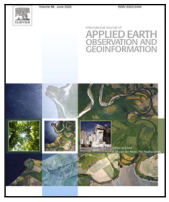
International Journal of Applied Earth Observation and Geoinformation 129 (2024) 103795

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



A framework for automated landslide dating utilizing SAR-Derived Parameters Time-Series, An Enhanced Transformer Model, and Dynamic Thresholding

Wandi Wang [a,b,*], Mahdi Motagh [a,b], Zhuge Xia [a,b,f], Simon Plank [d], Zhe Li [c], Aiym Orynbaikyzy [d], Chao Zhou [a,e], Sigrid Roessner [a]

[a] GFZ German Research Centre for Geosciences, Department of Geodesy, Section of Remote Sensing and Geoinformatics, Telegrafenberg, Potsdam, 14473, Germany
[b] Institute of Photogrammetry and Geoinformation, Leibniz University Hannover, Hannover, 30167, Germany
[c] Department of Electronic and Information Engineering, Hong Kong Polytechnic University, 430079, Hong Kong Special Administrative Region
[d] German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Wessling, 82234, Germany
[e] School of Geography and Information Engineering, China University of Geosciences, Wuhan, 430078, China
[f] Currently at the School of Earth Sciences and Engineering, Hohai University, Nanjing, 210098, China

ARTICLE INFO

Keywords:
Landslide
SAR
Anomaly detection
Deep-learning

ABSTRACT

Determining the timing of landslide occurrence is crucial for establishing an accurate, comprehensive and systematic landslide inventory while assessing the potential for reducing landslide risk. Unfortunately, many existing landslide inventories lack temporal information such as the precise time of landslide events. Optical and Synthetic Aperture Radar (SAR) sensors are the most commonly used remote sensing technologies for landslide detection. Unlike optical sensors, SAR sensors are not affected by cloudy conditions and provide valuable imagery regardless of sunlight availability. Therefore, SAR-derived parameters, i.e., SAR amplitude, interferometric coherence, and polarimetric features (alpha and entropy), offer a higher temporal resolution for detecting landslide occurrence times compared to optical data. Despite the advantages, there is currently no universally accepted automatic method for determining the time of landslide events using SAR data. This is due to the lack of anomaly labels and the high time-series volatility in detecting landslide occurrence times. Despite advances in deep-learning methods for anomaly detection in time-series, only a few of them can address these challenges in our case. In this paper, we propose an unsupervised multivariate transformed-based deep-learning model to automatically and efficiently estimate landslide occurrence times using multivariate SAR-derived parameters time-series analysis. The designed gated relative position can increase robustness and temporal context information, by learning global temporal trends in the time-series. Subsequently, the time-series of the anomaly score derived from the proposed Transformer model is analyzed using an adaptive thresholding strategy to dynamically and automatically mark anomalies related to the landslide occurrence. Our research focuses on collapsed landslides characterized by dramatic changes in ground surface topography, with a particular attention for the need of a prior knowledge about landslide boundaries. We assess the performance of the proposed methodology for several collapsed landslides including the July 21, 2020 Shaziba and 23 July, 2019 Shuicheng landslides in China, March 19, 2019 Takht landslide in Iran, June 15, 2018 Jalgyz-Jangak and May 25, 2018 Kugart landslides in Kyrgyzstan, July 7, 2018 Hitardalur landslide in Iceland, and January 25, 2019 Brumadinho landslide in Brazil. In comparison to commonly used neural networks like the LSTM algorithm, our proposed framework leads to a more accurate estimate for the time of landslide failure using time-series of SAR-derived parameters. Furthermore, our results suggest the great potential of SAR data to narrow the time period detected from optical data when used in conjunction with them.

* Corresponding author at: GFZ German Research Centre for Geosciences, Department of Geodesy, Section of Remote Sensing and Geoinformatics, Telegrafenberg, Potsdam, 14473, Germany.
E-mail address: wandi@gfz-potsdam.de (W. Wang).

https://doi.org/10.1016/j.jag.2024.103795
Received 23 September 2023; Received in revised form 11 March 2024; Accepted 24 March 2024
Available online 1 April 2024
1569-8432/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Landslides are among the most devastating natural disasters on Earth (Brabb, 1991; Kilburn and Petley, 2003), causing massive loss of life and damages worldwide. However, there are currently no particularly effective defense techniques to protect vulnerable communities from landslide hazards, except for evacuation warnings based on spatial and temporal risk assessments (Kilburn and Petley, 2003). High-quality landslide inventories, which record among others location, magnitude, date, and potential triggering factors (Guzzetti et al., 2012) are crucial for the landslide risk assessment (Froude and Petley, 2018), and early warning (Van Westen et al., 2008; Pardeshi et al., 2013). However, the majority of existing landslide inventories primarily include point locations, triggers and scales, but do not include the exact date of landslide occurrence, a shortcoming that contributes to the uncertainty of susceptibility and risk assessment (Van Westen et al., 2008). As a result, it is a challenge to apply the incomplete landslide inventories to have a proper evaluation of landslide risk at local and regional scales (Zhou et al., 2022b; Du et al., 2020)

The failure time of landslides can be detected in various ways including continuous GNSS, field investigation, airborne and spaceborne remote sensing (Colesanti and Wasowski, 2006; Delacourt et al., 2007; Zhou et al., 2022a). MT-InSAR (Multi-temporal Interferometric Synthetic Aperture Radar) technology has demonstrated its capability in detecting slow-moving landslides (Lu et al., 2019a) and identifying precursors before the occurrence of catastrophic slope failures. However, for detecting catastrophic failures within our specified scope in this study, the use of this technology is impractical due to coherence loss (Crosetto et al., 2016). Long time-series of satellite observations acquired by a variety of remote sensing techniques, such as optical and synthetic aperture radar (SAR), provide opportunities for landslide occurrence time detection. For example, scholars exploited time-series of changes in NDVI (Normalized Difference Vegetation Index) derived from optical imagery to automatically determine a time window for landslide occurrence (Behling et al., 2014, 2016; Fu et al., 2023). However, results from this approach suffer from large uncertainty in determining the exact time period of landslide occurrence. Landslide-prone areas are usually exposed to rainfall and often covered by clouds during catastrophic failures while optical data are dependent on daylight and vulnerable to cloud cover. Therefore cloud-free images may not be easily available at the times needed to precisely detect the time of catastrophic failure (Plank et al., 2016; Fu et al., 2023).

Due to their cloud penetration and all-weather operational capability, the use of Synthetic Aperture Radar (SAR) techniques offers new opportunities to investigate instability processes related to landslides (Motagh et al., 2013; Xia et al., 2023; Wang et al., 2023; Xia et al., 2022) and for a more precise temporal characterization of landslide occurrence on a regional scale (Mondini et al., 2021). SAR, as an active remote sensing system, alternatively sends electromagnetic microwaves to ground targets and then receives corresponding backscattering echoes from the Earth. Time series analysis of SAR-derived parameters such as alpha, entropy, amplitude, and interferometric coherence can be utilized to detect any temporal trends or anomalies in backscattering related to landslide failure (Jacquemart and Tiampo, 2021; Jung et al., 2016; Mondini et al., 2019; Plank et al., 2016). For example, interferometric coherence is a measure of the similarity between two SAR images. Except for the influence from long spatial/temporal baselines, a reduction in interferometric coherence generally suggests that the scattering properties of the ground target have changed (Zebker et al., 1992; Jacquemart and Tiampo, 2021). So, this information can be exploited to identify the landslide in vegetated areas where the loss of coherence is more significant after the failure due to loss of vegetation (Jung et al., 2016). Similarly, the difference between pre and post-SAR amplitude images, indicating that the ground target and its scattering properties have changed, can be exploited for landslide detection (Mondini et al., 2019). When SAR systems with dual or full polarimetry channels are available, polarimetric features such as alpha and entropy, can be exploited to analyze changes in backscattering components (e.g., surface, volume, and double-bounce) and detect the areal extent of landslides (Plank et al., 2016).

Deijns et al. (2022) employed a binary segmentation change detection approach to identify the most significant change point in the time-series of amplitude and coherence and took this point for landslide dating. However, other than landslides, a number of unavoidable factors can disrupt and alter the surface backscattering mechanism in a SAR imagery, resulting in significant changes in the time-series of SAR parameters and false detection. These factors include changes in soil moisture, vegetation, rainfall, snowfall, anthropogenic activities, geometric layover and shadow (Lee et al., 1994; Czuchlewski et al., 2003; Oliver and Quegan, 2004; Mondini et al., 2021). The combined time-series of all SAR-derived parameters that are able to detect landslides may help tackle some of the problems related to false alarms in landslide detection arising from a single parameter time-series analysis (Czuchlewski et al., 2003; Shimada et al., 2014; Niu et al., 2021). For example, alpha and amplitude are sensitive to roughness and dielectric properties (Cloude and Pottier, 1997; Hajnsek et al., 2003; Verhoest et al., 2008; Zhu et al., 2019), while entropy and coherence are more sensitive to contribution ratios of various scatterings (Cloude and Pottier, 1997) and vegetation (Jin et al., 2014; Villarroya-Carpio et al., 2022), respectively.

This study explores the use of a deep-learning algorithm to accurately detect the landslide occurrence time using time-series of SAR parameters derived from dual-pol Sentinel-1 data. Here we mainly focus on landslides that have occurred in mountainous areas covered with vegetation. Our hypothesis is that such slopes are dominated by volume scattering before the failure due to the presence of vegetation, whereas the scattering mechanism changes to surface scattering as bare soil is exposed after the failure (Shibayama et al., 2015). To be noticed is that subsets of the landslide body are also sufficient for this assumption. In this case, SAR parameters characterizing various scattering mechanisms, such as interferometric coherence, polarimetric alpha, entropy, and amplitude from SAR data, can be exploited for landslide detection (Yonezawa et al., 2012; Uemoto et al., 2019; Burrows et al., 2020; Jung and Yun, 2020).

The anomalies in the time-series of SAR data due to landslides can be detected using various methods including statistical methods (Ramaswamy et al., 2000), probabilistic models (Shyu et al., 2003) , linear models (Tang et al., 2002), proximity-based models (He et al., 2003), and Artificial Neural Networks (Malhotra et al., 2015). Owing to the increasing number of SAR sensors and advances in computing resources, a trend in using machine learning and deep learning methods has emerged for SAR and InSAR time-series analysis. Here we introduced a gated relative position bias to the Transformer model (Vaswani et al., 2017), applied to detect outliers related to landslides in time-series of multi-band (coherence, alpha, entropy and amplitude) SAR parameters from Sentinel-1 data. In our improved model, multivariate inputs are processed globally rather than locally like step-by-step in chronological order to overcome the “forget” issue in popular methods like Long Short-Term Memory (LSTM) networks. Moreover, the combination of multi-head attention and a gated relative positional embedding can better capture the temporal relationship between data in the time-series. In addition, the weights for each of the four SAR parameters are dynamically adjusted using the attention weights matrix derived from the Transformer model, so that greater weights are given to those parameters, reflecting the evolution of the landslide surface.

2. Study areas and data

2.1. Study areas

In this study, seven large landslide failures were selected including the Takht landslide in Iran, Jalgyz-Jangak and Kugart landslides in

Table 1
Details information of study cases.

Name	Country	Location	Date (dd/mm/yyyy)	Triggers	Type
Shaziba	China	E109°29′85″ N30°36′05″	21/07/2020	rainfall	deep-seated landslide
Shuicehng	China	E104°40′24″ N26°15′27″	23/07/2019	rainfall	deep-seated landslide
Takht	Iran	E55°26′40″ N30°07′02″	19/03/2019	rainfall	shallow landslide
Jalgyz-Jangak	Kyrgyzstan	E41°04′11″ N73°12′05″	15/06/2018	rainfall	shallow landslide
Kugart	Kyrgyzstan	E41°05′11″ N73°25′44″	25/05/2018	rainfall	deep-seated landslide
Hitardalur	Iceland	W22°06′21″ N64°47′02″	07/07/2018	rainfall	deep-seated landslide
Brumadinho	Brazil	W44°05′00″ N20°04′02″	25/01/2019	rainfall dam collapse	deep-seated landslide

Table 2
Details of the Sentinel–1 data used.

Case	Data type	Orbit	Time intervals (dd/mm/yyyy)	Number of images
Shaziba	SLC	Ascending	29/03/2018-	219
	GRD	-	14/11/2021	-
Shuicheng	SLC	Ascending	12/02/2017-	165
	GRD	-	28/07/2022	-
Takht	SLC	Ascending	15/11/2016-	164
	GRD	Descending	23/07/2022	164
Jalgyz-Jangak	SLC	Ascending	27/02/2015-	161
	GRD	Descending	1/08/2022	161
Kugart	SLC	Ascending	27/02/2015-	161
	GRD	Descending	1/08/2022	161
Hitardalur	SLC	Ascending	27/02/2018-	363
	GRD	Descending	22/12/2021	363
Brumadinho	SLC	Ascending	01/05/2015-	193
	GRD	Descending	13/12/2021	193

Kyrgyzstan, Shaziba and Shuicheng landslides in China, the Hitardalur landslide in Iceland, and the Brumadinho landslide in Brazil. Table 1 lists all the information for the location and date of failure of these landslides extracted from previous studies and news reports (Table 1). In March 2019, the Takht landslide occurred in a forested region in northern Iran, 1 km from the Takht village, with a total affected area of 0.2 km^2. The Shuicheng landslide happened on 23 July 2019 in the Pingdi village of Shuicheng County in Guizhou Province, China. The sudden, high-speed debris movement caused severe damage to the village infrastructure with more than 20 houses buried and about 42 fatalities. The Shaziba landslide, which occurred on 21 July 2020 in Mazhe County, Hubei Province, China, destroyed more than 60 houses and some 8000 people had to be evacuated amid fears of further landslides. The Brumadinho landslide occurred on 25 January 2019 at the Córrego do Feijão iron ore mine, 9 km east of Brumadinho, Minas Gerais in Brazil causing 270 fatalities. The Hitardalur landslide occurred in Fagraskógarfjall mountain in Hítardalur Valley, Iceland, and is considered among the largest landslides in historical time in West Iceland. The Kugart landslide is located 1 km south of the Kyzylsu River in the Dshalal Abad region of Southern Kyrgyzstan and occurred as a reactivation of an older landslide on 25 May 2018. There are no reports claiming casualties for this event. The Jalgyz-Jangak landslide occurred on 15 June 2018 and is located in northwest-southeast direction around 1 km away from Jalgyz-Jangak village in the Dshalal-Abad region of Southern Kyrgyzstan. The runout zone of the landslide covered a road, fortunately, there were no casualties.

2.2. Data

In this study, time series of Sentinel-1 data are used which were obtained for each landslide case in both ground range detected (GRD) and single look complex (SLC) formats. Sentinel-1 in SLC format was used to calculate time-series of coherence, alpha, and entropy parameters, while Sentinel-1 GRD data were used to derive time-series of amplitude. Table 2 shows the details of the used Sentinel-1 data.

3. Methodology

The overall methodology developed in this study is illustrated in Fig. 1. A pre-processing step (I) is performed first to extract SAR parameters from Sentinel-1 data. The results are combined into multi-band time-series (step (II)) and then in the step (III) are used as inputs to an improved Transformer model. This section describes how four parameters derived from SAR data enable the detection of landslides, and thus how the improved model is built and detects anomalies related to landslide occurrence in the time-series of multi-variant SAR-derived parameters. The graphical representation of the methodological steps is depicted in Fig. 1.

3.1. SAR pre-processing

Sentinel–1 pre-processing was carried out automatically using the command line graph processing tool (GPT) of the open-source ESA Sentinel application platform (SNAP) software. Four GPT graphs, including all steps of the SLC and GRD processing, were applied in batch to the time series of data for each landslide using Linux bash scripts. The details of the four parameters are explained in the following.

3.1.1. Amplitude

Amplitude is the strength of the radar signal backscattered to the sensor, which is widely used for surface change detection. Amplitude can be calculated by combining real (Q) and imaginary (I) parts of the complex radar signal as below (Lee et al., 1994).

$$Amplitude = \sqrt{\left(I^2 + Q^2\right)} \quad (1)$$

In this study, σ^0 (sigma nought) transformed from amplitude is selected to detect landslides. The $\sigma°$ from amplitude is defined as below (Lee et al., 1994).

$$\sigma^0 = 10^* \log_{10}\left(DN^2\right) \quad (2)$$

where DN is the image pixel digital number measured in the SAR amplitude image. σ^0 is dependent on the ground surface roughness and dielectric constant (Choker et al., 2017; Santangelo et al., 2022). The higher values usually represent bare soil, while the lower values represent vegetated areas (Santangelo et al., 2022). Landslides can significantly alter the original surface roughness by removing vegetation, deforming ground, and accumulating debris. Therefore, SAR images taken of the same area before and after a landslide failure show meaningful differences in amplitude images (Santangelo et al., 2022).

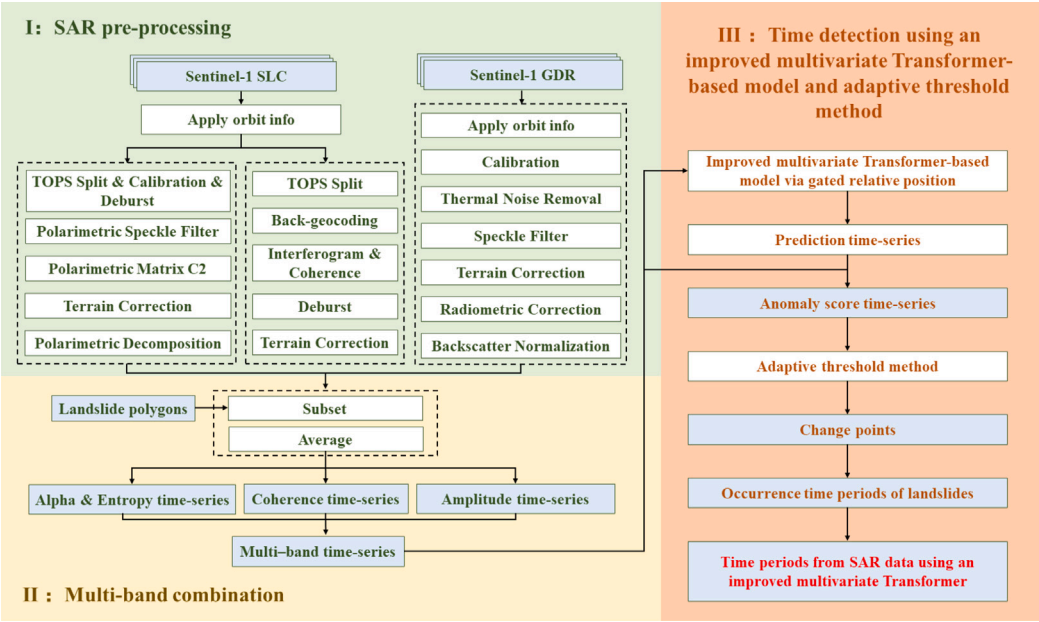


Fig. 1. Methodology framework developed for detection of landslide occurrence time.

As the hypothesis of this study, dominated scattering patterns change from volume scattering to surface scattering after failure. If the exposed soil dries rapidly after failure, landslides will allow volume scattering phenomena due to higher roughness and lower dielectric constant associated with the reduced soil moisture. In this case, the amplitude cannot reliably detect landslides.

3.1.2. Interferometric coherence

Interferometric coherence is a measure of the correlation or similarity between two or more SAR images acquired at different times over the same area (Moreira et al., 2013; Zebker et al., 1992). Interferometric coherence can be applied to identify the landslide areas where the loss of coherence is more significant than before the failure, due to the motion of vegetation elements and dielectric changes in vegetation and ground (Jung et al., 2016). The following formula describes the interferometric coherence estimation (Zebker et al., 1992; Touzi et al., 1999) which has been used in this study:

$$\gamma(x) := \frac{\sum_{i=1}^{N} z_1(i) \cdot z_2^*(i)}{\sqrt{\left(\sum_{i=1}^{N} |z_1(i)|^2\right) \cdot \left(\sum_{i=1}^{N} |z_2(i)|^2\right)}} \tag{3}$$

where N is the size of the square sliding window, x represents the center pixel of the sliding window, z_1 and z_2 denote the complex signals of the first and second image, and $*$ is complex conjugation. The bare ground is often characterized by high coherence values compared to areas covered by vegetation (Plank, 2014). For catastrophic landslides in vegetated areas, the interferometric coherence value initially decreases due to the loss of vegetation caused by the landslide. Subsequently, it increases and maintains a higher value over time, attributed to bare soil. Therefore, the time of the landslide can be detected by identifying changes between the pre- and post-failure interferometric coherence. The challenges, however, are that other factors such as rainfall, snowfall, and anthropogenic activities can cause temporal decorrelation, leading to false positive detection.

3.1.3. H/A decomposition in the dual polarization mode

H/A decomposition in the Dual Polarization mode was proposed by Cloude and Pottier (Cloude and Pottier, 1997). For a dual polarization SAR image, each pixel is represented by a 2*2 coherency matrix T_{dual} , which is nonnegative definite and Hermitian. The eigenvalue decomposition of T_{dual} is defined as:

$$T_{\text{dual}} = \begin{bmatrix} T_{11} & T_{12} \\ T_{12}^* & T_{22} \end{bmatrix} = U \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} U^H \tag{4}$$

The eigenvector matrix U is parameterized as:

$$U = \begin{bmatrix} e^{j\phi_i} \cos\alpha_1 \sin\alpha_1 & e^{j\phi_i} \cos\alpha_2 \sin\alpha_2 \\ e^{j\phi_i} \cos\alpha_1 \sin\alpha_1 & e^{j\phi_i} \cos\alpha_2 \sin\alpha_2 \end{bmatrix} \tag{5}$$

Eigenvectors help describe the different scattering processes and the eigenvalues indicate their relative magnitudes. $H-\alpha$ decomposition is based on the eigenvector and eigenvalue. The polarimetric entropy H and the scattering angle α are defined as:

$$H = \sum_{i=1}^{2} -P_i \log_2 P_i, 0 \le H \le 1 \text{ and } \alpha = \sum_{i=1}^{2} P_i \cos^{-1}\left(|u_{1i}|\right), 0 \le \alpha \le 90 \tag{6}$$

where, $P_i = \lambda / \sum_{j=1}^{2} \lambda_j, i = 1, 2$.

The randomness of scattering is represented by Entropy (H) in the range of 0 to 1, where $H = 0$ represents a single scattering (isotropic scattering) and $H = 1$ indicates a random mixture of scattering mechanisms (depolarizing target). $H(0 < H < 1)$ represents the degree of dominance of one particular scatter. The alpha (α) angle that indicates the average or dominant scattering mechanism is based on the eigenvectors. The lower values of alpha ($\alpha = 0$) represent surface scattering; dipole or volume scattering gives intermediate values ($\alpha = 45$) and higher values of alpha are given by dihedral reflectors (Cloude and Pottier, 1997). Therefore, alpha and entropy can identify the landslide, since the amounts and proportion of various scatterers would alter before and after the failure in the landslide regions.

However, the effectiveness of alpha and entropy may be limited in some cases such as when single scattering dominates before and after failure. In this case, entropy H is incapable of detecting landslide . When the bare soil after failure dries rapidly, it may enhance the surface roughness and reduce the dielectric constant (Bindlish and Barros, 2000). This process could potentially lead to a transition from surface scattering to volume scattering (De Jeu et al., 2008; Morrison and Wagner, 2019). In such cases, alpha proves ineffective in identifying the landslide.

In summary, each SAR-derived parameter can be exploited to detect landslides; however, each faces unique challenges in implementation due to varying factors such as rainfall, snowfall, vegetation, and anthropogenic activities. The combination of all the SAR-derived parameters

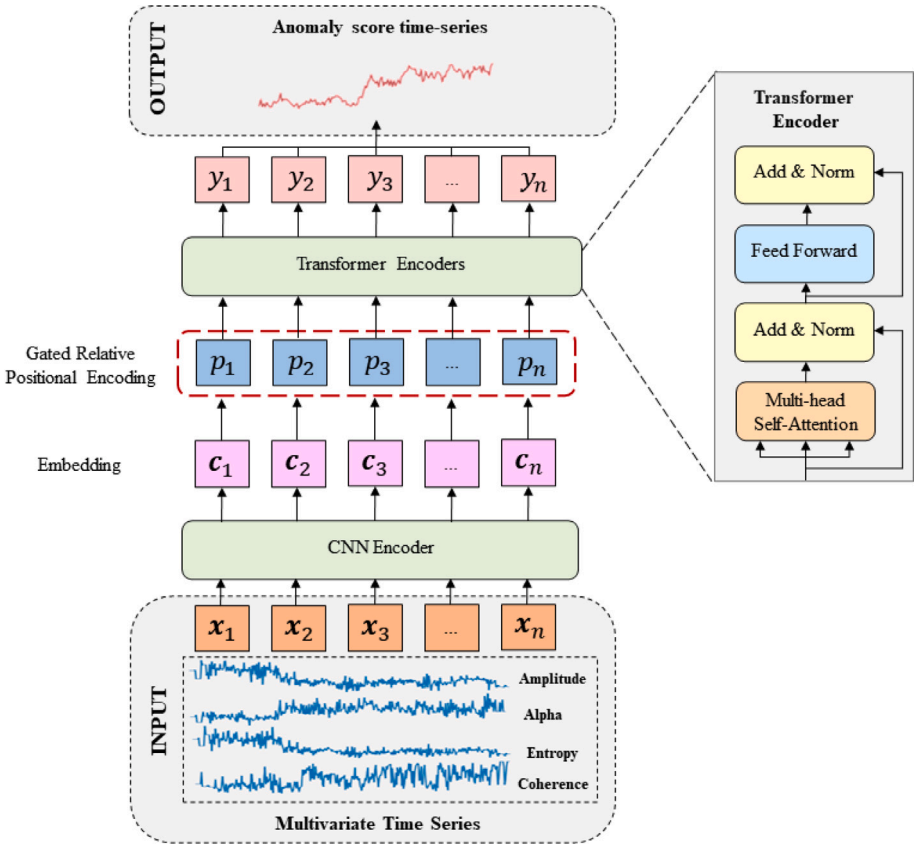


Fig. 2. Model architecture.

offers a comprehensive approach to overcome these challenges. (Lee et al., 1994; Czuchlewski et al., 2003; Oliver and Quegan, 2004; Shimada et al., 2014; Mondini et al., 2021).

3.2. Multi-band combination

Here all SAR parameters that are calculated from Sentinel-1 data are cropped to the known landslide boundary from the existing archives or extracted from the optical images. Advanced methods have been developed for accurately detecting landslide boundaries from remote sensing data (Behling et al., 2014; Lu et al., 2019b). Shaziba, Shuicehng, Takht, Hitardalur and Brumadinho landslide boundaries were manually extracted based on Google Map images. The Jalgyz-Jangak and Kugart's boundaries were acquired from a remote sensing-based landslide inventory map for Southern Kyrgyzstan compiled using RapidEye satellite imagery (Behling et al., 2014). Next, an average value is calculated for all pixels within the landslide boundary. In other words, each SAR parameter within the landslide boundary is ultimately represented by only a single mean value. Using other statistics including median, mode, 3 sigma limits, and quartile provided similar results as the average except for the mode (See Supplementary Material). Then, the time-series of all SAR parameters using the corresponding multi-temporal mean value was generated. Finally, four SAR parameters were combined into a multi-band as the input to our developed Transformer model.

3.3. Multivariate transformer with gated relative position

This section describes the complete architecture of our developed model layer-by-layer. Our model architecture uses the Transformer model as the backbone. The overall model architecture is presented in Fig. 2, which contains a convolutional feature encoder and a Transformer encoder.

3.3.1. Improved transformer model architecture

The input multivariate time-series $\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots, \boldsymbol{x}_n\}$ is a sequence of multivariate SAR parameters (Amplitude, Alpha, Coherence, and Entropy), and n denotes the length of the sequence. A CNN embedding layer is employed to transform this multivariate time-series into a sequence of low-dimensional dense vector $\boldsymbol{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3, \ldots, \boldsymbol{c}_n\}$. The convolutional outputs denote $\boldsymbol{c}_i$, which are then used as input into the Transformer model. The Transformer (Vaswani et al., 2017) is a neural network architecture that has been used in various natural language and vision processing tasks (Li et al., 2023), which mainly consists of multi-head self-attention (MHSA) and position-wise fully connected feed-forward networks. To further increase the temporal context information, a gated relative position bias is introduced to MHSA. MHSA is used to compute the attention weights between each position in the input sequence. The attention weights with a gated relative position bias are computed using the following equation:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}} + \boldsymbol{p}\right)\mathbf{V} \tag{7}$$

Here, **Q**, **K**, and **V** are the query, key, and value matrices, respectively. The $\boldsymbol{p}$ is the gated relative position bias matrix that we designed. The $\sqrt{d_k}$ represents the dimension of the **K** matrix. $softmax$ is the softmax function that normalizes the attention weights. Softmax ensures that the attention weights sum up to 1 and represents a valid probability distribution over the tokens. The **Q** matrix represents the information that the model is seeking or querying from the input sequence. The **K** matrix represents the information about other tokens in the input sequence, i.e., parts that the current model does not want to pay attention to. The **V** matrix holds the actual information or content at each position in the input sequence.

The $\boldsymbol{p}$ is the gated relative position bias added to the attention logits, which is encoded based on the offset between **Q** and **K**. The $\boldsymbol{p}$ matrix

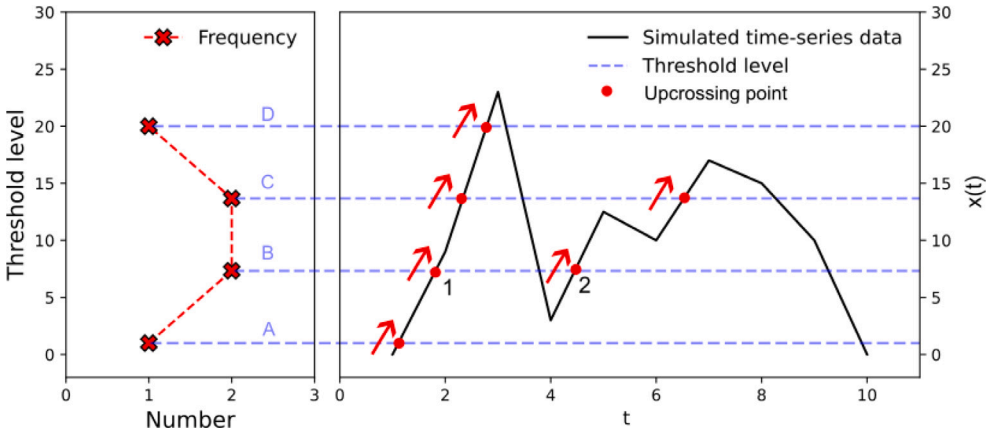


Fig. 3. Schematic of UCA method. The black line represents the simulated time-series data, and the red circles are the up-crossing points. The red crosses are the frequency of up-crossing points for each threshold level.

is computed by the following equation.

$$\boldsymbol{p} = \boldsymbol{d} + h^{(\text{update})}\ \boldsymbol{d} + \left(1 - h^{(\text{update})}\right)\tilde{\boldsymbol{p}} \tag{8}$$

where $h^{(\text{update})}$, $h^{(\text{reset})}$, and $\tilde{\boldsymbol{p}}$ can be calculated as follows.

$$\begin{aligned} h^{(\text{update})}, h^{(\text{reset})} &= \sigma\left(\mathbf{Q}\cdot\boldsymbol{u}\right), \sigma\left(\mathbf{Q}\cdot\boldsymbol{w}\right) \\ \tilde{\boldsymbol{p}} &= w h^{(\text{reset})}\ \boldsymbol{d} \end{aligned} \tag{9}$$

where $\boldsymbol{d}$ is a learnable scalar relative position bias, the vectors u, w are learnable parameters, σ is a sigmoid function, and w is a learnable value. The attention mechanism here can dynamically capture the differences between tokens indirectly through the dot product similarity computation, i.e., deviations. Tokens that are semantically similar or have strong relationships will have higher dot product scores and consequently higher attention weights, allowing the model to focus more on those tokens during processing.

The attention weights matrix output of the MHSA with a gated relative position bias is passed through a position-wise feed-forward network, which consists of two linear transformations with a non-linear ReLU activation function applied in between. Compared to traditional feed-forward neural networks, the Transformer itself introduces more complex dependencies and computations in gradient calculation and weight updates. What is especially important is that the gated relative position bias introduced allows the Transformer model to capture relative position information of time-series and directly contribute to the gradient computation and weight update process. The output $Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \ldots, \boldsymbol{y}_n\}$ from the feed-forward networks carries the refined and enriched information for each position in the input sequence. Our model is composed of a stack of 12 Transformer encoder layers. Here, only the encoder is used exclusively for time series problems because it efficiently captures temporal dependencies and reduces the computational complexity compared to the full model with both encoder and decoder.

The model utilizes 1024-dimensional hidden states and employs 12 attention heads. To prevent a significant increase in parameters such as weights and biases, we uniformly share relative position embeddings across all layers. We employ the Adam optimizer with a tri-stage learning rate schedule. Training is performed on a single GPU, utilizing a batch size of 32 and the LayerDrop rate is set to 0.05. For implementation, we utilized the PyTorch framework running on an NVIDIA 3090 GPU. Besides, we did not manually adjust the model parameters for each of the seven landslides in this study.

3.3.2. Anomaly score

Root Mean Squared Error ($RMSE$) between the true data and the predicted value is commonly used to assess the loss function related to prediction error.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{10}$$

where n is the number of timestamps of the given time-series, $\hat{y}_i$ is the predicted outputs and y_i is the real observation. In this study, the anomaly score defined as the normalized $RMSE$ was considered for anomaly detection. It ranges from 0 to 1, with the following equation.

$$\text{Anomaly score}_i = \frac{RMSE_i - min(RMSE)}{max(RMSE) - min(RMSE)}\quad i = 1, 2, \ldots, n. \tag{11}$$

here, $max(RMSE)$ and $min(RMSE)$ are the maximum and minimum of the $RMSE$, respectively. The higher the anomaly score at a certain point, the more probable that it is an anomaly. If the anomaly score of certain data points in a time-series increases significantly, it is highly probable that these points represent anomalous, indicative of changes related to landslides. Due to the nature of landsliding in vegetated slopes, the anomaly score of the post-failure timestamps is higher than that of the pre-failure timestamps, which can be used in the following section by an anomaly detection procedure for dating the landslide.

3.3.3. A adaptive threshold method to detect change point

An anomaly change point in a time series is defined as an excursion above a threshold. Therefore, if the value of a point in the time-series is greater than a fixed threshold, then this point can be marked as an anomaly. However, there are situations in which a fixed threshold approach is not working properly. For example, when a time-series has multiple change points, multiple thresholds are needed instead of a singular fixed threshold. To solve this problem, we propose an adaptive and dynamic threshold method to detect anomalies in the time-series of the anomaly score calculated in Section 3.3.2, which is based on up-crossing analysis (UCA). Fig. 3 illustrates details of the UCA for a simulated time-series (dark line). Let the time-series be denoted as $x(t)$ with time index t, $t = 1, 2, 3, \ldots, n$ and A, B, C, and D (blue dash lines) being the threshold levels. The intersection points in the up-direction of the time-series $x(t)$ and threshold level A, B, C, and D are defined as the up-crossing point $P(i)$ (red points in Fig. 3). The points in the $x(t)$ near the up-crossing points satisfies the following conditions:

$$x(i-1) < T \text{ and } x(i+1) > T \tag{12}$$

where T represents the threshold. The amount of the up-crossing points is defined as frequency (diagonal red crosses in Fig. 3). For example, the threshold level B intersects the time-series $x(t)$ with two up-crossing points that are noted as 1, and 2, corresponding to a frequency is 2.

Following the above analysis, the value in the time-series corresponding to the local minimum frequency of up-crossing is defined

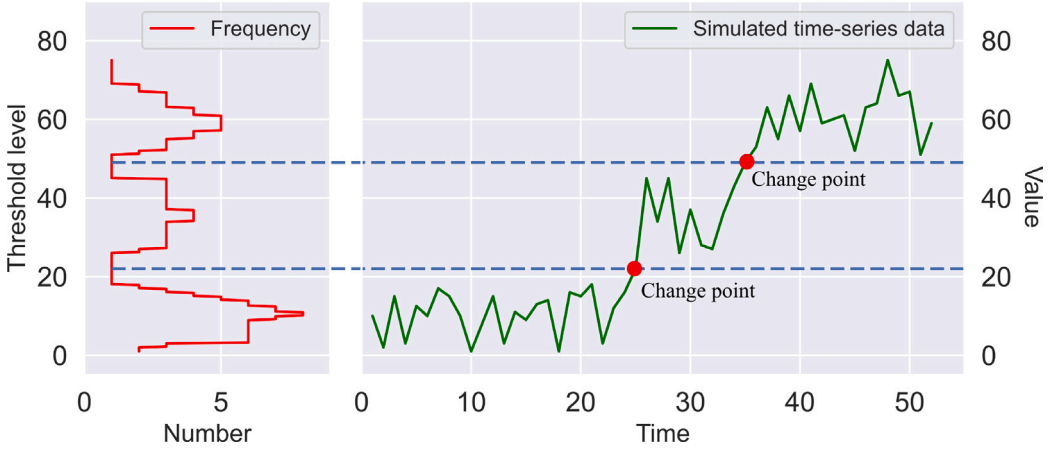


Fig. 4. Change point detection using adaptive threshold method. The green line represents the simulated time-series data, and the red circles are the change points. The red line represents the frequency of up-crossing points for each threshold level. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

as the adaptive threshold, which can be exploited to detect multiple change points. This has been illustrated in Fig. 4 with the simulated time-series that is characterized by multiple change points (green line in Fig. 4). Between the lowest and highest values of the simulated time-series, 200 values are equally sampled to be employed as threshold levels (like blue dash lines in Fig. 4). The frequency of the intersection points of these 200 threshold levels and simulated time-series data is shown in Fig. 4 in red line. From the frequency line (red line in Fig. 4), it is obvious that the frequency domain is characterized by 2 local minima. These two local minima in the frequency domain are corresponding to two change points (red points in Fig. 4) of the simulated time-series. Therefore, the threshold corresponding to the local minimum of the frequency can be adopted as the adaptive and dynamic threshold for detecting change points. Inevitable errors may arise in the process; however, the introduction of additional artificial judgment is not employed to preemptively mitigate these potential errors.

4. Results

4.1. Time-series of SAR-derived parameters

The interferometric coherence, amplitude, entropy and alpha derived from Sentinel-1 SAR data for the seven landslide cases are shown in Fig. 5. The boundaries of the seven landslides are delineated in red in Fig. 5. The Sentinel-2 and Google Earth images before and after each event are also shown as reference. The optical images show that all seven failures resulted in the loss of vegetation cover and the exposure of the underlying soil mass. In Fig. 5, we can see that all four SAR parameters exhibit discernible changes before and after the failure. Thus, their variations can be utilized to delineate the evolution of the landslide body. It is noteworthy that the magnitude of the change is not similar across all parameters. For instance, in the cases of Shuicheng, Shaziba, and Brumadinho landslides, the changes in coherence between pre- and post-failure are significantly greater than those observed in amplitude.

Fig. 6 illustrates the time-series of SAR parameters for the selected landslides in this study. As seen in Fig. 6, all time-series of SAR parameters are characterized by frequent spikes, troughs, as well as local and global changing trends. In the Shuicheng landslide (Fig. 6a), an offset in reaction to the landslide occurrence can be easily recognized in the time-series of all four SAR parameters owing to the significant shifts in value. However, several unexpected offsets also occur during data analysis for the time-series of coherence (in July 2020, July 2021, and June 2022) and amplitude (in June 2021). For the Shaziba landslide (Fig. 6b), time-series analysis of alpha and entropy clearly reveals the failure time due to a significant decline in their values following the landslide occurrence. The time-series of amplitude also shows a visible jump in the time-series before and after the failure. However, several other spikes, peaks, and troughs are also visible in time-series of alpha, entropy, and amplitude which will make automatic detection without a priori knowledge very challenging. Unlike the other three parameters, three clear sharp seasonal peaks in the time-series of coherence in Shaziba can be recognized before the main failure around the same time of year, and these peaks are likely to trigger false alarms during automatic detection. For the Takht landslide shown in Fig. 6c, an increase or decrease in the time-series of SAR parameters can be visually seen during the failure. However, the background noise in the time-series can easily lead to several false detections. In contrast to the Takht landslide, the time-series of coherence for the Jalgyz-Jangak and Kugart landslides (Fig. 6d-e) illustrate only a single jump point, which makes it more suitable for failure detection in this case than alpha and entropy. For time-series of amplitude, alpha and entropy, there is no clear dividing point between pre- and post-failure in Figs. 6d and 6e. Finally, from Fig. 6f-g, we found that the failure times of the Hitardalur and Brumadinho landslides can be best determined from the distinct jump points in the time-series of amplitude, rather than the coherence, alpha and entropy.

4.2. Result of landslide dating

Fig. 7 shows estimated time windows for landslide occurrence using our proposed strategy in this paper. The detailed time window is listed in Table 3. For comparison, the time window estimated using two cloud-free Sentinel-2 data nearest to the real failure time is also listed in Table 3. The time window was estimated using the time difference between the acquisition time of Sentinel-2 images. As seen in Fig. 7, the anomaly score produced by our improved model increased significantly after the failure for all cases analyzed with all post-failure anomaly scores being much higher than those from pre-failure. For example for the Shuicheng landslide, the anomaly score evaluated by our improved Transformer model significantly increased by 86% from 0.21 on July 22, 2019 before the failure to about 0.39 on August 01, 2019 after the failure. In the following, the anomaly score continues to remain high at around 0.8 after the failure. The duration of the detected time period and whether it coincidence with the real date of the landslide occurrence is used as a criterion to assess the accuracy of the time detection.

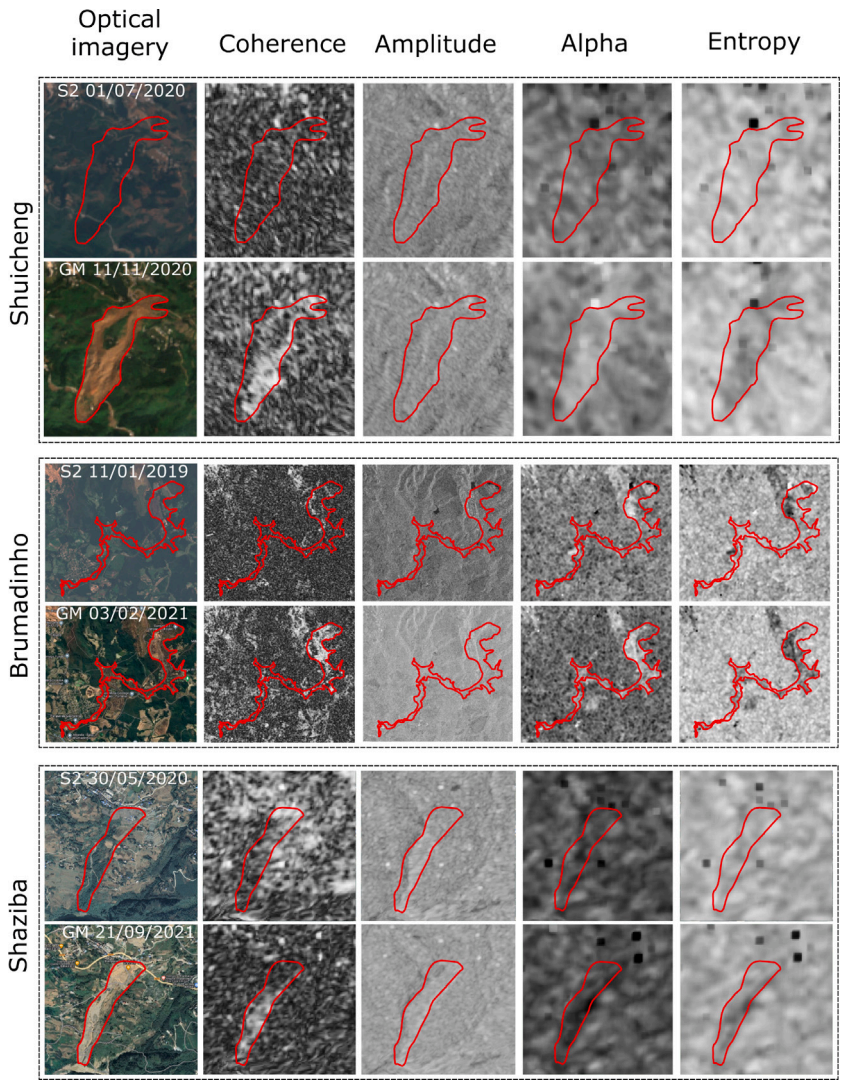


Fig. 5. Interferometric coherence, amplitude, entropy, and alpha images derived from Sentinel-1 before and after the failure of each landslide case. Landslide surfaces shown on Sentinel-2 images and Google Maps before and after the landslide failures, respectively. Google Maps was employed to display the landslide surface after the failure here since its resolution (1 m) is better than that of Sentinel-2 images (10 m). The acquisition times of the optical images are also shown in their corresponding pictures. Red lines indicate the landslide boundary. In each case, the first row shows the optical and SAR parameters images before the failure, while the second row displays these images after the failure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Our method estimates that the Shuicheng landslide occurred between 22.07 and 01.08.2019, narrowing down the time window in which the landslide occurred to 12 days compared to 45 days estimated from Sentinel-2 data. Similar observations for the anomaly score are clearly seen for the other 6 cases, through which our method can significantly narrow the time period down to the sensor revisit period. Only for the Jalgyz-Jangak landslide both optical and SAR data perform equally in identifying the occurrence time period (5 days for optical compared to 6 days for SAR data).

5. Discussion

5.1. Comparison of multi-SAR and single-SAR parameter

In this research, we proposed a novel methodology for the automated detection of the timing of landslides using SAR data. Our approach involves the integration of multiple SAR-derived parameters with an enhanced Transformer model. Analyzing SAR data from 7 cases, we found that the combination of the coherence, amplitude, and polarization parameters (entropy and alpha), has a significant potential for enhancing the precision of landslide dating in vegetated areas. These four SAR-derived parameters can be used for landslide detection because they are able to detect the transformation of the dominant scatterer on the landslide body from volume scattering to surface scattering triggered by the failure. Prior to the failure, the landslide surface is covered with vegetation with the dominating backscattering mechanism being volume scattering. After the failure the landslide surface transforms into bare soil, where surface scattering becomes dominant. However, it is crucial to consider that the alterations in the scattering mechanism can be induced not only by landslides but also by other factors such as precipitation, snowfall, soil moisture, plant growth, anthropogenic effects and so on. This indicates that some changes detected by SAR-derived parameters may not necessarily correspond to landslides but to other disturbances, leading to potential false alarms in automated analysis. Therefore, using time-series of a single SAR parameter does not always work for detecting the occurrence time period of landslides as it may lead to lots of false detection. For example, for the Shuicheng landslide, the jump point that occurred in June 2021 on the time-series of amplitude is a false alarm, as it is related to a continuous heavy rainfall during June 2021 that influenced soil moisture there. Rainwater that accumulates on the landslide body weakens the original dominated surface scattering. Thus

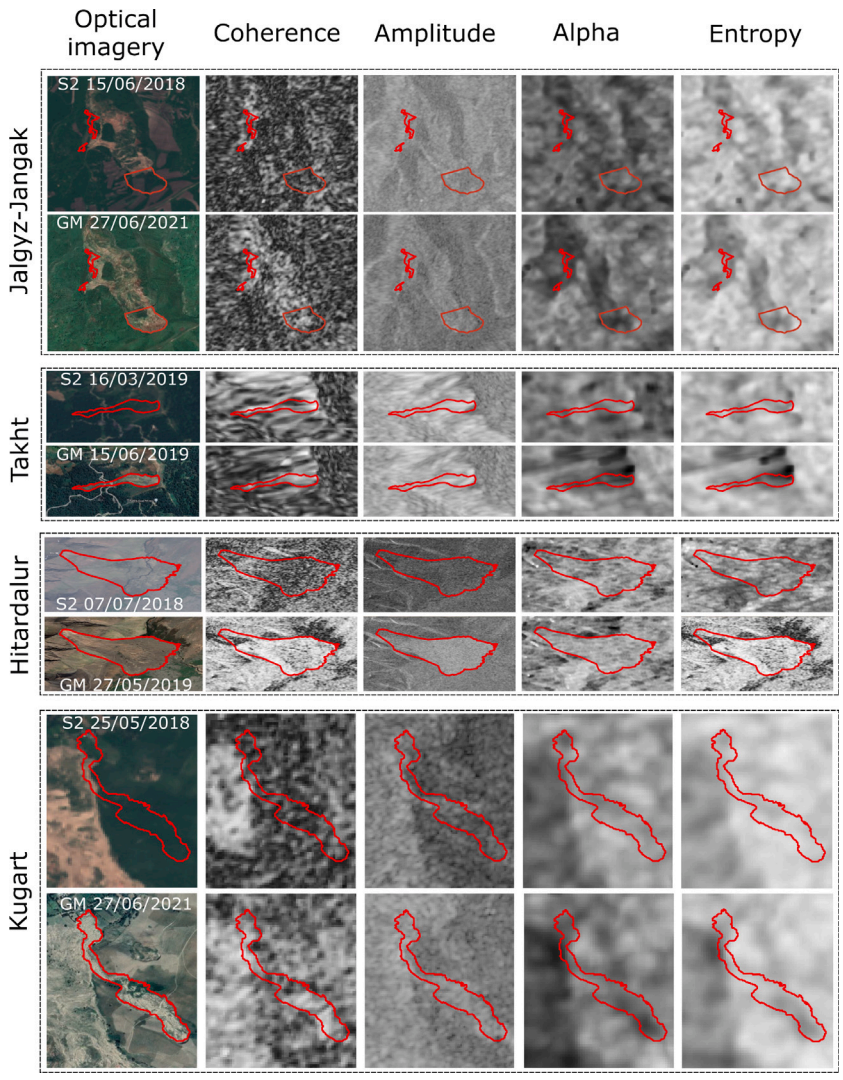


Fig. 5. (*continued*).

Table 3
The real occurrence date and detected time period of seven landslide cases.

Case	Real date	Time period from Sentinel-2 (dd/mm/yyyy)	Timespan from Sentinel-2 (days)	Time window detected by our method (dd/mm/yyyy)	Timespan from our method (days)
Shaziba	21/07/2020	30/05-03/08/2020	65	11/07-23/07/2020	12
Shuicheng	23/07/2019	01/07-15/08/2019	48	22/07-01/08/2019	12
Takht	19/03/2019	16/03-02/04/2019	17	18/03-24/03/2019	6
Jalgyz -Jangak	15/06/2018	13/06-18/06/2018	5	11/06-17/06/2018	6
Kugart	25/05/2018	24/05-28/06/2018	35	24/05-30/05/2018	6
Hitardalur	07/07/2018	20/06-17/07/2018	27	05/07-11/07/2018	6
Brumadinho	25/01/2019	11/01-01/02/2019	21	22/01-28/01/2019	6

the return signal to the radar receiver becomes weaker, resulting in an abrupt reduction in amplitude. In the case of the Shaziba landslide, the proportion of various types of scatterers and dominant scatterers showed minimal change after the failure. As a result, only alpha and entropy proved effective in detecting landslide occurrences. On the contrary, the coherence and amplitude values were not satisfactory for the time detection there, since there were more than two jump points in their time series. Among them, only one jump point corresponds to the landslide failure, whilst the others are related to other factors that affect temporal decorrelation and strength of the backscatter signal. Similar observations were made for other cases investigated in this study. Therefore, integrating four different SAR parameters is necessary to better date landslide occurrence. In the future, full polarization SAR and more polarization parameters are considered to improve the detection capabilities and accuracies of landslides' failure.

5.2. Comparison with the LSTM

In this study, we proposed an unsupervised multivariate Transformer model via gated relative position to automatically date landslides using the time-series of SAR-derived parameters. The effectiveness of the proposed method was validated with seven selected landslide cases. We also attempted to compare its results with those of the

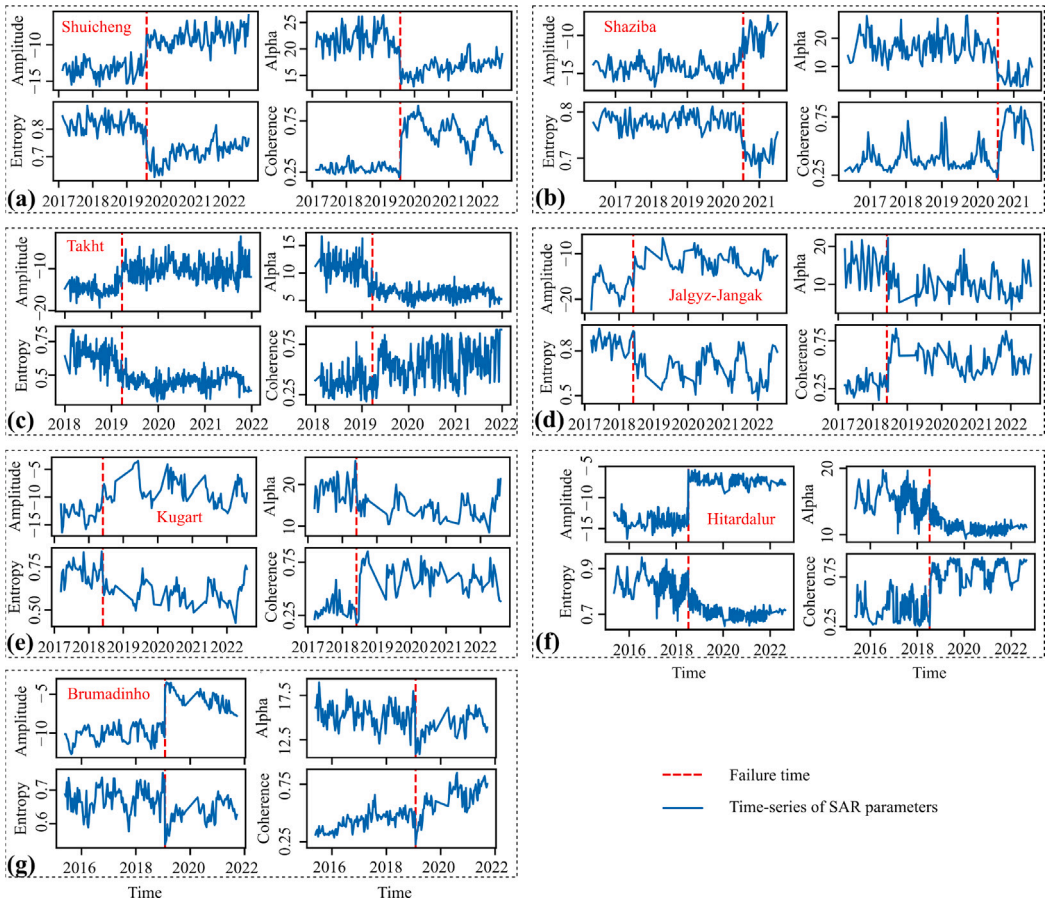


Fig. 6. Time-series of amplitude, alpha, entropy, and coherence for (a) Shuicheng, (b) Shaziba, (c) Takht, (d) Jalgyz-Jangak, (e) Kugart, (f) Hitardalur, and (g) Brumadinho landslide. The red dashed lines represent the real failure times.

most commonly used and state-of-the-art model, i.e., the LSTM model. Compared with the results of detected time using LSTM (Fig. 8), we can see that the failure times are not simply visible on the time-series of anomaly scores for all cases. In contrast to our transform model, the time-series of anomaly scores estimated by the LSTM after the failure are not necessarily higher than those before the failure and are characterized by unexpected waves and several peaks. This makes estimation of occurrence time using LSTM very challenging. For example, although the anomaly score value estimated by LSTM for the Shuicheng before the failure is slightly higher than that of post-failure values by about 0.3 and there was a very low anomaly score (approximately 0.16) near the real failure time, using the LSTM and proposed adaptive threshold method we infer that the landslide possibly occurred either in the period 11/03/2019-22/03/2019 or 25/08/2019-06/09/2019, even not coincidence with the real failure date. Similar to the Shuicheng landslide, the LSTM model failed to detect the correct occurrence period of the other six landslides. From the results of our proposed model (Fig. 7) and the LSTM model, it is obvious that the occurrence period of landslides detected by our method is much closer to the real date than the LSTM model. The possible explanation is that the Transformer with gated relative position bias has the ability to dynamically prioritize the most significant temporal relationships in the time-series and then identify the most impactful anomaly on a global scale. The gated relative position bias layer introduced in the Transformer can provide temporal features that are not available in LSTMs. This layer allows the Transformer to capture global dependencies across the entire sequence and dynamically adjust the importance of these dependencies based on their relative positions. This means that the Transformer can more effectively discern which parts of the sequence are most relevant for a given task, such as identifying anomalies in landslide time series data. This dynamic adjustment is a key difference from LSTM. Although LSTM inherently focuses on local temporal dependencies, its sequential processing nature limits its ability to weigh the importance of different sequence parts. For example, correlations and dependencies between data in the time-series of SAR parameters can be lost in the process of the LSTM model as it computes time-series step-by-step in chronological order, e.g., earlier information farther in the past from the current timestamp is forgotten and information probably is ignored after the current timestamp. Although the LSTM introduces a gate mechanism or uses a bidirectional model to alleviate this problem (Hochreiter and Schmidhuber, 1997; Pascanu et al., 2013; Malhotra et al., 2015, 2016; Nanduri and Sherry, 2016; Guo et al., 2018; Canizo et al., 2019), the temporal relationships are still lost between the data that are far away from each other in time-series. Therefore, this local attention mechanism limits the LSTM model from effectively analyzing time-series. Besides, the Transformer with gated relative position bias can provide a more nuanced understanding of the entire sequence, adapting its focus and attention based on the relative positioning of elements within the sequence. Therefore, the Transformer with gated relative position bias can identify the most impactful anomaly, which could explain why a single, more globally relevant anomalous time point can be detected, as shown in Fig. 7. In contrast, an LSTM, which focuses on local patterns, might detect multiple anomalies that are locally relevant but may not capture the global context as effectively, as shown in Fig. 8.

5.3. Role of dynamic weights

Another unique characteristic of our novel method lies in its ability to dynamically weight the four SAR parameters using the attention weights matrix derived from the Transformer model. This dynamic

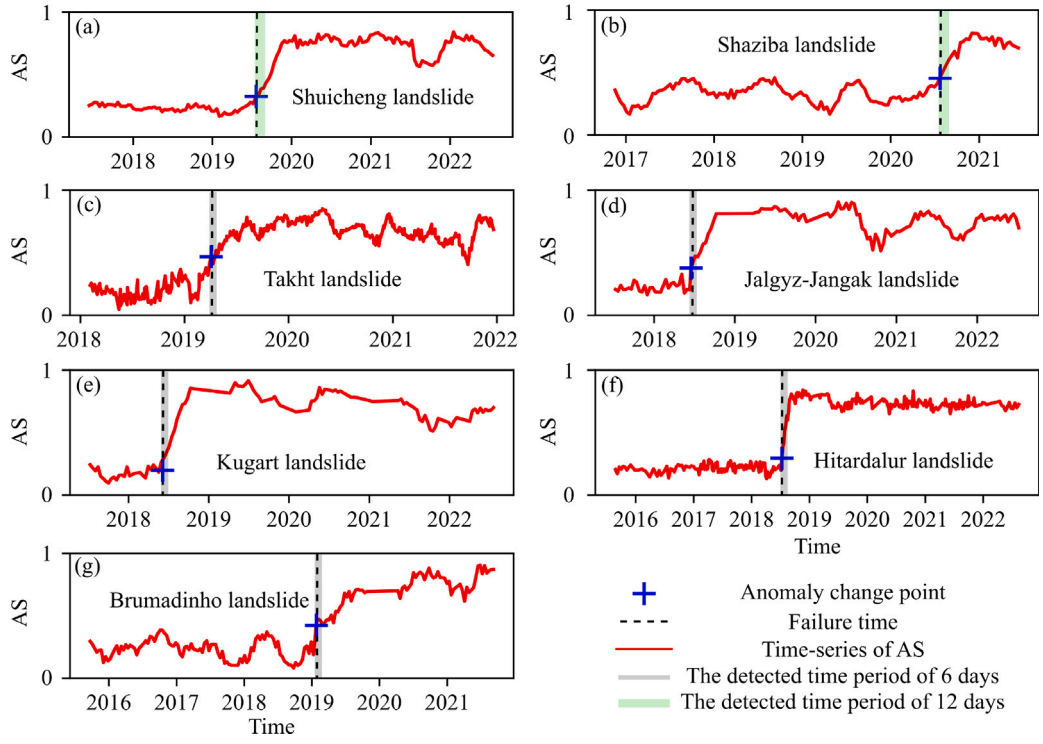


Fig. 7. Time detection results of seven landslide cases based on our improved Transformer model. The black dashed lines represent the failure times. The blue crosses represent anomaly change points in the time-series of anomaly score (AS). The anomaly change point is detected by the proposed adaptive threshold method from the time-series of anomaly score, and its acquired date is the start day of the occurrence period. Light green and gray bars represent the detected time period of 12 days and 6 days, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

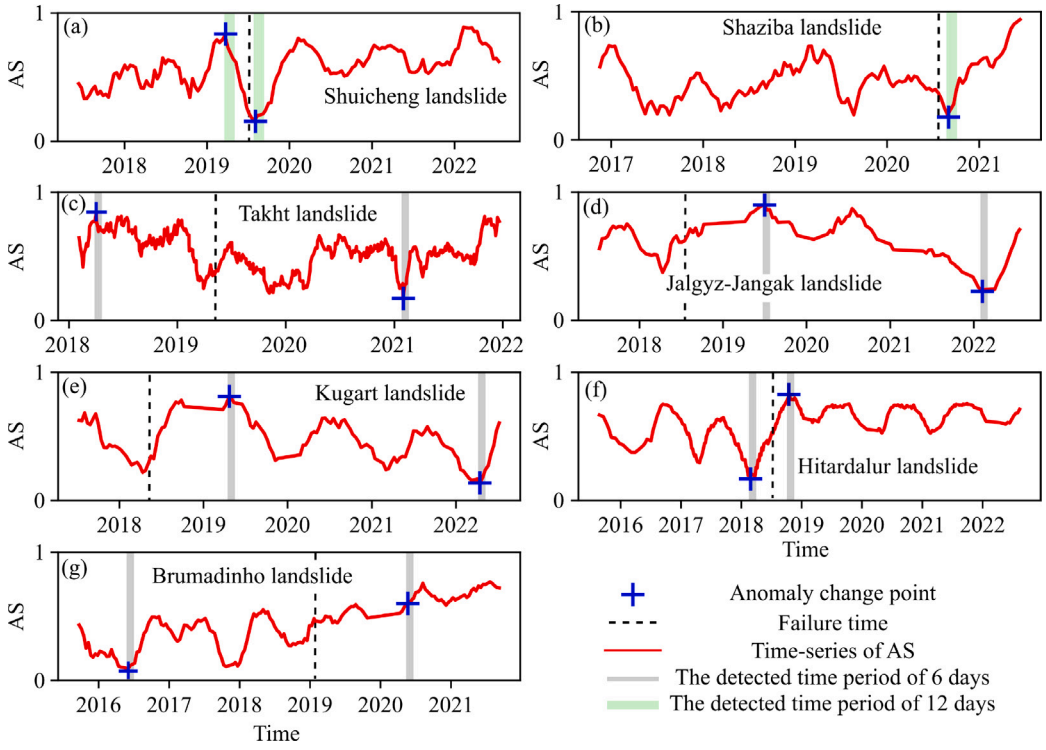


Fig. 8. Time detection results of seven landslide cases based on the LSTM model. The black dashed lines represent the failure times. The blue crosses represent anomaly points in the time-series of anomaly score (AS). The anomaly change point is detected by the proposed adaptive threshold method from the time-series of anomaly score, and its acquired date is the start day of the occurrence period. Light green and gray bars represent the detected time period of 12 days and 6 days, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

weighting improves the accuracy of landslide time detection. Taking the Takht landslide as an example, where all four parameters can be exploited to detect the failure time, but the coherence could not identify landslide occurrence time very well compared to the other three parameters. Assigning equal weights to all four parameters would result in false detection. In contrast to LSTM, our improved Transformer model can automatically and optimally increase or decrease the weights of coherence, alpha, entropy, and amplitude in an adaptive manner to enhance the precision of time detection. This adaptive approach is also beneficial for other case studies involving multi-variants time series analysis.

5.4. Robust performance of our methods

Our model also exhibits strong robustness against various noise on the time-series of four SAR parameters caused by rainfall, snowfall, vegetation, and anthropogenic activities. For example, two false alarms are observed on the time-series of coherence in the Shaziba landslide due to seasonality, and one false alarm is detected on the time-series of amplitude in the Shuicheng landslide due to rainfall. In addition, each time-series is characterized by irregularity and significant local waves, spikes, and peaks. Despite being disturbed by these noises, our model detected a time period of 6 days overlapping with the real time of the Shaziba and Shuicheng landslide occurrences in the correct manners (Fig. 7). In contrast, the LSTM model failed to detect the occurrence times of these two landslides due to such mentioned noises (Fig. 8).

5.5. The automated framework

The occurrence period of a single landslide is detected effectively because SAR data are highly responsive to landslide occurrences. However, the successful automated detection of multiple landslide occurrence periods is probably attributed to the implementation of the automated framework on the SAR data. For example, the Hitardalur and Brumadinho's landslide occurrence times are visible in time-series of amplitude as shown in Fig. 6 (f) and (g), whereas the Shaziba landslide occurrence time can be investigated from the coherence time-series. Landslides cannot be dated only using one fixed SAR parameter, which is limited by several reasons, such as different vegetation, rainfall, and snowfall. Multi SAR-derived parameters can help us better date landslides. However, manually reviewing the SAR-derived parameters for each landslide would be a time-consuming challenge. Several studies (e.g., (Fu et al., 2023; Deijns et al., 2022)) have utilized the threshold method, which requires manual intervention to determine landslide occurrence time, utilizing either optical data or SAR data coherence. Nevertheless, manual threshold selection may not be ideal for dating a considerable quantity of landslides. Compared with previous studies (Fu et al., 2023; Deijns et al., 2022), our framework can automatically detect the landslide occurrences period based on a deep learning model and four SAR-derived parameters.

5.6. Limitations

For our temporal detection methodology, the spatial detection of landslide boundaries is a prior foundation. In this work, the provided landslide spatial boundaries from existing archives and inventories were applied to subset the images of SAR-derived parameters. Therefore, it is essential to obtain certain prior knowledge from various perspectives before applying our method to narrow down the search for landslide occurrence time. Moreover, this framework has been successfully applied to landslides in mountainous areas covered by vegetation, where we can hypothesize that the slope is dominated by volume scattering before the failure and by surface scattering after the failure. This transition in the scattering mechanism causes measurable changes in SAR-derived parameters such as interferometric coherence, polarimetric alpha, entropy, and amplitude. The direct application of this method for dating landslides in non-vegetated regions could be challenging as the absence of vegetation limits the transition in the scattering mechanism. However, there are several techniques that can be employed to improve the applicability for non-vegetated areas. One potential approach could be to utilize object-based methods (Martha et al., 2010), such as landslide diagnostic features using textural characteristics based on SAR greyscale images, to detect the occurrence of landslides. This needs to be investigated in more detail in future work. Our methodology is well-suited for the detection of large, rapid failures characterized by significant surface changes, resulting in distinct boundaries after the failure due to the washout or erosion effects. However, its applicability is limited to slow-moving landslides and creeping motions, which do not show changes in the backscattering properties of the landslide area, making detection using SAR parameters impracticable. In comparison to optical data (Behling et al., 2014), applying Sentinel-1 SAR data has significant potential to narrow down the occurrence time period for most of the landslide cases, but there is still potential for further improvement. The primary issue affecting the accuracy of time detection is the periodicity of Sentinel-1 satellite imagery. Both Sentinel-1 A and Sentinel-1B satellite data do not yet completely cover the whole ground surface worldwide, and the revisiting time was 12 days for most regions until Dec. 2023; The Sentinel-1B satellite was no longer able to acquire data since Dec. 23, 2021. The accuracy of time detection can be increased by reducing the interval between image acquisitions from SAR satellites. However, judging by the archive data and imaging schedules, this is an unlikely option for all landslide cases. Moreover, there may also be missing images due to Sentinel-1 orbit errors, ground download, or calibration issues. Besides, the lack of imagery close to the failure is also a reason for not properly estimating the failure time. In the future, the combination of multiple satellite constellations, e.g., Radarsat−2, ALOS-2, TerraSAR-X, and NISAR can improve the sampling frequency of the dataset, leading to a higher temporal resolution when applying our method.

The SAR parameters used for landslide occurrence time period detection are also limited by the quality of the landslide boundaries and the resolution of the SAR data. The SAR parameters for a single timestamp of time-series are averaged spatially into every value over the landslide's body. Therefore, the quality of landslide boundary mapping and the SAR data resolution would greatly impact landslide detection. Utilizing higher resolution SAR satellites enhances the accuracy of the results obtained by our proposed methodology.

6. Conclusion

In this study, we propose a novel automated framework to detect landslide occurrence times using multivariate time-series analysis of Sentinel-1 SAR data based on a deep-learning approach.

The main steps in this framework are outlined below. Firstly, the time-series of interferometric coherence, polarization parameters (alpha and entropy), and SAR amplitude for the given landslide boundaries are derived and combined as the input data to implement the framework. Then, the improved multivariate Transformer model via gated relative position is adopted to successfully learn the patterns in the time-series, and highlight changes related to catastrophic failures via an anomaly score. Finally, a developed adaptive threshold method is used to automatically identify the jump point in the time-series of anomaly scores and estimates the time of the landslide.

This framework has been tested on seven landslide cases distributed all over the world with different locations and shapes, and results were compared with those obtained from optical data as well as from the LSTM model. Our study has yielded several noteworthy findings. For all cases, the experimental results are shown to verify the effectiveness of our method. Notably compared to optical data, the time windows detected from SAR data were significantly refined when compared to optical data. The proposed method outperformed the LSTM model. The time-series of anomaly scores predicted by LSTM was noisier than

our improved Transformer model. Consequently, employing LSTM for automatic landslide dating proved challenging.

In the future, the application field can be extended for the integration of multiple remote sensing acquisitions in improving the spatiotemporal accuracy of applying our method, as well as the exploitation of our framework in complementing date information for regional landslide inventories.

CRediT authorship contribution statement

Wandi Wang: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mahdi Motagh:** Writing – review & editing, Supervision, Funding acquisition. **Zhuge Xia:** Discussion. **Simon Plank:** Discussion. **Zhe Li:** Discussion. **Aiym Orynbaikyzy:** Discussion. **Chao Zhou:** Discussion. **Sigrid Roessner:** Discussion.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge the Copernicus program for free access to Sentinel-1 data. This work was supported by Helmholtz, Germany within the framework of the HIP project MultiSaT4SLOWS. The authors are grateful to two anonymous reviewers and the Editor, Jonathan Li, for their constructive comments. W.W. is supported by China Scholarship Council (CSC) Grant 202006450011.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jag.2024.103795.

References

Behling, R., Roessner, S., Golovko, D., Kleinschmit, B., 2016. Derivation of long-term spatiotemporal landslide activity—A multi-sensor time series approach. Remote Sens. Environ. 186, 88–104. http://dx.doi.org/10.1016/j.rse.2016.07.017.

Behling, R., Roessner, S., Kaufmann, H., Kleinschmit, B., 2014. Automated spatiotemporal landslide mapping over large areas using rapideye time series data. Remote Sens. 6 (9), 8026–8055. http://dx.doi.org/10.3390/rs6098026.

Bindlish, R., Barros, A.P., 2000. Multifrequency soil moisture inversion from SAR measurements with the use of IEM. Remote Sens. Environ. 71 (1), 67–88. http://dx.doi.org/10.1016/s0034-4257(99)00065-6.

Brabb, E.E., 1991. The world landslide problem. Episodes J. Int. Geosci. 14 (1), 52–61. http://dx.doi.org/10.18814/epiiugs/1991/v14i1/008.

Burrows, K., Walters, R.J., Milledge, D., Densmore, A.L., 2020. A systematic exploration of satellite radar coherence methods for rapid landslide detection. Nat. Hazards Earth Syst. Sci. 20 (11), 3197–3214. http://dx.doi.org/10.5194/nhess-20-3197-2020.

Canizo, M., Triguero, I., Conde, A., Onieva, E., 2019. Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study. Neurocomputing 363, 246–260. http://dx.doi.org/10.1016/j.neucom.2019.07.034.

Choker, M., Baghdadi, N., Zribi, M., El Hajj, M., Paloscia, S., Verhoest, N.E., Lievens, H., Mattia, F., 2017. Evaluation of the Oh, Dubois and IEM backscatter models using a large dataset of SAR data and experimental soil measurements. Water 9 (1), 38. http://dx.doi.org/10.3390/w9010038.

Cloude, S.R., Pottier, E., 1997. An entropy based classification scheme for land applications of polarimetric SAR. IEEE Trans. Geosci. Remote Sens. 35 (1), 68–78. http://dx.doi.org/10.1109/36.551935.

Colesanti, C., Wasowski, J., 2006. Investigating landslides with space-borne synthetic aperture radar (SAR) interferometry. Eng. Geol. 88 (3–4), 173–199. http://dx.doi.org/10.1016/j.enggeo.2006.09.013.

Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., Crippa, B., 2016. Persistent scatterer interferometry: A review. ISPRS J. Photogramm. Remote Sens. 115, 78–89. http://dx.doi.org/10.1016/j.isprsjprs.2015.10.011, Theme issue 'State-of-the-art in photogrammetry, remote sensing and spatial information science'.

Czuchlewski, K.R., Weissel, J.K., Kim, Y., 2003. Polarimetric synthetic aperture radar study of the Tsaoling landslide generated by the 1999 Chi-Chi earthquake, Taiwan. J. Geophys. Res.: Earth Surf. 108 (F1), http://dx.doi.org/10.1029/2003jf000037.

De Jeu, R.A., Wagner, W., Holmes, T., Dolman, A., Van De Giesen, N., Friesen, J., 2008. Global soil moisture patterns observed by space borne microwave radiometers and scatterometers. Surv. Geophys. 29, 399–420. http://dx.doi.org/10.1007/s10712-008-9044-0.

Deijns, A.A., Dewitte, O., Thiery, W., d'Oreye, N., Malet, J.P., Kervyn, F., 2022. Timing landslide and flash flood events from SAR satellite: a regionally applicable methodology illustrated in african cloud-covered tropical environments. Nat. Hazards Earth Syst. Sci. 22 (11), 3679–3700. http://dx.doi.org/10.5194/nhess-22-3679-2022.

Delacourt, C., Allemand, P., Berthier, E., Raucoules, D., Casson, B., Grandjean, P., Pambrun, C., Varel, E., 2007. Remote-sensing techniques for analysing landslide kinematics: a review. Bull. Soc. Géol. France 178 (2), 89–100. http://dx.doi.org/10.2113/gssgfbull.178.2.89.

Du, J., Glade, T., Woldai, T., Chai, B., Zeng, B., 2020. Landslide susceptibility assessment based on an incomplete landslide inventory in the Jilong Valley, Tibet, Chinese Himalayas. Eng. Geol. 270, 105572. http://dx.doi.org/10.1016/j.enggeo.2020.105572.

Froude, M.J., Petley, D.N., 2018. Global fatal landslide occurrence from 2004 to 2016. Nat. Hazards Earth Syst. Sci. 18 (8), 2161–2181. http://dx.doi.org/10.5194/nhess-18-2161-2018.

Fu, S., de Jong, S.M., Deijns, A., Geertsema, M., de Haas, T., 2023. The SWADE model for landslide dating in time series of optical satellite imagery. Landslides 20 (5), 913–932. http://dx.doi.org/10.1007/s10346-022-02012-4.

Guo, Y., Liao, W., Wang, Q., Yu, L., Ji, T., Li, P., 2018. Multidimensional time series anomaly detection: A gru-based gaussian mixture variational autoencoder approach. In: Asian Conference on Machine Learning. PMLR, pp. 97–112.

Guzzetti, F., Mondini, A.C., Cardinali, M., Fiorucci, F., Santangelo, M., Chang, K.T., 2012. Landslide inventory maps: New tools for an old problem. Earth-Sci. Rev. 112 (1–2), 42–66. http://dx.doi.org/10.1016/j.earscirev.2012.02.001.

Hajnsek, I., Pottier, E., Cloude, S.R., 2003. Inversion of surface parameters from polarimetric SAR. IEEE Trans. Geosci. Remote Sens. 41 (4), 727–744. http://dx.doi.org/10.1109/tgrs.2003.810702.

He, Z., Xu, X., Deng, S., 2003. Discovering cluster-based local outliers. Pattern Recognit. Lett. 24 (9–10), 1641–1650. http://dx.doi.org/10.1016/S0167-8655(03)00003-5.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780. http://dx.doi.org/10.1162/neco.1997.9.8.1735.

Jacquemart, M., Tiampo, K., 2021. Leveraging time series analysis of radar coherence and normalized difference vegetation index ratios to characterize pre-failure activity of the Mud Creek landslide, California. Nat. Hazards Earth Syst. Sci. 21 (2), 629–642. http://dx.doi.org/10.5194/nhess-21-629-2021.

Jin, H., Mountrakis, G., Stehman, S.V., 2014. Assessing integration of intensity, polarimetric scattering, interferometric coherence and spatial texture metrics in PALSAR-derived land cover classification. ISPRS J. Photogramm. Remote Sens. 98, 70–84. http://dx.doi.org/10.1016/j.isprsjprs.2014.09.017.

Jung, J., Kim, D.j., Lavalle, M., Yun, S.H., 2016. Coherent change detection using InSAR temporal decorrelation model: A case study for volcanic ash detection. IEEE Trans. Geosci. Remote Sens. 54 (10), 5765–5775. http://dx.doi.org/10.1109/TGRS.2016.2572166.

Jung, J., Yun, S.H., 2020. Evaluation of coherent and incoherent landslide detection methods based on synthetic aperture radar for rapid response: A case study for the 2018 Hokkaido landslides. Remote Sens. 12 (2), 265. http://dx.doi.org/10.3390/rs12020265.

Kilburn, C.R., Petley, D.N., 2003. Forecasting giant, catastrophic slope collapse: lessons from Vajont, Northern Italy. Geomorphology 54 (1–2), 21–32. http://dx.doi.org/10.1016/S0169-555X(03)00052-7.

Lee, J.S., Jurkevich, L., Dewaele, P., Wambacq, P., Oosterlinck, A., 1994. Speckle filtering of synthetic aperture radar images: A review. Remote Sens. Rev. 8 (4), 313–340. http://dx.doi.org/10.1080/02757259409532206.

Li, Z., Mak, M.W., Meng, H.M.L., 2023. Discriminative speaker representation via contrastive learning with class-aware attention in angular space. In: ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, IEEE, pp. 1–5. http://dx.doi.org/10.1109/ICASSP49357.2023.10096230.

Lu, P., Bai, S., Tofani, V., Casagli, N., 2019a. Landslides detection through optimized hot spot analysis on persistent scatterers and distributed scatterers. ISPRS J. Photogramm. Remote Sens. 156, 147–159. http://dx.doi.org/10.1016/j.isprsjprs.2019.08.004.

Lu, P., Qin, Y., Li, Z., Mondini, A.C., Casagli, N., 2019b. Landslide mapping from multi-sensor data through improved change detection-based Markov random field. Remote Sens. Environ. 231, 111235. http://dx.doi.org/10.1016/j.rse.2019.111235.

Malhotra, P., Ramakrishnan, A., Anand, G., Vig, L., Agarwal, P., Shroff, G., 2016. LSTM-based encoder-decoder for multi-sensor anomaly detection. http://dx.doi.org/10.48550/arXiv.1607.00148, arXiv preprint arXiv:1607.00148.

Malhotra, P., Vig, L., Shroff, G., Agarwal, P., et al., 2015. Long short term memory networks for anomaly detection in time series. In: ESANN, Vol. 2015. p. 89.

Martha, T.R., Kerle, N., Jetten, V., van Westen, C.J., Kumar, K.V., 2010. Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods. Geomorphology 116 (1–2), 24–36. http://dx.doi.org/10.1016/j.geomorph.2009.10.004.

Mondini, A.C., Guzzetti, F., Chang, K.T., Monserrat, O., Martha, T.R., Manconi, A., 2021. Landslide failures detection and mapping using Synthetic Aperture Radar: Past, present and future. Earth-Sci. Rev. 216, 103574. http://dx.doi.org/10.1016/j.earscirev.2021.103574.

Mondini, A.C., Santangelo, M., Rocchetti, M., Rossetto, E., Manconi, A., Monserrat, O., 2019. Sentinel-1 SAR amplitude imagery for rapid landslide detection. Remote Sens. 11 (7), 760. http://dx.doi.org/10.3390/rs11070760.

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., Papathanassiou, K.P., 2013. A tutorial on synthetic aperture radar. IEEE Geosci. Remote Sens. Mag. 1 (1), 6–43. http://dx.doi.org/10.1109/MGRS.2013.2248301.

Morrison, K., Wagner, W., 2019. Explaining anomalies in SAR and scatterometer soil moisture retrievals from dry soils with subsurface scattering. IEEE Trans. Geosci. Remote Sens. 58 (3), 2190–2197. http://dx.doi.org/10.1109/tgrs.2019.2954771.

Motagh, M., Wetzel, H.U., Roessner, S., Kaufmann, H., 2013. A TerraSAR-X InSAR study of landslides in southern Kyrgyzstan, central Asia. Remote Sens. Lett. 4 (7), 657–666. http://dx.doi.org/10.1080/2150704x.2013.782111.

Nanduri, A., Sherry, L., 2016. Anomaly detection in aircraft data using recurrent neural networks (RNN). In: 2016 Integrated Communications Navigation and Surveillance. ICNS, Ieee, pp. 5C2–1.

Niu, C., Zhang, H., Liu, W., Li, R., Hu, T., 2021. Using a fully polarimetric SAR to detect landslide in complex surroundings: Case study of 2015 Shenzhen landslide. ISPRS J. Photogramm. Remote Sens. 174, 56–67. http://dx.doi.org/10.1016/j.isprsjprs.2021.01.022.

Oliver, C., Quegan, S., 2004. Understanding Synthetic Aperture Radar Images. SciTech Publishing.

Pardeshi, S.D., Autade, S.E., Pardeshi, S.S., 2013. Landslide hazard assessment: recent trends and techniques. SpringerPlus 2, 1–11. http://dx.doi.org/10.1186/2193-1801-2-523.

Pascanu, R., Mikolov, T., Bengio, Y., 2013. On the difficulty of training recurrent neural networks. In: Dasgupta, S., McAllester, D. (Eds.), Proceedings of the 30th International Conference on Machine Learning. In: Proceedings of Machine Learning Research, vol. 28, no. 3, PMLR, Atlanta, Georgia, USA, pp. 1310–1318.

Plank, S., 2014. Rapid damage assessment by means of multi-temporal sar—a comprehensive review and outlook to sentinel-1. Remote Sensing 6 (6), 4870–4906. http://dx.doi.org/10.3390/rs6064870.

Plank, S., Twele, A., Martinis, S., 2016. Landslide mapping in vegetated areas using change detection based on optical and polarimetric SAR data. Remote Sens. 8 (4), 307. http://dx.doi.org/10.3390/rs8040307.

Ramaswamy, S., Rastogi, R., Shim, K., 2000. Efficient algorithms for mining outliers from large data sets. In: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data. pp. 427–438. http://dx.doi.org/10.1145/342009.335437.

Santangelo, M., Cardinali, M., Bucci, F., Fiorucci, F., Mondini, A.C., 2022. Exploring event landslide mapping using sentinel-1 SAR backscatter products. Geomorphology 397, 108021. http://dx.doi.org/10.1016/j.geomorph.2021.108021.

Shibayama, T., Yamaguchi, Y., Yamada, H., 2015. Polarimetric scattering properties of landslides in forested areas and the dependence on the local incidence angle. Remote Sens. 7 (11), 15424–15442. http://dx.doi.org/10.3390/rs71115424.

Shimada, M., Watanabe, M., Kawano, N., Ohki, M., Motooka, T., Wada, Y., 2014. Detecting mountainous landslides by SAR polarimetry: A comparative study using Pi-SAR-L2 and X-band SARs. Trans. Japan Soc. Aeronaut. Space Sci., Aerosp. Technol. Japan 12 (ists29), Pn_9–Pn_15. http://dx.doi.org/10.2322/tastj.12.pn_9.

Shyu, M.L., Chen, S.C., Sarinnapakorn, K., Chang, L., 2003. A Novel Anomaly Detection Scheme Based on Principal Component Classifier. Technical Report, Miami Univ Coral Gables Fl Dept of Electrical and Computer Engineering, http://dx.doi.org/10.1007/11539827_18.

Tang, J., Chen, Z., Fu, A.W.C., Cheung, D.W., 2002. Enhancing effectiveness of outlier detections for low density patterns. In: Advances in Knowledge Discovery and Data Mining: 6th Pacific-Asia Conference, PAKDD 2002 Taipei, Taiwan, May 6–8, 2002 Proceedings 6. Springer, pp. 535–548. http://dx.doi.org/10.1007/3-540-47887-6_53.

Touzi, R., Lopes, A., Bruniquel, J., Vachon, P.W., 1999. Coherence estimation for SAR imagery. IEEE Trans. Geosci. Remote Sens. 37 (1), 135–149. http://dx.doi.org/10.4095/219098.

Uemoto, J., Moriyama, T., Nadai, A., Kojima, S., Umehara, T., 2019. Landslide detection based on height and amplitude differences using pre-and post-event airborne X-band SAR data. Nat. Hazards 95, 485–503. http://dx.doi.org/10.1007/s11069-018-3492-8.

Van Westen, C.J., Castellanos, E., Kuriakose, S.L., 2008. Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. Eng. Geol. 102 (3–4), 112–131. http://dx.doi.org/10.1016/j.enggeo.2008.03.010.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Adv. Neural Inf. Process. Syst. 30, http://dx.doi.org/10.5040/9781350101272.00000005.

Verhoest, N.E., Lievens, H., Wagner, W., Álvarez-Mozos, J., Moran, M.S., Mattia, F., 2008. On the soil roughness parameterization problem in soil moisture retrieval of bare surfaces from synthetic aperture radar. Sensors 8 (7), 4213–4248. http://dx.doi.org/10.3390/s8074213.

Villarroya-Carpio, A., Lopez-Sanchez, J.M., Engdahl, M.E., 2022. Sentinel-1 interferometric coherence as a vegetation index for agriculture. Remote Sens. Environ. 280, 113208. http://dx.doi.org/10.1016/j.rse.2022.113208.

Wang, W., Motagh, M., Mirzaee, S., Li, T., Zhou, C., Tang, H., Roessner, S., 2023. The 21 July 2020 Shaziba landslide in China: Results from multi-source satellite remote sensing. Remote Sens. Environ. 295, 113669. http://dx.doi.org/10.1016/j.rse.2023.113669.

Xia, Z., Motagh, M., Li, T., Peng, M., Roessner, S., 2023. A methodology to characterize 4D post-failure slope instability dynamics using remote sensing measurements: A case study of the Aniangzhai landslide in Sichuan, Southwest China. ISPRS J. Photogramm. Remote Sens. 196, 402–414. http://dx.doi.org/10.1016/j.isprsjprs.2023.01.006.

Xia, Z., Motagh, M., Li, T., Roessner, S., 2022. The June 2020 Aniangzhai landslide in Sichuan Province, Southwest China: Slope instability analysis from radar and optical satellite remote sensing data. Landslides 1–17. http://dx.doi.org/10.1007/s10346-021-01777-4.

Yonezawa, C., Watanabe, M., Saito, G., 2012. Polarimetric decomposition analysis of ALOS PALSAR observation data before and after a landslide event. Remote Sens. 4 (8), 2314–2328. http://dx.doi.org/10.3390/rs4082314.

Zebker, H.A., Villasenor, J., et al., 1992. Decorrelation in interferometric radar echoes. IEEE Trans. Geosci. Remote Sens. 30 (5), 950–959. http://dx.doi.org/10.1109/36.175330.

Zhou, C., Cao, Y., Hu, X., Yin, K., Wang, Y., Catani, F., 2022a. Enhanced dynamic landslide hazard mapping using MT-insar method in the three Gorges Reservoir Area. Landslides 19 (7), 1585–1597. http://dx.doi.org/10.1007/s10346-021-01796-1.

Zhou, C., Cao, Y., Yin, K., Intrieri, E., Catani, F., Wu, L., 2022b. Characteristic comparison of seepage-driven and buoyancy-driven landslides in Three Gorges Reservoir area, China. Eng. Geol. 301, 106590. http://dx.doi.org/10.1016/j.enggeo.2022.106590.

Zhu, L., Walker, J.P., Ye, N., Rüdiger, C., 2019. Roughness and vegetation change detection: A pre-processing for soil moisture retrieval from multi-temporal SAR imagery. Remote sensing of environment 225, 93–106. http://dx.doi.org/10.1016/j.rse.2019.02.027.