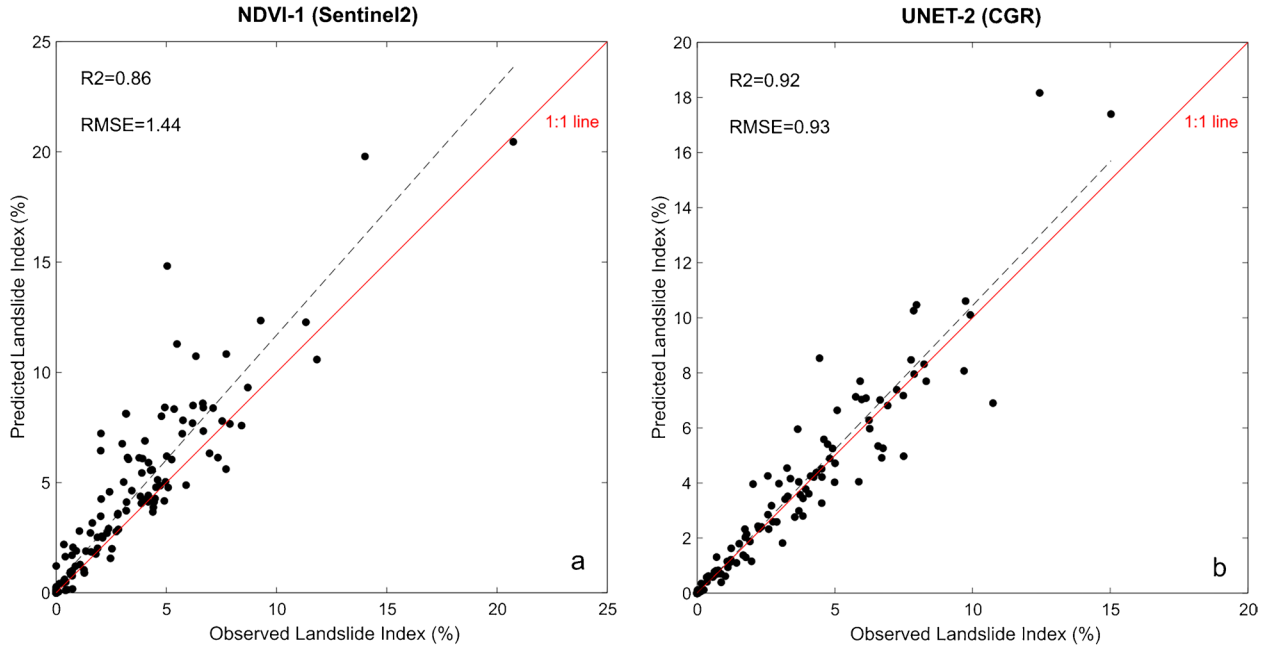


Fig. 11 Comparison of observed versus predicted landslide index for all tiles in Casola Valsenio using NDVI (**a**) and UNET (**a**) methods. Dashed lines represent the best-fit regression line, while red lines denote ideal prediction accuracy

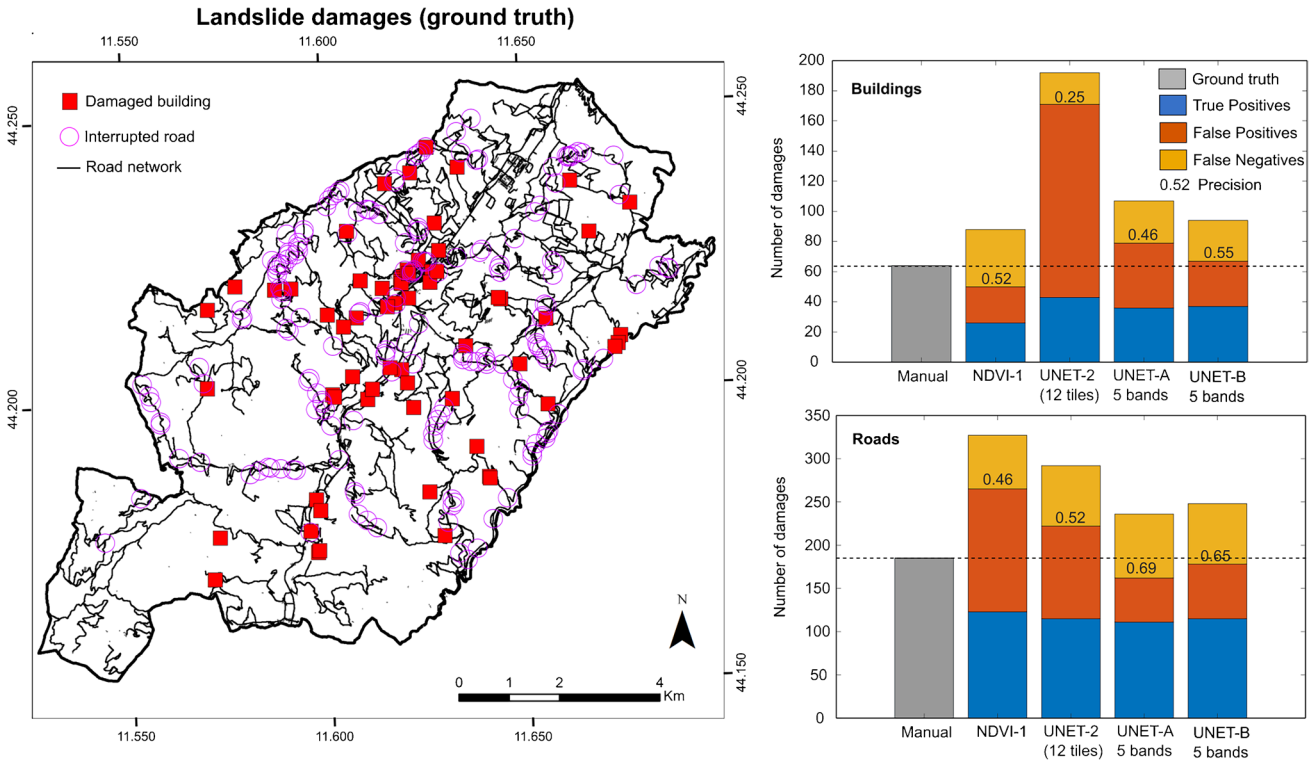


Fig. 12 Landslide-induced damages within the Casola Valsenio area. The right-side map illustrates the positions of damaged buildings and roads identified through post-event aerial image analysis. On the left, histograms show the comparison between the actual number of damaged buildings (upper) and roads (lower) versus predictions made by various automated techniques. NDVI-1 represents the NDVI change from Sentinel-2 images; UNET-2 denotes UNET application to CGR images; UNET-A/B are variations of UNET-2 incorporating an additional fifth band from NDVI changes using AGEA-CGR and Sentinel-2 images, respectively

landslides shaped by local terrain. However, in our case, the inclusion of these layers resulted in only marginal improvements. We tested the integration of slope, aspect, and curvature, derived from a pre-event 5 m DEM using standard GIS algorithms, as additional inputs to the UNET-2 model. The performance gain was minimal: for example, using 12 training tiles with a clustered layout, the F1-score increased from 0.64 (using only the 3-band CGR imagery) to 0.65 with the added topographic layers, representing a difference of just 0.01.

This limited gain is likely due to the relatively simple morphology of landslide boundaries in our study area, which can be effectively captured using spectral and NDVI data alone. A closer examination of the automated landslide maps suggests that much of the remaining discrepancy with the ground truth stems from the inherent difficulty of replicating expert-based mapping. Human interpreters often rely on contextual and geomorphic cues beyond pixel-level information, for example, recognizing landslide deposits beneath partially intact vegetation, or distinguishing recent slope failures from pre-existing bare soil. Such expert judgments remain difficult to reproduce with automated approaches.

A second limitation concerns the use of high-resolution aerial imagery. We resampled the original 0.2 m datasets to 2 m, thus not fully exploiting their potential. High-resolution images offer clear advantages for landslide mapping, notably the detection of smaller features (Malamud et al. 2004; Guzzetti et al. 2012). However, their use in NDVI and U-Net mapping introduces problems. NDVI changes derived from 0.2 m images generate numerous small polygons in forested areas because leaf shadows are misclassified as landslides. Spatial interpolation techniques (e.g., bilinear interpolation) mitigate this noise (Smith 1981; Bhattacharjee et al. 2014; Xu et al. 2015), but also blur edges that are important for precise delineation. More advanced denoising techniques like bilateral filtering (Annaby and Nehary 2021), DnCNN (Wang 2021), or CGANs (Kim and Lee 2020) can reduce noise while preserving edges, but have not yet been widely used in landslide research. Furthermore, processing U-Net with 0.2 m images was computationally very demanding. Although high-resolution images should theoretically improve deep learning performance (Bhuyan et al. 2023), they dramatically increase computational requirements and overfitting risk, particularly when the training dataset is limited (Wang et al. 2024).

Many published studies on landslide mapping use image patches ranging from 64×64 pixels to 224×224 pixels (Prakash et al., 2021; Ghorbanzadeh et al. 2021, 2022; Lu et al. 2023; Chandra et al. 2023). At a 10-m resolution of Sentinel satellite images, these patches represent large areas (from 0.4 to 5 km^2), likely containing enough landslide features for training, and are easily manageable with standard hardware. In contrast, the same patch sizes in 0.2 m aerial images correspond to much smaller areas less than 50×50 m, which might not contain enough landslide features for a proper training. Since the computational complexity of convolutional neural networks increases quadratically with spatial resolution (Ronneberger et al. 2015), high-resolution images are hard to process with standard hardware. Several recent studies address this issue with strategies such as using Transformers (Zamir et al., 2022), Deep Multi-Feature Learning (Su et al. 2021), or specialized hardware (Civit-Masot et al. 2021). These approaches are still primarily used in computer vision or medical applications and, to our knowledge, have not yet been tested for automated landslide mapping.

Another important topic is the reliability of the "ground truth," that is the landslide inventory used to train and test the U-Net and NDVI approaches. We treated the manual map as an accurate representation of the MORLE-induced landslides, but manual mapping inevitably involves subjectivity (Guzzetti et al. 2012). For example, in many cases, we included areas beneath partially intact vegetation to capture the full extent of deposits (Berti et al. 2025). Distinguishing shallow landslides from erosion, or identifying debris passing over bedrock cliffs, was also challenging. Automated methods cannot replicate such subjective expert judgement. Uniform guidelines for landslide mapping, as advocated by Novellino et al. (2024), may help standardize inventories used for training.

Despite these limitations, our findings confirm that U-Net outperforms NDVI change detection in accurately identifying landslides. NDVI frequently misclassifies agricultural fields as landslides, whereas U-Net reduces false positives by using both spectral information and shape-based recognition. Both approaches achieved relatively high F1-scores even with limited training data, suggesting that manually mapping only about 2–3% of the area may be sufficient for training and achieving near-optimal performance. This demonstrates the feasibility of applying such methods immediately after a MORLE. However, automated identification of affected buildings and roads remains difficult because it requires very high precision, particularly in areas with numerous small landslides. For the 2023 Romagna MORLE, automated mapping of impacted buildings and roads is still not reliable enough for operational use.

Finally, it should be stressed that irrespective of whether an automated or manual approach is employed, the timely availability of high-resolution images following a MORLE is crucial for effectively managing the crisis. Policies regulating access to high-resolution satellite data should therefore consider enabling their rapid sharing with public institutions involved in emergency response. Ongoing research efforts related to the Romagna 2023 disaster entail refining UNET for broader area application, analyzing landslide susceptibility, and crafting predictive models for landslide initiation and spread. The authors extend a warm invitation to colleagues for collaborative participation, aiming to extract valuable insights from this event.

Conclusions

In our study area, U-Net clearly outperforms NDVI change detection in accurately identifying landslides, mainly due to its lower susceptibility to false positives and its ability to exploit both spectral and morphological information. Both methods achieve relatively high F1-scores even with limited training data, suggesting that manual mapping of about 2–3% of the area is sufficient for effective model training. This supports the feasibility of rapid automated mapping in the aftermath of a MORLE.

However, automated mapping of affected buildings and roads remains challenging, as it requires very high precision and is still not sufficiently reliable for operational emergency response. The accuracy of automated methods is also constrained by the subjectivity of manual inventories, the computational limitations of processing high-resolution imagery, and the difficulty of reproducing expert-based geomorphological interpretation.

Overall, while automated approaches offer promising capabilities for rapid landslide detection, particularly when timely high-resolution imagery is available, they should be used with awareness of their current limitations. Continued advances in model architectures, data availability, and standardized mapping guidelines will be essential to further improve their operational value in post-MORLE scenarios.

Acknowledgements

We acknowledge the support of the Po River District Basin Authority (Collaboration Agreement for Updating the Framework of the Distressed Areas Affected by the May 2023 Event), whose keen interest and involvement in the subject matter greatly enhanced the development of this research.

Data Availability

The landslide inventory is freely accessible in the Zenodo repository (https://doi.org/10.5281/zenodo.13742643, Pizziolo et al., 2024).

Declarations

Conflict of interest The authors declare no competing interests.

AI and AI-assisted technologies in the writing process During the preparation of this work, the authors used ChatGPT 4 (chat.openai.com) to enhance the grammar and syntax, as well as to refine the sentence structure. All the content is original, and no concepts, ideas, or interpretations were produced by this tool. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Amatya P, Kirschbaum D, Stanley T, Tanyas H (2021) Landslide mapping using object-based image analysis and open source tools. Eng Geol 282:1–10

Annaby MH, Nehary EA (2021) Bilateral filters with adaptive generalized kernels generated via Riemann-Lebesgue theorem. J Signal Process Syst 93(11):1301–1322. https://doi.org/10.1007/s11265-021-01707-6

Behling R, Roessner S, Kaufmann H, Kleinschmit B (2014) Automated spatiotemporal landslide mapping over large areas using RapidEye time series data. Remote Sens Basel 6:8026–8055. https://doi.org/10.3390/rs6098026

Berti M, Pizziolo M, Scaroni M, Generali M, Critelli V, Mulas M, Tondo M, Lelli F, Fabbiani C, Ronchetti F, Ciccarese G, Dal Seno N, Ioriatti E, Rani R, Zuccarini A, Simonelli T, Corsini A (2025) RER2023: the landslide inventory dataset of the May 2023 Emilia-Romagna meteorological event. Earth Syst Sci Data 17:1055–1074. https://doi.org/10.5194/essd-17-1055-2025

Bhattacharjee S, Mitra P, Ghosh SK (2014) Spatial interpolation to predict missing attributes in GIS using semantic kriging. IEEE Trans Geosci Remote Sens 52(8):4771–4780

Bhuyan K, Meena SR, Nava L et al (2023) Mapping landslides through a temporal lens: an insight toward multi-temporal landslide mapping using the U-Net deep learning model. GISci Remote Sens. https://doi.org/10.1080/15481603.2023.2182057

Brath A, Casagli N, Marani M, Mercogliano P, Motta R (2023) Rapporto della Commissione tecnico-scientifica istituita con deliberazione della Giunta Regionale n. 984/2023 e determinazione dirigenziale 14641/2023, al fine di analizzare gli eventi meteorologici estremi del mese di maggio 2023. Regione Emilia-Romagna, Technical Report, p 147. https://www.regione.emilia-romagna.it/alluvione/rapporto-della-commissione-tecnico-scientifica. Access 7 Apr 2025

Burgess DW, Lewis P, Muller JP (1995) Topographic effects in AVHRR NDVI data. Remote Sens Environ 54(3):223–232. https://doi.org/10.1016/0034-4257(95)00155-7

Chandra N, Sawant S, Vaidya H (2023) An efficient U-Net model for improved landslide detection from satellite images. PFG J Photogramm Remote Sens Geoinf Sci 91:13–28. https://doi.org/10.1007/s41064-023-00232-4

Chen H, He Y, Zhang L et al (2023) A landslide extraction method of channel attention mechanism U-Net network based on Sentinel-2A remote sensing images. Int J Digit Earth 16:552–577. https://doi.org/10.1080/17538947.2023.2177359

Chung CF, Fabbri AG (2005) On mineral potential maps and how to make them useful. In: Cheng Q, Bonham-Carter G (eds). GIS and spatial analysis. Proceedings of IAMG '05, The annual conference of the international association of mathematical geology, vol 1. Geomatics Research Laboratory, Toronto, Canada, pp 533–538

Civit-Masot J, Luna-Perejón F, Corral JMR, Domínguez-Morales M, Morgado-Estévez A, Civit A (2021) A study on the use of Edge TPUs for eye fundus image segmentation, Engineering Applications of Artificial Intelligence, 104. https://doi.org/10.1016/j.engappai.2021.104384

Comert R, Avdan U, Gorum T, Nefeslioglu HA (2019) Mapping of shallow landslides with object-based image analysis from unmanned aerial vehicle data. Eng Geol 260:105264

Crozier MJ (2005) Multiple-occurrence regional landslide events in New Zealand: hazard management issues. Landslides 2:247–256. https://doi.org/10.1007/s10346-005-0019-7

Ferrario MF, Livio F (2023) Rapid mapping of landslides induced by heavy rainfall in the Emilia-Romagna (Italy) region in May 2023. Remote Sens 16:122. https://doi.org/10.3390/rs16010122

Fiorucci F, Ardizzone F, Mondini AC, Viero A, Guzzetti F (2019) Visual interpretation of stereoscopic NDVI satellite images to map rainfall induced landslides. Landslides 16:165–174

Ghorbanzadeh O, Didehban K, Rasouli H et al (2020) An application of Sentinel-1, Sentinel-2, and GNSS data for landslide susceptibility mapping. ISPRS Int J Geo-Inf 9:561. https://doi.org/10.3390/ijgi9100561

Ghorbanzadeh O, Crivellari A, Ghamisi P et al (2021) A comprehensive transferability evaluation of U-Net and ResU-Net for landslide detection from Sentinel-2 data (case study areas from Taiwan, China, and Japan). Sci Rep 11:14629. https://doi.org/10.1038/s41598-021-94190-9

Ghorbanzadeh O, Xu Y, Ghamisi P, Kopp M, Kreil D (2022) Landslide4Sense: reference benchmark data and deep learning models for landslide detection. IEEE Trans Geosci Remote Sens 60:1–17. https://doi.org/10.1109/TGRS.2022.3215209

Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R (2017) Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens Environ 202:18–27. https://doi.org/10.1016/j.rse.2017.06.031

Guzzetti F, Mondini AC, Cardinali M et al (2012) Landslide inventory maps: new tools for an old problem. Earth-Sci Rev 112:42–66. https://doi.org/10.1016/j.earscirev.2012.02.001

He R, Zhang W, Dou J, Jiang N, Xiao H, Zhou J (2024) Application of artificial intelligence in three aspects of landslide risk assessment: a comprehensive review. Rock Mechanics Bulletin 3:4. https://doi.org/10.1016/j.rockmb.2024.100144

Holbling D, Eisank C, Albrecht F, Vecchiotti F, Friedl B, Weinke E, Kociu A (2017) Comparing manual and semi-automated landslide mapping based on optical satellite images from different sensors. Geosci. https://doi.org/10.3390/geosciences7020037

Iakubovskii P (2019) Segmentation models. GitHub repository. https://github.com/qubvel/segmentation_models. Accessed 5 Dec 2025

Joubert-Boitat I, Wania A, Dalmasso S (2020) Manual for CEMS-rapid mapping products. Technical Report, EUR 30370 EN, Publications Office of the European Union, Ispra. https://doi.org/10.2760/29876

Khalili MA, Voosoghi B, Calcaterra D, Kouchakkapourchali A, Di Muro C, Madadi S, Tufano R, Di Martire D (2024) Advanced machine learning strategies for landslide detection," IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, Athens, Greece: 1755–1759. https://doi.org/10.1109/IGARSS53475.2024.10641104

Kim H-J, Lee D (2020) Image denoising with conditional generative adversarial networks (CGAN) in low dose chest images. Nucl Instrum Methods Phys Res Sect A. https://doi.org/10.1016/j.nima.2019.02.041

Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak J, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88. https://doi.org/10.1016/j.media.2017.07.005

Lu P, Qin Y, Li Z, Mondini AC, Casagli N (2019) Landslide mapping from multi-sensor data through improved change detection-based Markov random field. Remote Sens Environ. https://doi.org/10.1016/j.rse.2019.111235
Lu W, Hu Y, Zhang Z (2023) A dual-encoder U-Net for landslide detection using Sentinel-2 and DEM data. Landslides 20:1975–1987. https://doi.org/10.1007/s10346-023-02089-5
Malamud BD, Turcotte DL, Guzzetti F, Reichenbach P (2004) Landslides, earthquakes, and erosion. Earth Planet Sci Lett 229(1–2):45–59. https://doi.org/10.1016/j.epsl.2004.10.018
Martinez SN, Schaefer LN, Allstadt KE, Thompson EM (2021) Evaluation of remote mapping techniques for earthquake-triggered landslide inventories in an urban subarctic environment: a case study of the 2018 Anchorage, Alaska Earthquake. Front Earth Sci. https://doi.org/10.3389/feart.2021.673137
McColl ST, Cook SJ (2024) A universal size classification system for landslides. Landslides 21:111–120. https://doi.org/10.1007/s10346-023-02131-6
Meena SR, Soares LP, Grohmann CH et al (2022) Landslide detection in the Himalayas using machine learning algorithms and U-Net. Landslides 19:1209–1229. https://doi.org/10.1007/s10346-022-01861-3
Milledge DG, Bellugi DG, Watt J, Densmore AL (2022) Automated determination of landslide locations after large trigger events: advantages and disadvantages compared to manual mapping. Nat Hazards Earth Syst Sci 22(2):481–508
Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D (2022) Image segmentation using deep learning: a survey. IEEE Trans Pattern Anal Mach Intell 44(7):3523–3542. https://doi.org/10.1109/TPAMI.2021.3059968
Mondini AC, Guzzetti F, Reichenbach P et al (2011) Semi-automatic recognition and mapping of rainfall induced shallow landslides using optical satellite images. Remote Sens Environ 115:1743–1757. https://doi.org/10.1016/j.rse.2011.03.006
Mondini AC, Chang KT, Chiang SH, Schlögel R, Notarnicola C, Saito H (2017) Automatic mapping of event landslides at basin scale in Taiwan using a Montecarlo approach and synthetic land cover fingerprints. Int J Appl Earth Obs Geoinf 63:112–121
Nava L, Bhuyan K, Meena SR et al (2022) Rapid mapping of landslides on SAR data by attention U-Net. Remote Sens. https://doi.org/10.3390/rs14061449
Nichol J, Wong M (2005) Detection and interpretation of landslides using satellite images. Land Degrad Dev 16:243–255. https://doi.org/10.1002/ldr.648
Notti D, Cignetti M, Godone D, Giordan D (2023) Semi-automatic mapping of shallow landslides using free Sentinel-2 images and Google Earth Engine. Nat Hazards Earth Syst Sci 23:2625–2648. https://doi.org/10.5194/nhess-23-2625-2023
Novellino A, Pennington C, Leeming K et al (2024) Mapping landslides from space: a review. Landslides. https://doi.org/10.1007/s10346-024-02215-x
Prakash N, Manconi A, Loew S (2020) Mapping landslides on EO data: performance of deep learning models vs. traditional machine learning models. Remote Sens 12:346. https://doi.org/10.3390/rs12030346
Prakash N, Manconi A, Loew S (2021) A new strategy to map landslides with a generalized convolutional neural network. Sci Rep 11:9722. https://doi.org/10.1038/s41598-021-89015-8
Qi W, Wei M, Yang W et al (2020) Automatic mapping of landslides by the ResU-Net. Remote Sens. https://doi.org/10.3390/rs12152487
Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. Lecture notes in computer science, Springer, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
Scheip CM, Wegmann KW (2021) Hazmapper: a global open-source natural hazard mapping application in Google Earth Engine. Nat Hazards Earth Syst Sci 21(5):1495–1511. https://doi.org/10.5194/nhess-21-1495-2021
Shafapourtehrany M, Rezaie F, Jun C, Heggy E, Bateni SM, Panahi M, Özener H, Shabani F, Moeini H (2023) Mapping post-earthquake landslide susceptibility using U-Net, VGG-16, VGG-19, and metaheuristic algorithms. Remote Sens 15:4501. https://doi.org/10.3390/rs15184501
Shahabi H, Rahimzad M, Ghorbanzadeh O, Piralilou ST, Blaschke T, Homayouni S, Ghamisi P (2022) Rapid mapping of landslides from Sentinel-2 data using unsupervised deep learning. 2022 IEEE Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), Istanbul, Turkey, pp 17–20. https://doi.org/10.1109/M2GARSS52314.2022.9840273
Smith PR (1981) Bilinear interpolation of digital images. Ultramicroscopy 6(1):201–204. https://doi.org/10.1016/S0304-3991(81)80199-4
Soares LP, Dias HC, Grohmann CH (2020) Landslide segmentation with U-Net: evaluating different sampling methods and patch sizes. arXiv. https://doi.org/10.48550/arXiv.2007.06672
Stumppf A, Kerle N (2010) Object-oriented mapping of landslides using random forests. Remote Sens Environ 115(10):2564–2577. https://doi.org/10.1016/j.rse.2011.05.013
Su Z, Chow JK, Tan PS et al (2021) Deep convolutional neural network–based pixel-wise landslide inventory mapping. Landslides 18:1421–1443. https://doi.org/10.1007/s10346-020-01557-6
Tanoli JI, Ningsheng C, Regmi AD, Jun L (2017) Spatial distribution analysis and susceptibility mapping of landslides triggered before and after Mw7.8 Gorkha earthquake along Upper Bhote Koshi, Nepal. Arab J Geosci 10:277. https://doi.org/10.1007/s12517-017-3026-9
Tehrani FS, Calvello M, Liu Z (2022) Machine learning and landslide studies: recent advances and applications. Nat Hazards 114:1197–1245. https://doi.org/10.1007/s11069-022-05423-7
Tharwat A (2018) Classification assessment methods. Appl Comput Inform 17:168–192. https://doi.org/10.1016/j.aci.2018.08.003
Trigila A, Iadanza C, Spizzichino D (2008) IFFI project(Italian landslide inventory) and risk assessment. Proceedings of the first world landslide forum, pp 18–21
Tucker CJ (1979) Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens Environ 8(2):127–150. https://doi.org/10.1016/0034-4257(79)90013-0
Wang W (2021) An improved denoising model for convolutional neural network. J Phys Conference Series, 1982, 2021 2nd International Conference on Artificial Intelligence and Information Systems (ICAIIS 2021) 28–30 May 2021, Chongqing, China. https://doi.org/10.1088/1742-6596/1982/1/012169
Wang Y, Yang L, Liu X (2024) An improved semantic segmentation algorithm for high-resolution remote sensing images based on DeepLabv3+. Sci Rep 14:9716. https://doi.org/10.1038/s41598-024-60375-1
Williams JG, Rosser NJ, Kincey ME et al (2018) Satellite-based emergency mapping using optical imagery: experience and reflections from the 2015 Nepal earthquakes. Nat Hazards Earth Syst Sci 18:185–205. https://doi.org/10.5194/nhess-18-185-2018
Xu T, Liu S, Xu L, Chen Y, Jia Z, Xu Z, Nielson J (2015) Temporal upscaling and reconstruction of thermal remotely sensed instantaneous evapotranspiration. Remote Sens 7(3):3400–3425. https://doi.org/10.3390/rs70303400
Yang W, Wang Y, Sun S et al (2019) Using Sentinel-2 time series to detect slope movement before the Jinsha River landslide. Landslides 16:1313–1324. https://doi.org/10.1007/s10346-019-01178-8
Yang X, Zuo X, Xie W, Li Y, Guo S, Zhang H (2022) A correction method of NDVI topographic shadow effect for rugged terrain. IEEE J Sel Top Appl Earth Obs Remote Sens. https://doi.org/10.1109/JSTARS.2022.3193419
Zamir SW, Arora A, Khan SH, Hayat M, Khan FS, Yang M (2022) Restormer: Efficient transformer for high-resolution image restoration. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), LA, 19–24 USA, pp 5728-5739

Matteo Berti (✉) · Giuseppe Ciccarese · Nicola Dal Seno · Elena Ioriatti · Rodolfo Rani · Alessandro Zuccarini
Department of Biological, Geological, and Environmental Sciences, University of Bologna, Bologna, Italy
Email: matteo.berti@unibo.it

Marco Pizziolo · Michele Scaroni · Mauro Generali
Regione Emilia-Romagna, Area Geologia, Suoli E Sismica, Bologna, Italy

Stefano Olivucci · Giampiero Gozza · Paolo Formicola
Regione Emilia-Romagna, Settore Innovazione Digitale, Dati, Tecnologia e Polo Archivistico, , Bologna, Italy

Vincenzo Critelli · Marco Mulas · Melissa Tondo · Francesco Lelli · Cecilia Fabbiani · Francesco Ronchetti · Alessandro Corsini
Department of Chemical and Geological Sciences, University of Modena and Reggio-Emilia, Modena, Italy

Tommaso Simonelli
Autorità Di Bacino Distrettuale del Fiume Po, Parma, Italy