JGR Solid Earth

RESEARCH ARTICLE
10.1029/2021JB022957

Special Section:
Machine learning for Solid Earth observation, modeling and understanding

Pinpointing Early Signs of Impending Slope Failures From Space

Shuo Zhou[1], Antoinette Tordesillas[1], Emanuele Intrieri[2], Federico Di Traglia[3], Guoqi Qian[1], and Filippo Catani[4]

[1]School of Mathematics and Statistics, University of Melbourne, Melbourne, VIC, Australia, [2]Department of Earth Sciences, University of Florence, Florence, Italy, [3]National Institute of Oceanography and Applied Geophysics-OGS, Sgonico, Italy, [4]Department of Geosciences, Machine Intelligence and Slope Stability Laboratory, University of Padova, Padova, Italy

Key Points:

- Ground deformation data from Sentinel-1 encode rich information and can be used to predict the location of impending landslides
- In imbalanced satellite data, point-level clustering dynamics diminish but outlying motions at grid cell level become distinctive
- The third quartile of ground displacement in a cell accurately identifies the location of the 2017 Xinmo landslide almost 1 year in advance

Correspondence to:
A. Tordesillas,
atordesi@unimelb.edu.au

Citation:
Zhou, S., Tordesillas, A., Intrieri, E., Di Traglia, F., Qian, G., & Catani, F. (2022). Pinpointing early signs of impending slope failures from space. *Journal of Geophysical Research: Solid Earth*, *127*, e2021JB022957. https://doi.org/10.1029/2021JB022957

Received 4 AUG 2021
Accepted 12 JAN 2022

Author Contributions:
Conceptualization: Antoinette Tordesillas
Data curation: Emanuele Intrieri, Federico Di Traglia
Formal analysis: Shuo Zhou, Antoinette Tordesillas
Funding acquisition: Antoinette Tordesillas
Investigation: Shuo Zhou, Antoinette Tordesillas

Abstract A promised potential of spaceborne interferometric synthetic aperture radars (InSAR) is a capability for regularly monitoring ground deformation with millimeter accuracy, for timely forecasting of impending natural hazards such as landslides. The main limitation in InSAR being actually capable of unleashing this potential for hazard prediction is that key precursory ground displacements are, in the majority of cases, a very small subset of the entire big data set provided by the method over large regions. Consequently, pinpointing a single impending failure may become very difficult or impossible. We develop a data-driven framework that can handle such imbalanced spatiotemporal data based on the concept of outlying aspects mining, to find a subset of features out of a collection of potential features, which best distinguishes the landslide source area from the others. We show that the identified feature subspace can be used to find anomalous areas across multiple spatial scales, such that Sentinel-1 satellite monitoring points which persistently lie in these areas can accurately detect the location of the Xinmo landslide (China) almost 1 year in advance—without false alarms. In a second case study, we identify the area affected by rockfalls on Stromboli volcano, a task that is generally infeasible with traditional methods applied to InSAR data. With continuing improvements in the spatial and temporal resolution from the new generation of satellites, such as Sentinel-1, this approach opens the door to reliable and early prediction of failure over a broad range of slope instabilities.

Plain Language Summary The utility of Earth observation satellite data is well recognized in early warning for natural hazards such as landslides. With the revisiting times now down to 6 days, sufficient information can be extracted to predict the location of an impending landslide. However, one still needs to know where to focus the analysis at first, especially when the target area is very small relative to the size of the monitoring domain and lies in among other unstable high-risk slopes. We address this problem by learning the most unique characteristics related to the landslide source area directly from the displacement data captured by Sentinel-1 satellites, and building a new prediction method to find the temporally persistent outlying areas over multiple spatial scales as the final prediction. The results from two case studies—2017 Xinmo landslide and Stromboli volcano rockfalls in 2015–2016—show that the location of the impending failure can be accurately identified well before the time of failure, by tracking the abnormality in a simple feature: the third quartile in the ground displacement of small regions. This approach provides an important advance for timely risk assessment from satellite data, especially for remote areas that are difficult, if not impossible, to access.

1. Introduction

Satellite interferometric radars (InSAR) are ushering in a new decade of large-scale monitoring data at unsurpassed spatial and temporal resolution to support real-time early warning systems (EWS) for impending disasters (Carlà et al., 2019; Intrieri et al., 2018, 2019; Yang et al., 2019). Game-changing opportunities have been anticipated for monitoring and forecasting landslides and related hazards (Intrieri et al., 2018; Yang et al., 2019). Realizing such opportunities, however, depends crucially on algorithms that can efficiently sift through these big and complex data and distill the information into timely and accurate intelligence for risk management decision-making (Anantrasirichai et al., 2018; Cigna & Tapete, 2021; Sansosti et al., 2014; Zinno et al., 2016). Having such data-driven approaches deployed directly on large-scale spaceborne deformation data can benefit landslide-monitoring practitioners in many ways, including: (a) new opportunities can be identified in monitoring and forecasting landslide hazards for remote areas that are difficult, if not impossible, to access for in situ surveys; (b) more timely prediction of landslides can be delivered at lower cost, compared to traditional slope-based

Methodology: Shuo Zhou, Antoinette Tordesillas, Guoqi Qian
Project Administration: Antoinette Tordesillas
Resources: Antoinette Tordesillas
Software: Shuo Zhou
Supervision: Antoinette Tordesillas
Validation: Shuo Zhou
Visualization: Shuo Zhou, Emanuele Intrieri, Federico Di Traglia
Writing – original draft: Shuo Zhou, Antoinette Tordesillas
Writing – review & editing: Shuo Zhou, Antoinette Tordesillas, Emanuele Intrieri, Federico Di Traglia, Guoqi Qian, Filippo Catani

methods which depend on the availability of detailed geophysical characteristics of the impending failure; (c) a potential hazard subset area can be extracted from the vast amount of monitoring space well in advance for detailed slope stability assessment; (d) the need for predefined assumptions on the type of failure or the underlying trigger mechanism can be eliminated, since automatic intelligence are directly and continually learned from the streaming deformation data. By and large, such algorithms are still lacking, given the challenges imposed by the imbalanced, high-dimensional, spatiotemporal data that invariably result from observations of slope dynamics at large spatial scales (Carlà et al., 2019; Intrieri et al., 2019; Yang et al., 2019).

Observations of precursory displacements give rise to high-dimensional, spatiotemporal data (Intrieri et al., 2018; Yang et al., 2019). Clues to where and when failure happens are hidden in the spatiotemporal structures or correlations embodied in the data (Gao et al., 2019; Intrieri et al., 2018; Singh & Tordesillas, 2020; Tordesillas, Zhou, et al., 2020; Zhou et al., 2020). Traditional landslide forecasting methods employed in early warning systems do not account for these structures, as they typically focus only on the analysis of individual univariate time series from selected locations of the studied domain, as discussed in recent reviews (Carlà et al., 2019; Intrieri et al., 2019; Stähli et al., 2015) and several applications (Intrieri et al., 2017; Intrieri & Gigli, 2016; Kothari & Momayez, 2018). Furthermore, even after the advent of persistent scatterers interferometry and ground-based interferometric synthetic aperture radar (GB-InSAR) that are able to produce nearly spatially continuous displacement maps, the whole spatial data set is rarely exploited (Dick et al., 2015). The difficulty is that spatiotemporal data analytics is still in its infancy and methods that can robustly extract and characterize spatiotemporal structures, while remaining computationally feasible, are still being developed (Wang et al., 2020). In recent studies, progress have been made in resolving this issue by developing a new class of data-driven tools that characterize the emerging spatiotemporal patterns to predict the ultimate failure pattern early in the precursory regime (Das & Tordesillas, 2019; Singh & Tordesillas, 2020; Tordesillas et al., 2018, 2021; Tordesillas, Zhou, et al., 2020; Wang et al., 2020; Zhou et al., 2020). By design, these tools exploit the whole-of-slope radar data. Application of these in the analysis of ground motion radar data gave promising results (Das & Tordesillas, 2019; Singh & Tordesillas, 2020; Tordesillas et al., 2018, 2021; Wang et al., 2020; Zhou et al., 2020). In particular, analysis of thousands of displacement time series data for a developing rockslide in an operational open pit mine, obtained using GB-InSAR, gave a precise prediction of the ultimate location and geometry of the failure region, weeks in advance of the event (Tordesillas et al., 2018, 2021). Building on this capability, a subsequent stochastic spatiotemporal forecasting algorithm identified imminent failure 10 hr prior to the collapse (Wang et al., 2020).

A limitation of some of these new data-driven tools, specifically those which use standard clustering methods to identify the failure location (Das & Tordesillas, 2019; Zhou et al., 2020) is that they work well only for balanced data. In the case of landslide monitoring, this means that measurement points in the potential unstable slope have to account for a significant proportion of the entire data set in order to detect the unique precursory failure dynamics in space and time. Typically, balanced data are only available on small-scale analysis (e.g., GB-InSAR data where one is already aware about the location of the impending landslide and so the imbalance can be avoided by a priori spatial subsetting of data). However, this is only an exception but not the rule, since the usual InSAR data provides large area coverage containing only a tiny fraction of measurement points within the unknown failure location (e.g., the Sentinel-1 data in Carlà et al., 2019). Direct application of the existing algorithms designed for balanced data to characterize the early prefailure dynamics from a regional-scale satellite data is infeasible, since the dynamics in the data from the points of interest, namely those that will later end up in the failure region, is heavily obfuscated due to the overwhelming number of stable points amidst many other unstable moving sites that did not eventuate into a catastrophic collapse. Accordingly, this study seeks to address this issue by developing a new framework that can handle such severe imbalance in monitoring data, such that an accurate and timely prediction to the location of the impending landslide can be made directly from the imbalanced satellite data.

To achieve this, we first apply a data-driven tool on the Xinmo landslide, in order to identify the most outlying aspects (a subset of features out of a collection of potential features) that best distinguishes a given target object from the others. The tool is called outlying aspects mining (OAM; Duan et al., 2015; Vinh et al., 2016), and the output feature subset is referred to the optimal feature subspace (OFS), which constitutes the most unique characteristics of the target object. To deploy OAM, we uniformly divide the monitoring area in the Xinmo Sentinel-1 data into multiple grid cells and treat the cell that is the closest spatially to the landslide source area as the *target* cell, in order to learn the OFS. This is done for a wide range of grid cell sizes to ensure the search method is robust to scale variations. We then show that the resultant OFS can be used to find the temporally persistent outlying

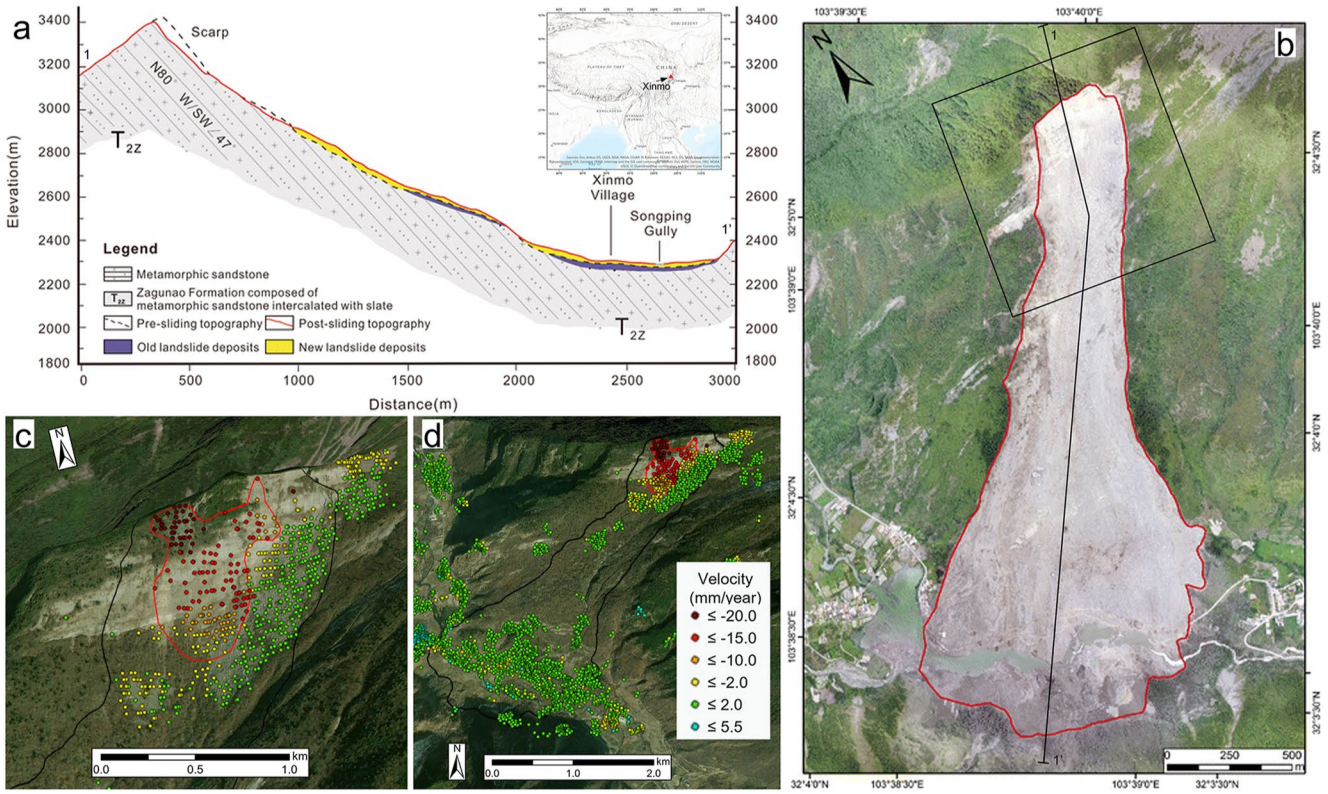

Figure 1. Xinmo landslide. (a) Cross-section of the landslide (from Fan et al., 2017) with the location of the site in the inset. (b) Postevent image taken by unoccupied aerial vehicle (UAV) on 26 June 2017 (Fan et al., 2017). (c) Subset of the Sentinel-1 data showing (outlined in red) the detail of the deformation over the area that first detached on 24 June 2017 (Intrieri et al., 2018). (d) Perspective view of the whole landslide highlighting how most of the landslide body was experiencing no deformation before the failure with the exception of the area lined in red near the crest of the mountain.

cell across a range of spatial scales during the monitoring period. In turn, our method can accurately identify almost a year in advance the location of the Xinmo landslide, such that the outlyingness is further confirmed by the persistent spatiotemporal dynamics observed at this location from December 2016. We further validate the performance of our prediction method on the Stromboli rockfalls, showing that the rockfall affected areas can be effectively detected, which is beyond the capability of traditional methods applied to satellite data.

We emphasize that the proposed framework is a general approach for landslide monitoring and forecasting from large-scale spaceborne deformation data. It has broad potential applicability in different settings since no assumptions on the type of failure or the underlying trigger mechanism are needed. The prediction component of the framework delivers an early forecast of the *location and geometry* of the impending failure. Finally, predicting the time of failure is outside the scope of this study.

2. Study Areas and Data Used

We test our methodology on two very different landslides: the 2017 Xinmo rock avalanche (Figure 1) and the 2015–2016 Stromboli rockfall sequence (Figure 2). These two case studies represent optimal examples for the study of catastrophic collapse phenomena in rock bodies at the regional (Xinmo) and slope (Stromboli) scale, using satellite interferometric radar technique. The suitability and applicability of our method to other types of landslides such as soil-based failures are outside the scope of this investigation.

2.1. Xinmo Landslide

On 24 June 2017, a rock avalanche occurred on the mountain above Xinmo village, Sichuan, China, in a bedrock mainly composed of metamorphic sandstone and phyllite. Initially, a volume of 4.3 million m^3 mass detached

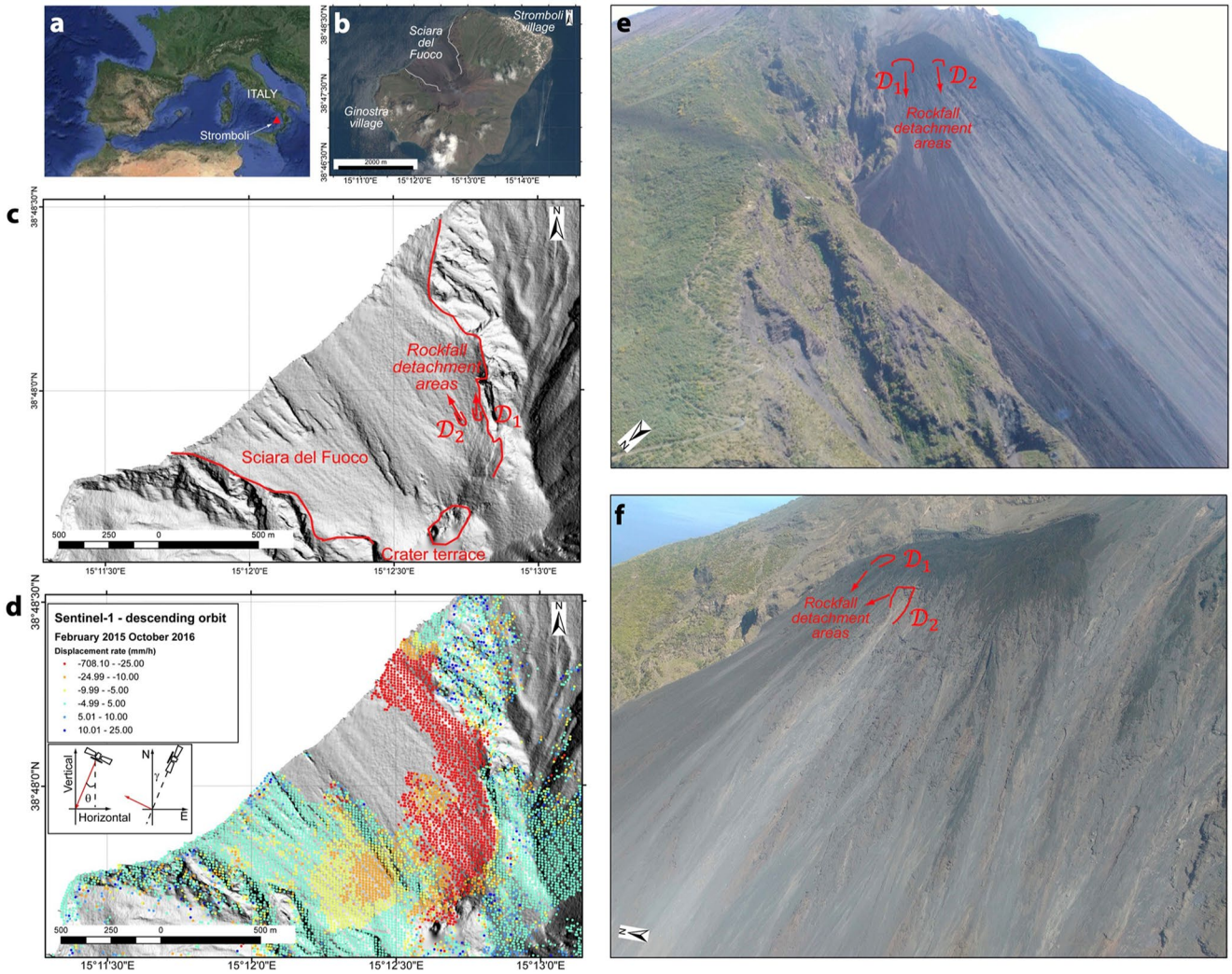

Figure 2. Stromboli rockfalls. (a, b) Geographic location of the test site. (c) Main geological-geomorphological features Stromboli. (d) Sentinel-1 (descending orbit) SqueeSAR displacement data (from Di Traglia, Nolesini, Solari, et al., 2018) during February 2015 to October 2016. (e, f) Photos taken by helicopter of the area affected by rockfalls observed from different points of view (photo: Federico Di Traglia).

from near the crest of the mountain and gained momentum along the slope. The volume of the unstable mass increased up to 13 million m^3 as the rock avalanche entrained new material until it eventually reached Xinmo village with an estimated velocity of 250 km/hr (Fan et al., 2017). This caused the death of 83 people and the destruction of 64 houses. The impact also generated a seismic shaking with a Richter magnitude $M_L = 2.3$ (Fan et al., 2017). The landslide also partially obstructed the Songping river. This area is known to be tectonically active and several strong earthquakes (with associated landslides) occurred in the last decades, such as the 1933 Diexi earthquake ($M = 7.5$), the 1976 Songpan-Pingwu earthquake swarm ($M = 7.2$), and the 2008 Wenchuan earthquake ($M = 8.0$). Detailed studies (Fan et al., 2017) revealed that partially interconnected cracks, probably generated by the 1933 Diexi earthquake, were present in the location of the source area. The trigger is to be brought back to prolonged rainfall increasing hydrostatic pressures within the cracks, possibly causing slow creep deformations leading to crack opening and propagation, and eventually to the mass collapse, which started as a large rockslide in the source area and then evolved into an extremely rapid (Varnes & Cruden, 1996) flow-like event, classifiable as a rock avalanche (according to the classification of Hungr et al., 2014).

2.2. Stromboli Rockfalls

Stromboli volcanic island is located in the Tyrrhenian Sea, off the northern coast of Sicily, Italy. The volcano collapsed several times, forming two depressions on its NW (Sciara del Fuoco; SdF) and SE flank, showing bilateral flank instability. The SdF depression is filled with volcaniclastic deposits and lavas, emitted from a summit crater terrace located at c. 750 m a.s.l., and from vents within the SdF. Present-day volcano slope instability comprises mobilization of coarse-grained and fine-grained sediments directly or indirectly related to the eruptive activity (Calvari et al., 2016; Di Traglia et al., 2020). Slope failures will produce a wide spectrum of mass movement phenomena, from small rockfall (continuously affecting the SdF) to (debris) rotational or (rock) planar slides, evolving in (rock/debris) avalanches (Di Traglia, Bartolini, et al., 2018) and eventually triggering tsunamis (Di Traglia, Nolesini, Ciampalini, et al., 2018). After the last flank eruption (August–November 2014), the gravitational readjustment of the lava breccia, sometimes evolving into rockfalls, has been detected using various methods (Di Traglia et al., 2020; Di Traglia, Nolesini, Ciampalini, et al., 2018; Schaefer et al., 2019). In particular, frequent rockfalls detached from two different source areas (D_1 and D_2, Figure 2), mostly in summer 2015 and summer 2016. During the summer, the rockfalls are more frequent due to phenomena of cyclic thermal stressing (Collins & Stock, 2016). In this period, the volcanic activity was low and most of the slope phenomena were not induced by the eruptive activity but by the slope dynamics of the volcano's flank (Di Traglia et al., 2020; Schaefer et al., 2019).

2.3. Sentinel-1 Data

The analyses carried out in this study use SAR images acquired by Sentinel-1 constellation in C-band (6.5-cm wavelength) with the Interferometric Wide swath mode, which collects data with a 250-km swath at 5-m × 14-m spatial resolution. The images are acquired along the descending orbit (incidence angle of 40.78°) and have been processed using SqueeSAR algorithm (Ferretti et al., 2011) to retrieve reliable displacement information from the monitoring region. The resultant Xinmo data spans from 9 October 2014 to 19 June 2017, i.e., 5 days before the failure, and comprises over 130,000 measurement points (MPs) which are spread across a 460 km^2 area. Of these, around 170 MPs (~0.1%) lie in the *landslide source area*: the region characterized by accelerating deformation in Intrieri et al. (2018). The extracted Stromboli data consist of 8,373 displacement time series from 23 February 2015 to 15 October 2016 (Di Traglia, Nolesini, Solari, et al., 2018), of which only around 20 MPs are located in the rockfall detachment areas shown in Figure 2c. Both Xinmo and Stromboli contain displacement data of 44 timestamps: the most frequent time interval in Xinmo is 24 days, whereas Stromboli has a better temporal resolution of 12 days.

3. Method

An overview of the proposed framework is given in Figure 3. There are two phases: *characterization* (Figure 3a) and *prediction* (Figure 3b). The characterization phase aims at finding unique features associated with the landslide source area. In this phase, we have displacement data of all MPs (over 130,000) in the monitoring domain and the knowledge of when and where the landslide occurred to train our model. The source area is a relatively small region of ~170 MPs that are characterized by accelerating deformation (Intrieri et al., 2018). The characterization phase outputs the optimal feature subspace (OFS) that best distinguishes the landslide source from other areas. This is achieved by analyzing the data at each timestamp as (a) partitioning the monitoring space into multiple grid cells; (b) deriving new features for each cell based the displacement data of its member MPs; (c) examining how distinctive each candidate feature subspace is via outlying aspects mining (OAM); and afterward, (d) summarizing the scores over all timestamps to obtain the final evaluation metric. Built upon the identified OFS, the prediction phase finds the temporally persistent outlying grid cell across a wide range of spatial scales during the monitoring period to narrow down the search perimeter and pinpoints the area where the landslide will likely occur. Two levels of predictions are generated in the prediction phase. The first is the *intermediate* prediction at each timestamp which represents the potential area of high landslide risk. Specifically, at each timestamp, each MP in the entire monitoring region receives its prediction support score ($\rho_i^{(t)}$), indicating its likelihood of failure. $\rho_i^{(t)}$ quantifies the proportion of times an MP lies in the most outlying cell, across all grid partitions of different spatial scales. The intermediate prediction is then given as the group of MPs with the highest prediction support score at that time. Temporal signals such as change of maximum prediction support score with the increase of time and the spatial distance between intermediate predictions of consecutive timestamps can be tracked. And

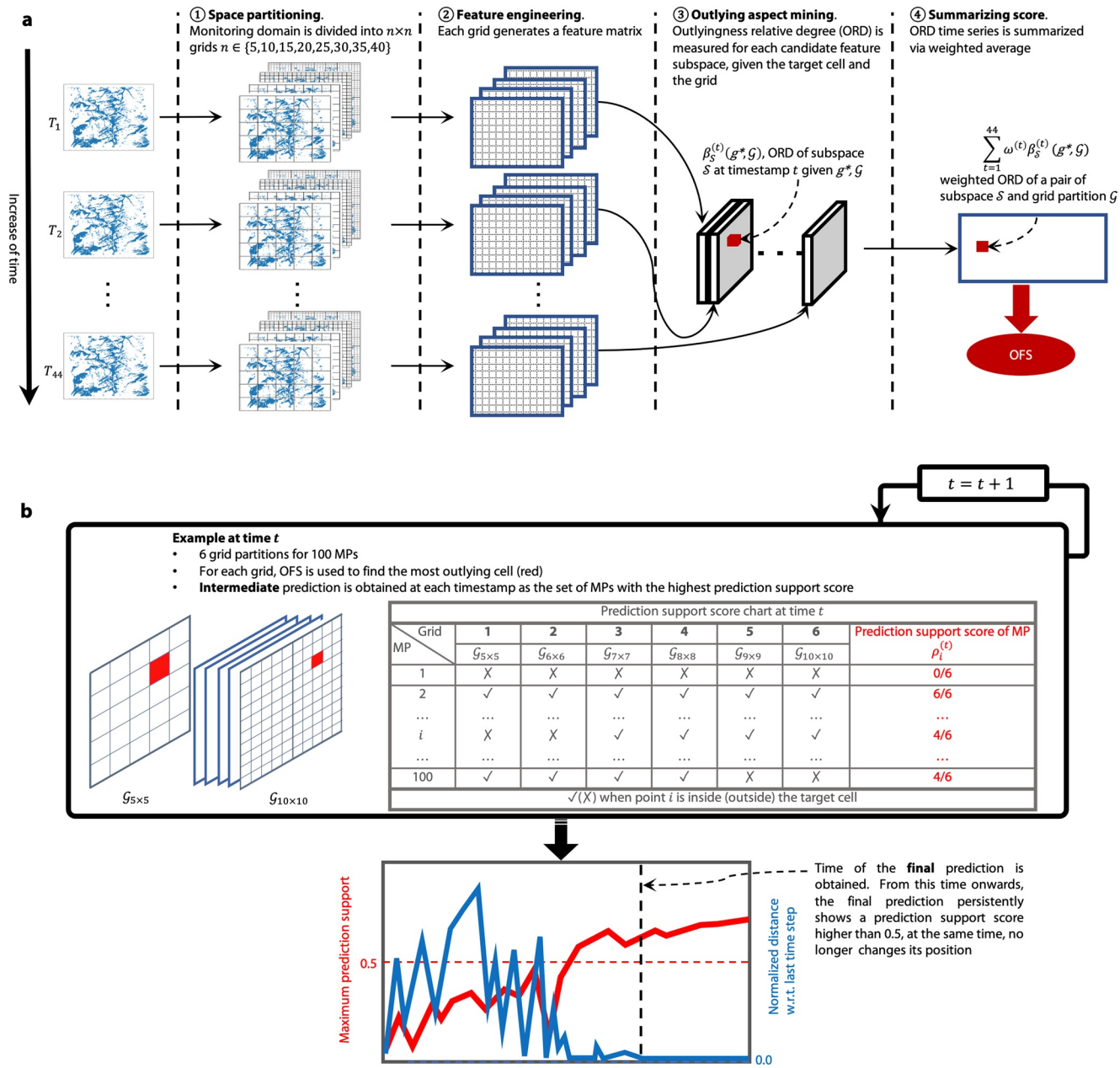

Figure 3. An overview of the proposed framework. (a) Characterization phase that identifies the optimal feature subspace (OFS) based on the Sentinel-1 data. (b) Prediction phase applies the identified OFS to predict the landslide location from the given imbalanced data. A toy example at time t with 100 measurement points (MPs) is used for illustration. Each monitoring point i is assigned a prediction support score $\rho_i^{(t)}$ that quantifies the proportion of times it lies in the most outlying cell, across all grid partitions. An intermediate prediction is given at t by outputting the set of MPs with maximum prediction support. Then the final prediction of landslide location is a product of two dynamical trends marked by a temporally persistent: high prediction support score above 0.5; and a low to zero value of normalized distance compared to earlier time stages.

the trends in these signals facilitates the identification of the second level prediction—the *final* prediction. As a general guide, we regard the final prediction as the intermediate prediction that is marked by a temporally persistent high prediction support score (>0.5), at the same time, showing no changes in its position. A list of notations used can be found in Table 1.

Table 1
List of Notations

Term	Definition
M	Total number of valid grid cells in a grid partition
N	Total number of features
I	Total number of measurement points (MPs)
J	Total number of grid partitions
g_i	The ith grid cell of a grid partition
g^*	Target cell that is the closest to the actual landslide source area
f_i	The ith feature
$\mathcal{G}$	Grid partition of M nonoverlapping cells $\{g_1, g_2, ..., g_M\}$
$\mathcal{G}_{i\times j}$	Grid partition that divides the monitoring space into an $i \times j$ grid
$\mathcal{F}$	The set of all features $\{f_1, f_2, ..., f_N\}$
$\mathcal{S}$	The feature subspace that is a subset of $\mathcal{F}$
$\alpha_{\mathcal{S}}^{(t)}(g_i, \mathcal{G})$	Outlyingness degree (OD): quantifies the outlyingness of cell g_i at timestamp t given the feature subspace $\mathcal{S}$ and the grid partition $\mathcal{G}$
$\beta_{\mathcal{S}}^{(t)}(g^*, \mathcal{G})$	Outlyingness relative degree (ORD): quantifies the outlyingness of subspace $\mathcal{S}$ at timestamp t given the target cell g^* and the grid partition $\mathcal{G}$
$\omega^{(t)}$	Weight of the ORD score at time t
T_{end}	The final timestamp
T_{hl}	Half-life parameter such that the weight of historical data is reduced every T_{hl} days
$\rho_i^{(t)}$	Prediction support score of the ith MP at timestamp t: the proportion of times ith MP lies in the most outlying cell, across all grid partitions
$\rho_{\max}^{(t)}$	Maximum prediction support score at timestamp t

3.1. Space Partitioning

The monitoring region is divided into multiple $n \times n$ uniformly sized grids, where $n \in \{5, 10, 15, 20, 25, 30, 35, 40\}$, such that at each timestamp eight grids are generated with different cell sizes. For each grid partition, the cell with the least inter-median distance to the landslide source is defined as the target cell and all other cells are the ones to be distinguished from the target cell. Grid cells in $\mathcal{G}_{5\times5}$ and $\mathcal{G}_{40\times40}$ are small areas of 18.4 and 0.3 km^2, respectively, which are similar in size to the regions studied in Intrieri et al. (2018) that can fully encapsulate the landslide source. To reduce the amount of computation, we exploit the fact that the MPs are not uniformly distributed in the monitoring region by eliminating those cells that contain <100 MPs (a number comparable to the average cell density for the finest grid $\mathcal{G}_{40\times40}$).

3.2. Feature Engineering

At each timestamp, a grid cell is characterized by a list of 32 features (Table 2) derived from the displacement time series. These features can be categorized into three types: (a) $f_1 - f_{15}$, basic statistics related to measurements at the current timestamp; (b) $f_{16} - f_{22}$, time-related features based on information over the last five timestamps (the same time interval as in Intrieri et al., 2018); and (c) $f_{23} - f_{32}$, clustering-based features assigning the MPs of each grid cell to two clusters—one cluster corresponding to the sliding region, the other to the stable cluster—following Zhou et al. (2020) and Tordesillas, Zhou, et al. (2020). Essentially, $f_1 - f_{22}$ capture the basic spatio-temporal dynamics in each cell while the clustering-based features encode more advanced insights demonstrated in Zhou et al. (2020). In their work, a comparative study has been conducted between the landslide source area and other potential risky regions that are experiencing instabilities to reveal distinctive dynamics that are unique in the landslide source. By performing a clustering analysis, Zhou et al. (2020) shows that while approaching to the failure time, larger and more consistent kinematic separation can be found between the stable and sliding clusters in the landslide source and a better linear fitting can also be found by applying the Fukuzono's method

Table 2
List of Used Features That Describe a Given Grid Cell

No.	Description
f_1	Mean displacement
f_2	Mean velocity
f_3	Mean acceleration
f_4	Standard deviation in displacement
f_5	Standard deviation in velocity
f_6	Standard deviation in acceleration
f_7	First quartile in displacement
f_8	Second quartile in displacement
f_9	Third quartile in displacement
f_{10}	First quartile in velocity
f_{11}	Second quartile in velocity
f_{12}	Third quartile in velocity
f_{13}	First quartile in acceleration
f_{14}	Second quartile in acceleration
f_{15}	Third quartile in acceleration
f_{16}	Cumulative displacement within the past five timestamps
f_{17}	Mean of average displacement within the past five timestamps
f_{18}	Standard deviation of average displacement within the past five timestamps
f_{19}	Mean of average velocity within the past five timestamps
f_{20}	Standard deviation of average velocity within the past five timestamps
f_{21}	Mean of average acceleration within the past five timestamps
f_{22}	Standard deviation of average acceleration within the past five timestamps
f_{23}	Mean displacement in the sliding region
f_{24}	Mean velocity in the sliding region
f_{25}	Mean acceleration in the sliding region
f_{26}	Silhouette score (Rousseeuw, 1987) w.r.t. displacement
f_{27}	Silhouette score w.r.t. velocity
f_{28}	Silhouette score w.r.t. acceleration
f_{29}	Clustering similarity compared to the past timestamp measured by normalized mutual information (Vinh et al., 2010)
f_{30}	Proportion of MPs in the sliding region
f_{31}	Slope of the linear regression fitted in the inverse of average velocity (Fukuzono, 1985) in the sliding region over the past five timestamps
f_{32}	r-squared of the linear regression fitted in the inverse of average velocity in the sliding region over the past five timestamps

(Fukuzono, 1985) over MPs in the sliding cluster. K-means (Lloyd, 1982), the most popular clustering algorithm that partitions data samples into K nonoverlapping clusters based on their distribution in the feature space, is used for each cell to obtain the clustering assignment based on their displacement at each timestamp. That is, the following minimization problem is solved to find the clustering assignment for MPs within the ith grid cell at time t, $C^{(t)}(g_i) = \left\{c_1^{(t)}, c_2^{(t)}, \ldots, c_K^{(t)}\right\}$ as

$$C^{(t)}(g_i) = \underset{C^{(t)}(g_i)}{\arg\min} \sum_{k=1}^{K} \sum_{x^{(t)} \in c_k^{(t)}} \left(x^{(t)} - \mu_k^{(t)}\right)^2, \tag{1}$$

where $x^{(t)}$ is the displacement of a single MP in g_i at time t, $c_k^{(t)}$ is the kth cluster in the cell, and $\mu_k^{(t)} = \frac{1}{|c_k^{(t)}|}\sum_{x^{(t)}\in c_k^{(t)}} x^{(t)}$ is the corresponding cluster center. The number of clusters, K, is set to 2 for all grid cells and all timestamps. Based on this clustering assignment, several features are extracted such as the Silhouette scores (Rousseeuw, 1987) that quantify the clustering quality with respect to different features, the normalized mutual information (NMI; Vinh et al., 2010) that measures the clustering persistence between consecutive timestamps, and the goodness of the linear fitness in the inverse velocity method (Fukuzono, 1985) in the identified sliding cluster and so on.

3.3. Outlying Aspects Mining

At each timestamp, we find the OFS that best characterizes the target cell g^*. That is, we explore a set of promising feature subspaces and measure how outlying g^* is in the given feature subspaces, to find the subset of features where g^* is a significant outlying object among all cells. Specifically, we use the same search mechanism as Duan et al. (2015), which brute-forces all feature subspaces up to a user-specified maximum dimensionality. We limit the maximum number of features a subspace can have to 3. That is, 5,488 feature combinations (from the 32 engineered features listed in Table 2) that have less or equal to 3 features are examined. This allows us to identify the feature subspace that characterizes the target cell well, while saving computation time.

An essential component of OAM is the scoring function that quantifies the outlyingness degree (OD) of objects in the given feature subspace. Without loss of generality, we assume that the lower the OD score, the more outlying the object is. We denote such scoring function by α such that $\alpha_{\mathcal{S}}^{(t)}(g_i, \mathcal{G})$ denotes the OD of a query cell g_i with respect to other grid cells in $\mathcal{G}$ given the feature subspace $\mathcal{S}$ at timestamp t.

In this work, we choose isolation path (iPath; Vinh et al., 2016) as the method for measuring the OD of a query cell. This method builds upon the isolation forest (Liu et al., 2008) outlier detection algorithm. In isolation forest, multiple binary random trees are constructed, with an inner node of a tree indicating a random split on a random chosen feature, while a leaf node standing for one single object that is isolated from the others. The iPath score of a query point is the average path length from the tree root to the corresponding leaf node. Intuitively, in a feature subspace where the query object is an outlier compared to the others, the query object is expected to be easily isolated via a small number of splits since they are far away from the majority, hence, has a lower iPath score.

The direct use of OD as the performance metric gives rise to difficulties in summarizing the values from different timestamps, since their value ranges can vary from one timestamp to another. To resolve this, instead of using the raw OD score, we propose the following outlyingness relative degree (ORD) score as the performance metric

$$\beta_{\mathcal{S}}^{(t)}(g^*, \mathcal{G}) = \frac{\left|\left\{g_i | \alpha_{\mathcal{S}}^{(t)}(g_i, \mathcal{G}) < \alpha_{\mathcal{S}}^{(t)}(g^*, \mathcal{G})\right\}\right|}{M}, \tag{2}$$

where $\left|\left\{g_i | \alpha_{\mathcal{S}}^{(t)}(g_i, \mathcal{G}) < \alpha_{\mathcal{S}}^{(t)}(g^*, \mathcal{G})\right\}\right|$ is the number of cells that have smaller OD scores than g^*, and M is the number of valid cells in $\mathcal{G}$. The ORD score lies between 0 and 1: the best (worst) performing subspace is assigned an ORD score of 0 (1).

3.4. Summarizing Scores

To answer the question on which is the best subspace that characterizes g^*, we propose to summarize the ORD time series, $\beta_{\mathcal{S}}^{(1)}(g^*, \mathcal{G}), \ldots, \beta_{\mathcal{S}}^{(t)}(g^*, \mathcal{G}), \ldots, \beta_{\mathcal{S}}^{(T_{end})}(g^*, \mathcal{G})$, of subspace $\mathcal{S}$ and grid partition $\mathcal{G}$, to a single score by the weighted average. We refer to this weighted outlyingness relative degree as the WORD score. An exponential weighting (Aggarwal et al., 2004) function is employed to control the relative importance between the historical and more recent information. This ensures that identifying the cell which overlaps with MPs in the landslide source area becomes increasingly important as the landslide time draws near. Formally, this can be expressed as

$$\omega^{(t)} = 2^{-(T_{end}-t)/T_{hl}}, \tag{3}$$

where $T_{end} - t$ measures how far away (in days) t is from the final timestamp (recall that T_{end} is known in the characterization phase), and T_{hl} is the so-called half-life parameter (Aggarwal et al., 2004). One can easily see that the weight of the final timestamp is 1, and the weights of previous timestamps are reduced by half every T_{hl} days. With the help of the time-weighted summary of ORD scores, one pair of feature subspace and grid partition

Table 3
Characterization Results on Different Grid Partitions

Grid partition	Avg. WORD	Precision	Recall
$\mathcal{G}_{5\times5}$	0.2393	0.4714	1.0000
$\mathcal{G}_{10\times10}$	0.1553	0.9706	0.9538
$\mathcal{G}_{15\times15}$	0.2329	0.9485	0.5318
$\mathcal{G}_{20\times20}$	0.1172	0.9821	0.9538
$\mathcal{G}_{25\times25}$	0.1584	1.0000	1.0000
$\mathcal{G}_{30\times30}$	0.0902	1.0000	0.4624
$\mathcal{G}_{35\times35}$	0.1469	1.0000	0.8439
$\mathcal{G}_{40\times40}$	0.1411	1.0000	0.9538

Note. Average WORD score among all feature subspaces quantifies the effectiveness in mining the outlying aspects (the lower, the better). Precision and recall validate the quality of each grid partition based on clustering its corresponding target cell (the higher the value, the better the performance).

now has a single WORD value as the performance metric. This enables a straightforward comparison among different grid partitions, since we can further summarize the performance of a grid partition as the average WORD score over its feature subspaces (or vice versa, when comparing subspaces by averaging over multiple grid partitions).

3.5. Predicting the Landslide Location Using OFS

The algorithm presented above is designed to identify the OFS that best distinguishes the target cell. The OFS contains up to three features from the list of 32 engineered features. Here, we discuss how to apply the discovered OFS to predict the source location of an unknown impending landslide.

A depiction of the proposed implementation is given in Figure 3b. At each timestamp, an intermediate prediction is obtained. Specifically, suppose there are I MPs in the whole monitoring region. Multiple grids of different spatial scales are applied to the monitoring domain and for each grid, the most outlying cell is found relative to the OFS and is assigned the label of *failure*, while all other cells are labeled *stable*. A prediction support score, $\rho_i^{(t)}$, is then assigned to the ith MP, $i = 1, 2, \ldots, I$, which is given by the ratio of the number of times the ith MP lies in the failure cell across all grids to the total number of grids analyzed, J. This procedure can be thought of as a system comprising J predictors (grid partitions) whose task is to predict the landslide location based on a "voting scheme". Each predictor votes the most outlying cell as its prediction based on the OFS and its particular spatial scale. The prediction support is the ratio of agreement among the J predictors and signifies the confidence level of the prediction. We denote the maximum prediction support among all MPs at time t by $\rho_{\max}^{(t)}$ while the intermediate prediction of the likely location of the landslide is defined as the set of MPs with support equal to $\rho_{\max}^{(t)}$.

Next, a second level examination is performed to produce the final prediction. This inspects the temporal persistence in two signals: the change of $\rho_{\max}^{(t)}$ with the increase of time, and the spatial distance between intermediate predictions of consecutive timestamps. The distance between two intermediate predictions is measured as the distance between their centers (the median positions among their member MPs), and we normalize the horizontal and vertical coordinates of all MPs into [0, 1] to ensure that this distance is bounded in $[0, \sqrt{2}]$. The final prediction is the region with persistent outlyingness. Among all intermediate predictions produced during the monitoring period, the one with a temporally persistent high prediction support (>0.5), at the same time, with close to zero changes in its position, is regarded as the final prediction.

3.6. Implementation

ORD is measured based on iPath using the Python scikit-learn package (Pedregosa et al., 2011) with the following default parameters: the number of trees in iPath is set to 100, and the half-life parameter T_{hl} for computing the WORD score is set to 30 days. To measure the clustering performance of the target cell in each grid partition, K-means is used with the number of clusters set to 2 by using the displacement at the corresponding timestamp as the feature, following Zhou et al. (2020). For predicting the landslide location, we use a total of 46 $n \times n$ grid partitions where $n = 5, 6, \ldots, 50$. A general guide is to use a wide spectrum of grid partitions with different sizes such that cells in the fine-grained grids have comparable size to the impending failure site (e.g., in Xinmo, a cell in the grid partition with $n \approx 40$ has similar size as the landslide source area), whereas coarse-grained grids introduce a boarder spatial coverage, to ensure that the prediction is robust to scale variations.

4. Results

4.1. Characterization Results on Xinmo

According to Table 3, even though the number of cells to be considered is much larger in the fine grid partitions, distinguishing the target cell from the others is a more challenging task for the coarse grid partitions since details of the kinematic differences among MPs are smeared out in these cases. For example, the average WORD

scores of $\mathcal{G}_{5\times5}$ (0.2393) and $\mathcal{G}_{15\times15}$ (0.2329) are nearly doubled, compared to $\mathcal{G}_{20\times20}$ (0.1172) and $\mathcal{G}_{30\times30}$ (0.0902). Overall, the top-3 best performing grid partitions are $\mathcal{G}_{20\times20}$, $\mathcal{G}_{30\times30}$, and $\mathcal{G}_{40\times40}$, in which it is relatively easy to distinguish the target cell from the rest. A further consideration is that grid partitions should be fine enough so that the developing separation between the sliding and stable zones inside the target cell can be detected by the clustering analysis (Das & Tordesillas, 2019; Singh & Tordesillas, 2020; Tordesillas et al., 2018; Tordesillas, Zhou, et al., 2020). This is quantified by precision and recall (Powers, 2011) based on the ground truth, which is identified by clustering the displacement within a small region around the landslide source at the final timestamp, 5 days before the landslide (Intrieri et al., 2018). With respect to precision, most grid partitions perform well: precision is around or above 0.95. The exception is $\mathcal{G}_{5\times5}$ (0.4714), where a group of MPs above the actual landslide area is erroneously clustered as the sliding cluster (see the clustering visualization in Figure 4a). This reinforces the earlier point that using too coarse a grid can blur the ground motion contrasts and compromise the accuracy of the prediction. With respect to recall, three grid partitions, $\mathcal{G}_{15\times15}$, $\mathcal{G}_{30\times30}$, and $\mathcal{G}_{35\times35}$, demonstrate significantly lower recall compared to others. This is because the actual landslide source is not fully encapsulated in the target cell of these grid partitions (Figure 4a): instead, the source overlaps with the target cell as well as its adjacent cells. That said, this "splitting" of the landslide source does not necessarily lead to poor performance in distinguishing the target cell from the other competing grid cells. In fact, among the eight grid partitions studied, the target cell in $\mathcal{G}_{30\times30}$ has the lowest WORD score. This suggests that the target cell in $\mathcal{G}_{30\times30}$, which encapsulates only 80 out of the 170 MPs in the landslide source area, is nevertheless well distinguished from all the other cells. In summary, the splitting of the landslide source should not lead to erroneous results, as long as a grid partition divides the entire domain into small enough cells such that MPs in and around the landslide region show clear kinematic contrasts.

Based on the aforementioned results, we limit the comparison among subspaces by using only the WORD scores produced by the top-3 best performing grid partitions: $\mathcal{G}_{20\times20}$, $\mathcal{G}_{30\times30}$, and $\mathcal{G}_{40\times40}$. The top-10 outlying feature subspaces are identified via averaging the WORD scores over the top-3 grid partitions (Figure 4b). Among them, the 75% percentile (or third quartile) of the displacement, f_9, is the most informative feature to use to distinguish the target cell from the others. When f_9 is used solely, the target cell is always identified as the most outlying cell after 19 May 2016, across all top-3 grid partitions. In addition to having the lowest average WORD score among all studied subspaces, f_9 also contributes to all the other top performing subspaces. When any of these other top subspaces are used, the target cell is consistently identified as the most outlying cell from 30 July 2016 onwards. Considering individual features presented in the top-10 outlying feature subspaces, apart from f_9, basic statistics in the spatial (f_1—mean displacement of a cell) and temporal (f_{17}—mean of average displacement of a cell within the past five timestamps) domains, and clustering patterns (f_{26}—clustering quality as measured by Silhouette score, f_{29}—clustering persistence as measured by NMI) are also important aspects that characterize the target cell (Figure 4c). Compared to f_9, however, the contribution of these features to the outlyingness of the target cell is limited. We illustrate this with an example comparison between f_{17} and f_9 (Figure 4d). As the landslide time draws near, the target cell progressively distances itself from all the other cells—mainly to the right—the direction of increasing f_9.

4.2. Distinguishing the Xinmo Landslide Source

In this section, we show the output of the proposed prediction algorithm on Xinmo by using the identified OFS. Strictly speaking, this can only be considered as a "proof of concept" for the prediction method, since we are employing the same data used in the characterization phase. More specifically, this analysis serves to demonstrate the unique spatiotemporal dynamics of the landslide source, and how this can be exploited to distinguish the source from the other moving unstable areas in this vast monitoring domain. The effectiveness of our prediction algorithm is validated on an entirely different sites, the 2015–2016 rockfalls on Stromboli.

Two regimes of deformation, before and after 30 July 2016, emerge during the monitoring period preceding the Xinmo landslide. Frequent changes of positions can be found in the intermediate predictions, i.e., the collection of MPs with the maximum prediction support at different timestamps, during Regime 1 (pre-30 July 2016), before converging permanently in the landslide source area in Regime 2 (post-30 July 2016). The intermediate prediction first appearing at 30 July 2016 accurately identifies the location of landslide source almost 1 year in advance of the time of failure—24 June 2017—and then emerges into the final prediction at 3 December 2016 (~7 months in advance), from then the maximum prediction support maintains a high value of above 0.5. Regime

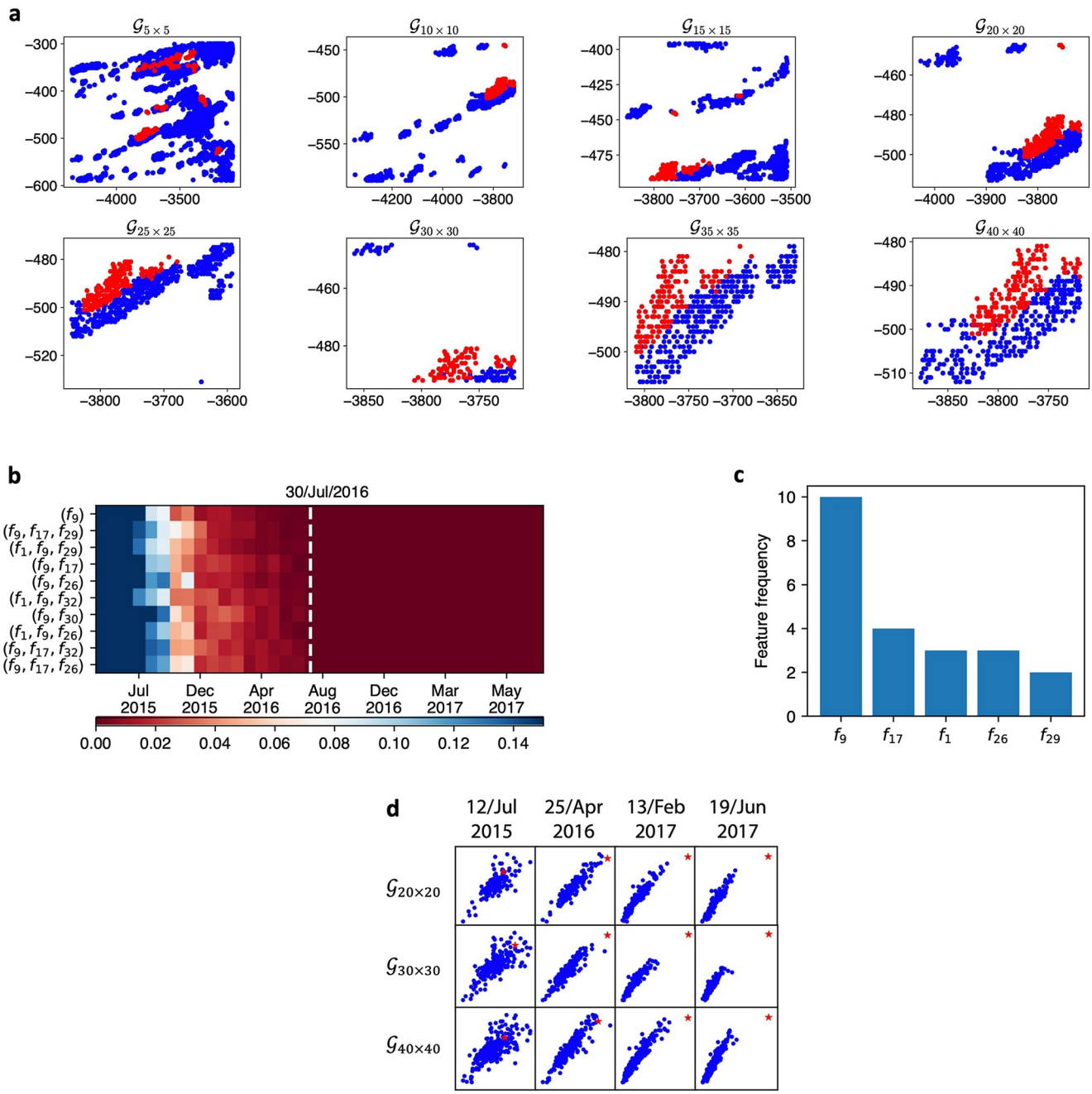

Figure 4. The third quartile in displacement, f_9, is the optimal feature subspace (OFS) to characterize the landslide source area in Xinmo. (a) Visualizations of K-means clustering on the displacement at the final timestamp for the target cell g^* with respect to different grid partitions. Red indicates the predicted sliding cluster, as determined by its higher mean displacement than that of the blue cluster. Note the red cluster in $\mathcal{G}_{25\times25}$ perfectly encapsulates the landslide source area in Intrieri et al. (2018) (see Figures 1c and 1d), since both the precision and recall are equal to 1. (b) Outlyingness relative degree (ORD) time series of the top-10 feature subspaces identified by iPath for the top-3 (out of 8) grid partitions. The lower the score, the better the performance. From top to bottom, the subspaces are ranked by their average WORD scores over the top-3 grid partitions. All top subspaces identify the target cell as the most outlying cell after 30 July 2016, which is almost 1 year in advance of the Xinmo landslide. (c) The five most frequent features in the top-10 subspaces. (d) Visualization of grid cells in the feature space defined by the two most frequent features f_9 (x axis) and f_{17} (y axis). For all grid partitions, the target cell (red star) is getting more isolated in the lead up to failure. Compared f_9 to f_{17}, the target cell distances itself mainly rightwards from all the other cell: f_9 pulls out the target cell better than f_{17}.

1 is marked by a low prediction confidence: $\rho^{(t)}_{\max}$ fluctuates below 0.5 (Figure 5b—red line). This concurs with the relatively low ground motions observed for most of the entire Sentinel-1 data set with none of the MPs manifesting persistent outlying motion. In laboratory tests of samples driven to failure and rockfall studies (Singh & Tordesillas, 2020; Tordesillas et al., 2018; Tordesillas, Zhou, et al., 2020; Zhou et al., 2020), this lack of a persistent

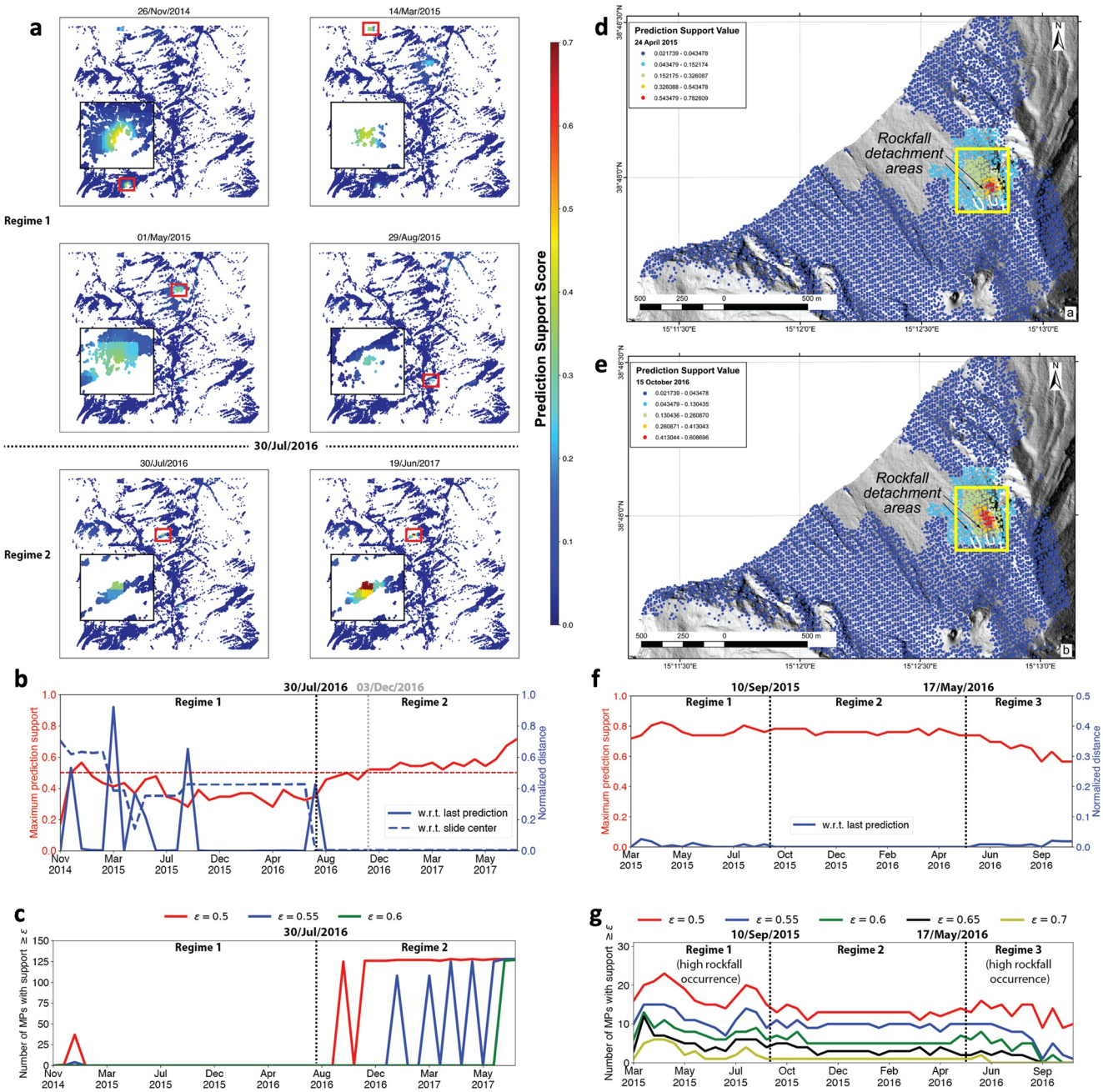

Figure 5. The proposed algorithm accurately identifies the location of the Xinmo landslide almost 1 year in advance and identified the location of rockfalls in Stromboli. (a) Prediction support score $\rho_i^{(t)}$ at different timestamps in Xinmo (the higher the value of a measurement point (MP) is, the more likely it is in the failure location). Within each subplot, the lower-left is a zoomed-in visualization of the red box, a small area around the predicted landslide region (MPs with $\rho_i^{(t)} = \rho_{\max}^{(t)}$), which overlaps the actual landslide source area (compare with Figures 1c and 1d) from 30 July 2016 onwards. (b) Changes in $\rho_{\max}^{(t)}$ and normalized distances with time in Xinmo. The intermediate prediction is attained at the correct position from 30 July 2016 onwards, with $\rho_{\max}^{(t)}$ increasing as the landslide time draws near. (c) The number of MPs with $\rho_i^{(t)}$ higher than a given threshold ε in Xinmo. More than 120 MPs consistently receive high confidence ($\rho_i^{(t)} \geq 0.5$) in Regime 2. (d, e) Prediction support score $\rho_i^{(t)}$ in Stromboli area on 24 April 2015 and 15 October 2016. Note that the 2015–2016 rockfall detachment areas fall within the zone characterized by higher $\rho_i^{(t)}$. (f) Changes in $\rho_{\max}^{(t)}$ and the normalized distance to previous timestamp with respect to time in Stromboli. (g) The number of MPs with $\rho_i^{(t)} \geq \varepsilon$ in Stromboli. The periods of observed higher rockfall occurrence are reported in Di Traglia, Nolesini, Solari, et al. (2018) and Schaefer et al. (2019).

and coherent pattern in outlying motion among a group of MPs is typical of the stable regime. As shown in Figure 5b (blue solid line), the spatial proximity in the intermediate predictions between consecutive timestamps spikes to a normalized distance as high as 0.9 (the spatial coordinates of MPs are normalized to a value range of 0–1, so the maximum normalized distance in $\sqrt{2}$). In a monitoring area of 460 km^2, this corresponds to a jump of

around 17 km. The dramatic changes in the intermediate predictions suggest Regime 1 is relatively stable. Indeed, no anomalous motions are identified in the actual landslide source area (Figure 5b—blue dashed line): the median position of MPs in the intermediate prediction never coincides with that of the landslide source area.

In contrast, after 30 July 2016, a group of MPs begin to manifest increasing outlying motions. This is evident in the essentially steady increase in $\rho^{(t)}_{\max}$ as more and more predictors, varying in spatial scales, identify MPs from this area to be in their most outlying cell. This sees a boost in $\rho^{(t)}_{\max}$ from 0.35 at the start of Regime 2 to 0.7 in the tertiary creep phase, 5 days before the landslide. Note the pinpoint accuracy and persistence of the prediction from 30 July 2016 with both inter-median distances to actual landslide source center and to the last prediction remaining at zero for Regime 2 (Figure 5b). The majority of high confidence scoring ($\rho^{(t)}_i \geq 0.5$) MPs in this intermediate prediction (Figure 5c) correspond to those identified to be in the landslide source using traditional time series analysis (Intrieri et al., 2018). Overall, findings suggest that the last intermediate prediction, which no longer changes in position from 30 July 2016 onwards, is manifesting outlying behavior across an increasing range of spatial scales as the landslide time draws near.

4.3. Robustness to False Alarms on Xinmo

As demonstrated above, our prediction algorithm temporarily identifies a few sites outside the actual landslide source area as the intermediate predictions during Regime 1. Let A1–A4 denote the small regions centered around these intermediate predictions at 26 November 2014, 14 March, 1 May, and 28 August 2015, respectively, and A* be the area around the true landslide source that is first detected at 30 July 2016. During Regime 1, accelerating motions manifest in A1–A4 and similar magnitude or even higher values in displacement or velocity can be frequently found, compared to the motion in A* (Figures 6a and 6b). For example, considerably higher displacements than A* can be found for A1 and A3 from the beginning of monitoring period to June 2015. The displacement in A4 is indistinguishable from A* until July 2016. During April 2015, the highest velocity can be found in A2, differentiating itself from the other areas. Given these anomaly deformation signals presented in A1–A4, conventional threshold-based approaches (Allasia et al., 2013; Barla & Antolini, 2016; Cigna et al., 2013; Crosta & Agliardi, 2002; Intrieri et al., 2012; Macciotta et al., 2016; Wang et al., 2010) can be easily misled and trigger false alarms, if the displacement/velocity thresholds were not carefully chosen, which mistakenly predict the active motions in these sites as landslide events, whereas in reality they have not developed into catastrophic failures. Unlike these methods which rely on subjective/expert decision-making, our framework provides an automatic and objective method for eliminating false alarms by examining their persistence in spatiotemporal dynamics (Figure 5b) in a way that leaves only the actual location of landslide as the final prediction. To further illustrate that A1–A4 are indeed false alarms where landslides do not occur, we perform an in-depth clustering analysis in these regions, along with A*, for the time period between the beginning of monitoring, to the timestamps when they are eliminated (i.e., when a different intermediate prediction is obtained by our algorithm).

Findings in Zhou et al. (2020) and Tordesillas, Zhou, et al. (2020) suggested that the small region around the landslide source exhibits distinct spatiotemporal dynamics that can be revealed by clustering the displacement of its member MPs. Specifically, while approaching the time of failure, clustering patterns in such small-scale region tend to converge to a stable solution such that the sliding area that is characterized by relative larger mean displacement no longer changes its position. At the same time, the kinematic separation among different clusters increases. In contrast, such distinct spatiotemporal pattern is missing in false alarm regions which briefly show instabilities but do not eventuate.

Consistent with these findings, our results in Figure 6 confirm that A1–A4 are false alarms without the unique spatiotemporal dynamics presented in the true landslide area (A*). Specifically, no clear clustering patterns in the spatial domain can be found in A1–A3. The fast-moving cluster is mixed with other MPs in the stable cluster. The spatial distribution of MPs in the active cluster of A4 and A*, on the other hand, is much compact and clearly separated from the stable cluster (Figure 6c). Different from the temporal signals in A4, A* demonstrates a unique increase of clustering persistence (NMI, Figure 6d) to above 0.8 from February 2016. During the same time, the kinematic separation between the active and stable clusters in A* is further enhanced (Silhouette score, Figure 6e). Such unique characteristics in A* validate the capability of our algorithm in eliminating false alarms and pinpointing the correct landslide source.

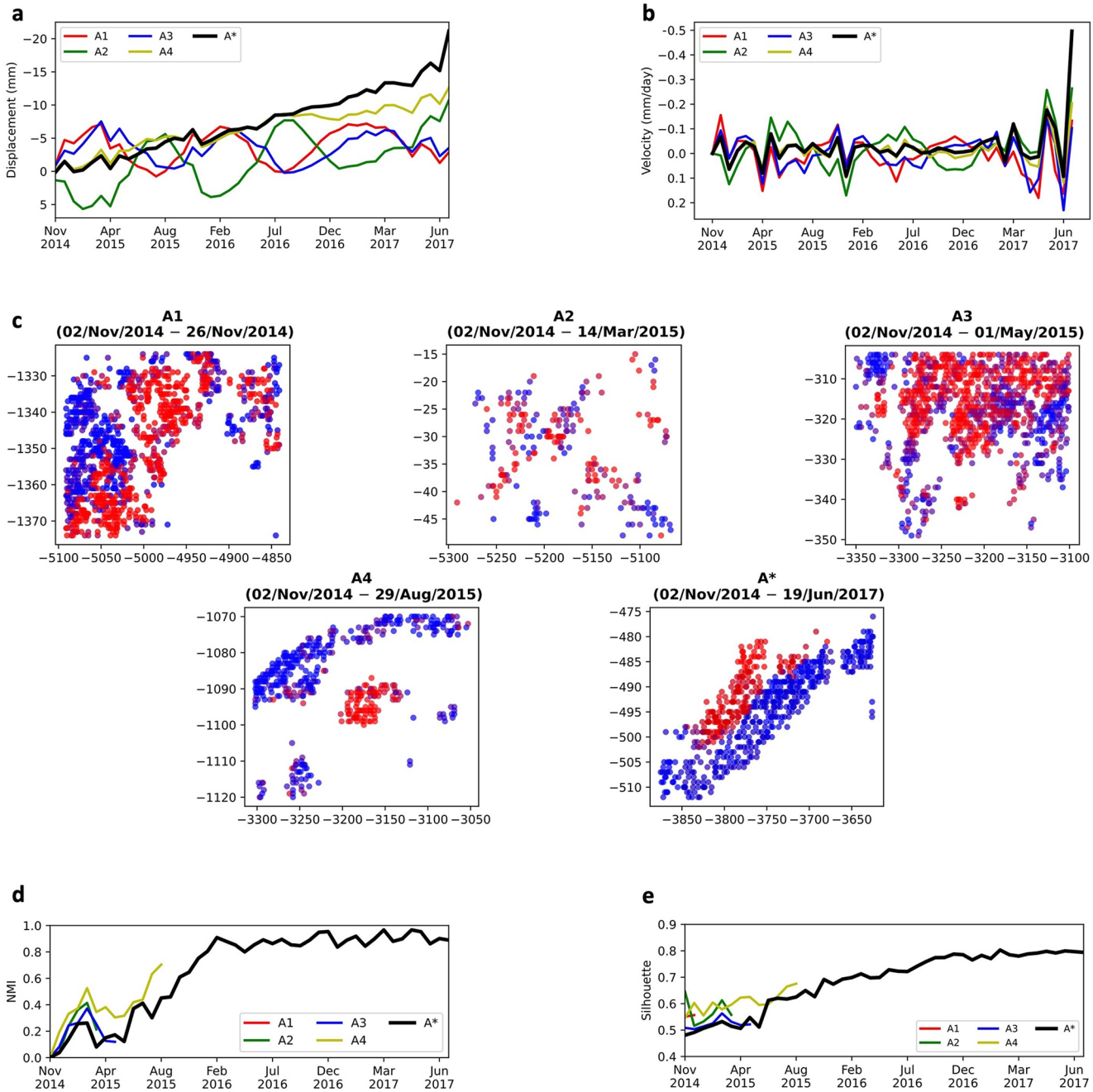

Figure 6. Clustering analysis performed on the areas of false alarm (A1–A4) and the landslide source area (A*) in Xinmo. (a, b) Average displacement and velocity among measurement points (MPs) within each area during the monitoring period. (c) Visualization of clusters generated by *K*-means for each site (accumulated from the beginning of monitoring to the timestamps when they are eliminated by our prediction algorithm). Red cluster is the fast-moving cluster with relative higher average displacement than the blue one. (d) Change of clustering persistence measured by normalized mutual information (NMI). (e) Change of clustering quality measured by Silhouette score.

4.4. Robustness to Parameters on Xinmo

The performance of our framework is insensitive to the change of parameters. For example, in the characterization phase, changes in parameters of the weighting function (3), which aims at controlling the relative importance between the historical and more recent information, may affect the order among different feature subspaces. We show that their effects are less significant. By stopping the summarization of ORD time series at an early timestamp (e.g., end of 2016, Figure 7a), f_9 remains as the highest ranked feature subspace among all candidates,

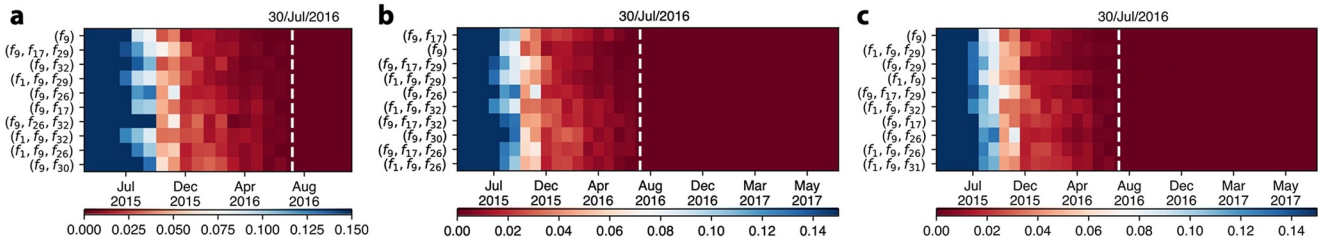

Figure 7. Robustness of parameters. (a) Top-10 feature subspaces when T_{end} is set to 27 December 2016. (b, c) Outlyingness relative degree (ORD) time series of the top-10 feature subspaces identified by using a half-life parameter in the weighting function (3) as 7 and 60 days, respectively.

meanwhile, presents itself in all other top-10 feature subspaces. Similarly, even though changing the half-life parameter from 30 days to 7 and 60 days (Figures 7b and 7c) slightly alters the order among top performing feature subspaces, compared to Figure 4b, we still see several subspaces repeating themselves in the top-10 best performing subspaces (e.g., (f_9), (f_9, f_{17})). More importantly, f_9 remains constantly appearing in all of the top performing feature subspaces, proving its general optimality in distinguishing the target cell from the other competing regions.

4.5. Prediction Results on Stromboli

The Stromboli data span a long analysis period of 2 years, and covers several seasons for the same environmental conditions that apply to the whole slope. Displacement values and displacement rates are high in many other areas of the slope (recall Figure 2d). In the period considered, our model does not identify false alarms in the other areas of the slope, but only where the collapses actually occurred. The intermediate predictions from our algorithm correspond to the area that is characterized by strong erosion in Di Traglia et al. (2020) (Figures 5d and 5e). The area of subaerial part of the flank of the volcano is about 10^6 m^2, while the rockfall detachment areas are smaller (10^3–10^4 m^2), and they can occur throughout the slope (Di Traglia, Bartolini, et al., 2018). Moreover, the physically based analysis of the slope stability identifies many areas characterized by greater instability (mainly due to the high topographical gradient of the slope; Schaefer et al., 2019). Therefore, it is impossible to know a priori where collapses will occur, while the method proposed here is able to identify the exact location of the block release areas. To our knowledge, detecting such regions is beyond the capability of traditional methods on satellite data, especially when the monitoring domain encompasses many fast-moving areas that are subject to thermal contraction and movement, and the area affected by rockfalls is not the one characterized by the highest displacements. Generally, $\rho^{(t)}_{\max}$ remains relatively high throughout the monitoring period (above 0.7 until 5 June 2016 to just below 0.6 at the end; Figure 5f). The normalized distance in intermediate predictions between consecutive timestamps is under 0.03, suggesting the instabilities persisted while remaining localized to a small fraction of MPs (around 20 MPs out of the total of 8,373 MPs). The temporal pattern of high confidence $\rho^{(t)}_i \geq 0.5$ MPs are consistent with the high rockfall frequencies reported in Di Traglia, Nolesini, Solari, et al. (2018) and Schaefer et al. (2019), especially in two of the three observed regimes of deformation (Figure 5g). Regime 1 (23 February 2015–10 September 2015) is when the most outlying ground motions can be observed among multiple unstable sites (MPs) that likely interact (Singh & Tordesillas, 2020; Zhou et al., 2020). Regime 2 (10 September 2015–17 May 2016) differentiates itself with no change to the intermediate prediction, which usually indicates the increasing risk of collapse in that location (Das & Tordesillas, 2019; Singh & Tordesillas, 2020; Zhou et al., 2020). Indeed, a brief period of activity ensues during 17 May to 5 June, at the start of Regime 3 (17 May 2016 to 3 October 2016), when the number of high confidence MPs see a small rise to a peak, before MPs in the region decelerate for the rest of the monitoring period.

4.6. Validation of Results on Stromboli

Similar to the analysis performed in Section 4.3, we further validate the effectiveness of our method by performing the clustering analysis on a narrowed-down region around the intermediate predictions generated by our algorithm (the yellow boxes in Figures 5d and 5e). The group of MPs with highest prediction support $\rho^{(t)}_{\max}$ matches the location of red cluster in K-means (Figure 8a), which is the cluster that demonstrates the highest instability compared to other clusters. Additionally, as shown in Figure 8b, the clustering persistence (similarity of

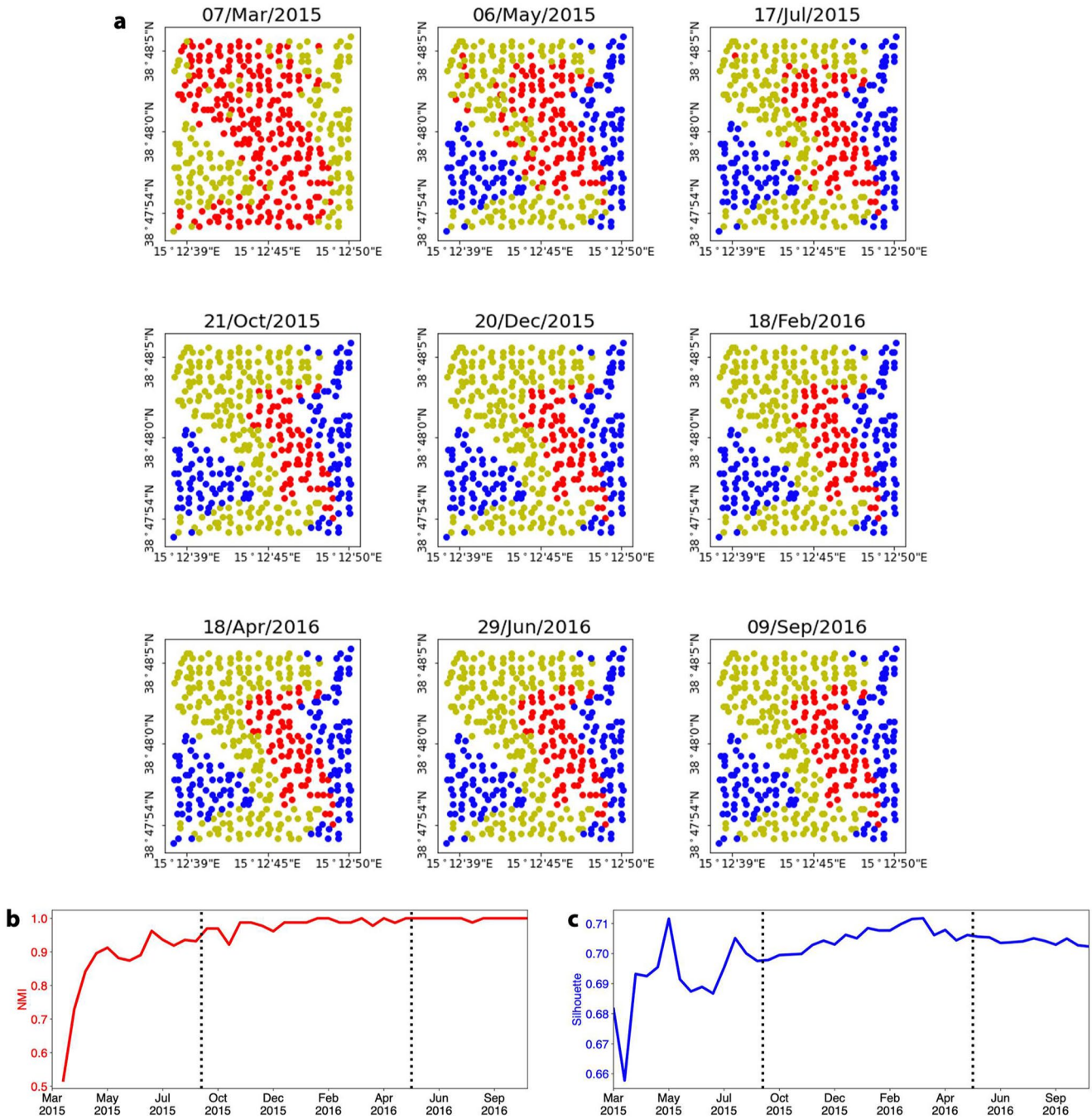

Figure 8. Clustering analysis on the narrowed-down region around the intermediate predictions during the monitoring period of Stromboli. (a) Visualization of clusters generated by *K*-means at different time stages. The colors represent the clustering membership, where the mean displacement in each cluster is decreasing from red to yellow to blue. (b) Change of clustering persistence measured by normalized mutual information (NMI). (c) Change of clustering quality measured by Silhouette score. The dashed lines in (b) and (c) mark the beginning of Regime 2 (10 September 2015) and Regime 3 (17 May 2016).

clustering assignment between consecutive time steps measured by NMI) dramatically increases from 0.5 to above 0.9 during Regime 1, suggesting that the failure is gradually developing during this phase. From the beginning of Regime 2, the clustering assignment generated by *K*-means barely changes over time, with the area above the rockfall detachment area, D_1, being always marked as the most unstable cluster that is likely to collapse. The clustering quality (kinematic separation among clusters measured by Silhouette score in Figure 8c), on the other hand, highlighted a difference between Regimes 2 and 3. In Regime 2, a steady increase in Silhouette can be found to

its peak at around 0.71, followed by a drop during the transition to Regime 3. This suggests an increasingly high likelihood of collapse in the unstable cluster at around March 2016, which is supported by the high rockfall frequency from that time, as reported in Di Traglia, Nolesini, Solari, et al. (2018) and Schaefer et al. (2019). Overall, the clustering analysis performed in the narrowed-down space confirms that the proposed prediction algorithm can correctly identify the location of rockslide via directly analyzing the motion information of MPs within the entire monitoring domain captured by satellite and finding the most outlying area across multiple spatial scales.

Ideally, one should use the OFS associated with a specific site (identified by applying the characterization phase on the historical failure events on the target site) as the input feature for the prediction model. However, in reality such information may not always be available. We show that under this circumstance, the specific OFS obtained from Xinmo, f_9, which captures the key deformation dynamics in the vast amount of monitoring region, can be an effective surrogate of the OFS for the new site. We use Stromboli to verify that this is true via statistical hypothesis testing, where the null hypothesis is: *f_9, the input of prediction model, is not indicative of the key deformation dynamics on Stromboli*. If we have significant statistical evidence to reject this null hypothesis, we would have high confidence on using f_9 for prediction. Similarly, the output of the prediction on Stromboli can also be verified. Specifically, we design the following tests:

- Test 1: input validation via permutation test (detailed in Appendix A). For a given timestamp t and a specific grid partition $\mathcal{G}$, the null hypothesis above is equivalent to saying that the f_9 values of cells in $\mathcal{G}$ are distributed homogeneously over the monitoring domain. We choose the sample variance in f_9 as the test statistic. Inherently, if the null hypothesis holds true, there will be a very small chance of this sample variance deviating from its presumed probability distribution under the null hypothesis. We generate permutations of the test statistic values to quantify this small chance. Specifically, 10% of MPs are randomly selected and their positions are randomly shuffled to form a permutation of all MPs; this is repeated 200 times to get 200 permuted data sets. A test statistic value is calculated based on each permuted data set, resulting in an empirical distribution of the test statistic, from which the p-value of the test can be calculated as the proportion of permutations whose test statistic values are larger than or equal to the test statistic value of the original sample. The p-value gives the observed significance of rejecting the null hypothesis. The lower the p-value, the stronger evidence we have to reject the null hypothesis. If p-value <0.05, we say the null hypothesis is rejected at 0.05 level of significance, i.e., at 95% confidence level
- Test 2: output validation via one-tailed t-test. At a given timestamp t, let MPs in the monitoring domain be divided into two groups: the high and low risk regions as determined by their prediction support scores with a threshold value of 0.5. Here, the null hypothesis is equivalent to saying that there is no significant difference between the movements in these two groups. One-tailed t-statistic is used for the testing, and the null hypothesis is rejected if the mean displacement in the high-risk group is significantly greater than that of the low-risk group

For Test 1, as demonstrated in Figure 9a, the test statistic from the original data frequently locates on the right-hand side of the empirical distribution constructed from the permutations, indicating extremely high variance in f_9 over all cells, suggesting nonhomogeneously distributed feature values in the spatial space. Additionally, the trend in the p-values conforms to the evidence uncovered by the prediction support score: close to zero p-values can be found in Regime 1 where a relatively large number of MPs demonstrate high prediction support, whereas slightly higher p-values (still <0.05 in most cases) apply in the later timestamps, especially in Regime 3 where the maximum support decreases (Figure 9b). Results from Test 1 give strong statistical evidence to reject the null hypothesis, showing that f_9 is a feature that carries meaningful spatial structure. In terms of Test 2, we see close to zero p-values throughout the monitoring period (Figure 9c), suggesting that for all timestamps, the group of high-risk MPs ($\rho_i^{(t)} \geq 0.5$) demonstrate significantly higher movement than the other regions, thus validating our prediction algorithm.

5. Discussion

5.1. A "Zoomed-In" View for Balanced Slope Failure Data

Pinpointing the location of an impending failure from space is a challenging task. To understand this, it is instructive to recall recent lessons learned on the fundamental dynamics of precursory failure for a balanced monitoring data. Having a balanced data on failure means essentially that a "zoomed-in" view of the unknown failure

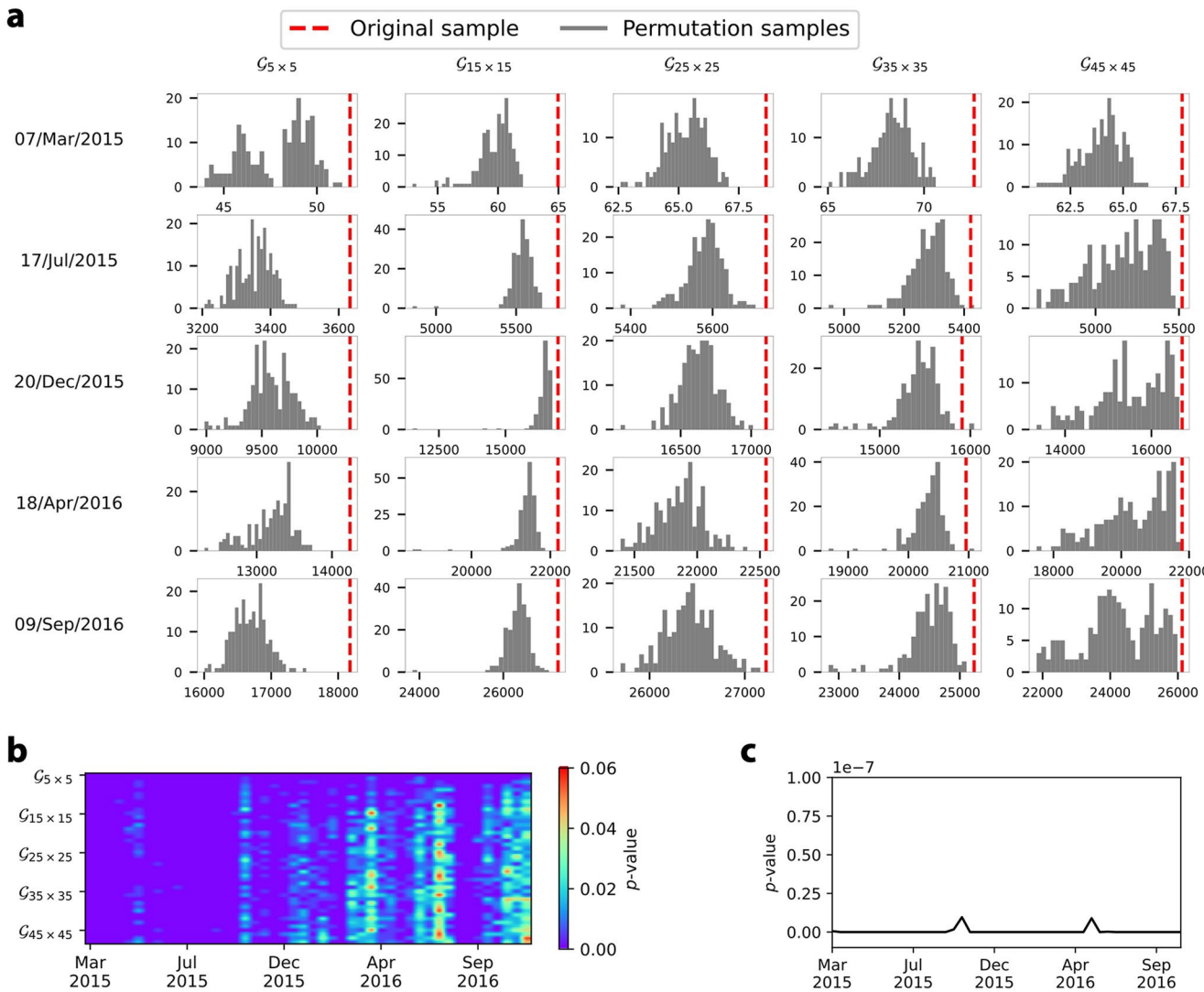

Figure 9. Further validation on Stromboli rockfalls via statistical tests. (a) Test statistics for Test 1. For each subplot, *x* axis gives the test statistic values and *y* axis is the frequency of each bin in the histogram. Red dashed line indicates the test statistic for the original sample. (b) *p*-Values for Test 1. (c) *p*-Values for Test 2.

location was captured. That is, the sizes of the unstable and stable regions are comparable so that the number of monitoring points is not significantly different for the two regions. Prior studies of failure in laboratory samples and of landslides/rockfalls in the field scale, based on kinematic data from simulations and physical measurements (Singh & Tordesillas, 2020; Tordesillas et al., 2021), uncovered a nontrivial spatiotemporal clustering pattern in the precursory failure regime.

In slopes, MPs in the impending failure region distinguish themselves near the failure time by moving in unison at increasing displacements (Singh & Tordesillas, 2020; Tordesillas et al., 2021). As a result, machine learning clustering algorithms like *K*-means can readily identify the ultimate failure region in the terminal stages when failure is imminent. More importantly, prior to this time, the spatiotemporal dynamics of this clustering pattern tracked accurately the evolution of failure and gave clues to the ultimate failure location. Chronologically, when the whole granular mass/body is stable, this clustering pattern is essentially nonexistent. But in the subsequent regime, i.e., the precursory failure regime, when instabilities develop and spread in the granular mass, a clustering pattern emerges. This pattern is transient initially but then ultimately persists in space and time close to and during failure. As reported in recent studies (Singh & Tordesillas, 2020; Tordesillas et al., 2021), such spatiotemporal clustering pattern has been discovered in a wide range of failures, despite their scales (laboratory versus slope versus field) and types (brittle versus ductile).

The spatiotemporal dynamics of kinematic clustering can be directly related to the fundamental physics of granular media deformation and failure, especially to the process of force transmission (Tordesillas et al., 2021; Tordesillas, Kahagalage, et al., 2020). Forces in granular bodies are mainly transmitted along preferred paths which are known as force chains. Force chains are akin to highways in a road network. Like any transmission network with preferred paths, some links in the network become prone to congestion thus forming bottlenecks. That is, compared to links elsewhere, the links in the bottleneck transmit forces close to their *capacity* which is the force that needs to be overcome to break the link. In Tordesillas, Kahagalage, et al. (2020); Tordesillas et al. (2021), it was shown that the path that percolates through the body containing such vulnerable links, i.e., force bottleneck, is in fact the path of least resistance to failure and is where the macrocrack in small samples eventually developed at failure. In monitoring data, this path delimits the landslide boundary at the time of collapse (Tordesillas et al., 2021).

The force bottleneck is an emergent structure that is highly dynamic in the prefailure regime. This is because a stable granular body is a highly redundant structure. Such a body has many paths available to accommodate stress reconfigurations that then relieve stress build up in the vulnerable bottleneck sites. This has the effect of diverting damage away from the preexisting bottleneck, a process that generally leads to a shift in the location of the bottleneck. However, as damage spreads, and the time of failure draws near, redundancies in transmission paths diminish until damage can no longer be avoided in the bottleneck area. This then marks the time when the bottleneck becomes spatially persistent, signaling the onset of imminent failure. Previously disconnected cracks begin to propagate rapidly and coalesce along the persistent bottleneck which is now becoming physically incised in the body. At the same time that material connections along this path break apart, the region of impending failure becomes increasingly detached from the rest of the slope. MPs in this active region will start to distinguish themselves by moving as a coherent whole (in near rigid-body motion) at increasingly higher velocities than those of the MPs in the stable zone. This explains why a persistent kinematic clustering can be observed in the displacement-state-space, with the persistent bottleneck delimiting the boundary of the landslide at the time of failure (Tordesillas et al., 2021).

5.2. A "Zoomed-Out" View for Imbalanced Sentinel Data

Observing the above processes from space, however, is like "zooming-out" from the target area. The kinematic signal from the failure location is drowned out by the signal from all the other MPs which comprise the vast majority of MPs in the analyzed region. Keep in mind that the total number of MPs generally increases with the growth of monitoring region, while the number of MPs in the failure location remains fixed. Eventually, if we zoom-out far enough, the fraction of MPs in the failure location relative to the whole monitoring region may dwindle down to <1%, as is the case in the Xinmo data. The failure location is now significantly more difficult to pinpoint. The method of deploying ordinary clustering algorithms in raw displacement no longer works in these scenarios, as the signal from the target area is drowned out by the vast amount of information from all the MPs, leading to inaccurate clustering and wrong prediction of the failure location. A further complication arises when the monitoring region encapsulates multiple unstable sites which similarly exhibit large movements, but which do not lead to catastrophic failure. In fact, over the past few decades, the monitoring area of the Xinmo data has been a site of multiple historical seismic events (Fan et al., 2017; Intrieri et al., 2018). The recent study in Tordesillas, Zhou, et al. (2020) also confirms that multiple sites exhibit comparable large movements to that in the actual landslide source area. Also, as demonstrated in Figures 5a and 6, in the early stages of the monitoring period, other sites were identified as areas of high risk of failure, before the prediction settled to the true landslide source area with high confidence (Figure 5b).

To address the issues relating to data imbalance, we shifted our analysis from the MP level—as deployed on past studies of balanced data (e.g., Das & Tordesillas, 2019; Singh & Tordesillas, 2020; Tordesillas, Zhou, et al., 2020; Zhou et al., 2020)—to the cell level. In the characterization phase, we treat the target cell (see definition in Section 3.1) that encapsulates the landslide source as a distinct object and compare it to the other cells in the monitoring domain, in order to uncover the unique features associated with the failure location. This is achieved by applying our outlying aspect mining algorithm to identify the OFS that maximizes the outlyingness of the target cell. Our strategy explores a vast set of features derived from the raw displacement data which by itself can no longer reliably identify the failure location (recall Figure 3a, and Sections 3.1–3.4). In the prediction phase, we designed an algorithm (Figure 3b, Section 3.5) that can predict the location of the impending failure region

by first ranking the MPs according to the prediction support score—the number of times an MP is identified in the most outlying cell across all "zoom levels" (grid cell sizes)—then producing the final prediction based on the temporal persistence in signals extracted from our analysis.

5.3. Summary of Key Findings

We summarize the key findings from our study as follows:

1. Our study confirmed that the ground deformation data of sufficient spatial and temporal resolution encode rich dynamics underlying the catastrophic slope failure and can be a valuable source for landslide monitoring and prediction
2. Our characterization procedure revealed that the OFS, 75% percentile of the displacement within a given cell, is a distinguishing characteristic that differentiates the landslide source area from the vast amount of region monitored by Sentinel-1. This is a cheap statistic to compute given the ease of extracting ground motion data from satellite images
3. Results from the characterization step confirmed that standard clustering approaches which are designed for slope-scale data can be unreliable in distinguishing the dynamics of the landslide source in imbalanced satellite data, given the inadequate performance of clustering-based features. Instead, by shifting the analysis from MP level to cell level, we found that the relative outlyingness in the motion among cells of different spatial scales has become a more trustworthy indicator in identifying the location of an impending failure
4. A prediction procedure was proposed based on the OFS for identifying the temporally persistent outlying cell across a range of spatial scales during the monitoring period. The simplicity of the identified OFS enhances the efficiency of the proposed prediction procedure. Typically, only a few seconds on a standard desktop PC are needed to deliver a prediction. This is near-real-time prediction, especially compared to the temporal resolution of several days in Sentinel-1
5. The proposed prediction procedure is straightforward to implement in practical early warning systems and has broad applicability in identifying potential landslides location in different settings. Specifically, the temporal signals in the changes in maximum prediction support, and the spatial distance to the previous intermediate prediction can be useful indicators in determining the final location of an impending collapse
6. The predictions deliver more than susceptibility maps because they are informed by continuous monitoring and provides indications not just of potential instabilities but of regions that are approaching imminent failure. Such warning-oriented information can provide actionable intelligence to early warning systems that are based on deformation rate thresholds. In conjunction with the fast increasing global availability of InSAR data at short time intervals, our method may also provide the missing paradigm to switch from susceptibility mapping (spatial prediction) to spatiotemporal prediction mapping at the regional scale, a capability that remains elusive
7. Our analysis shed new light on the use of regional-scale satellite deformation data to advance existing landslide-monitoring practices. Our method can be considered as a new alternative that is easy and cheap to apply, especially when other types of geophysical failure indicators are not readily available. At the very least, it can identify a potential hazard area from the vast amount of monitoring space, for direct field inspection, or for the application of targeted forecasting algorithms designed for smaller and more balanced data such as pointwise time-series inverse velocity methods (Intrieri et al., 2018)
8. The proposed framework is a general toolkit of potentially broad applicability and transferability. Although publicly available failure data are rare, agencies engaged in design and management of early warning systems have a growing repository of data on past slope failures for different sites. With the rapid development in remote sensing and Earth observation systems, we expect for this trend to continue unabated. Agencies may not openly share these historical records of failure for understandable reasons, but they do exist. One example of this is in open pit mining (e.g., Bingham Canyon mine; Gaida et al., 2021). This is precisely why we designed our platform to be one complete package—with the ability to accommodate the changing conditions that may intensify certain landslide triggers (e.g., climate change that may lead to increased intensity, duration, and frequency of rainfall (Johnston et al., 2021; Justice, 2020), earthquakes (Masih, 2018, etc.). Specifically, if historical data on past failure events on the vulnerable sites are available, we recommend that our entire platform be implemented for improved decision-making. That is, we recommend that both characterization and prediction components be implemented on historical failure data—especially on the most recent failures to update the OFS, as illustrated in Figure 10

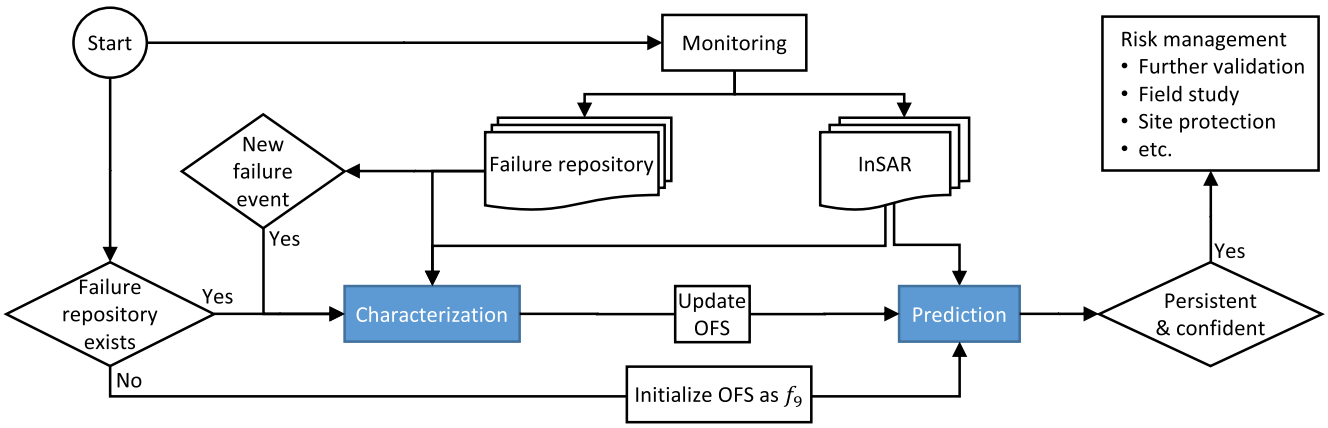

Figure 10. Integrating the proposed framework into real-world landslide monitoring and early warning systems.

6. Conclusion

We developed a method for early detection of impending failures from regional-scale Sentinel-1 data and tested its applicability using data on the 2017 Xinmo landslide and the 2015–2016 Stromboli rockfall. Both data suffer from severe imbalances, viz the fraction of measurement points in the failure region is disproportionately small with respect to the total number of points in each data set. This challenges the applicability of existing clustering-based approaches that are designed for balanced data, since the clustering dynamics in the failure location become less significant, as drowned out by the vast number of MPs in the monitoring region. To address this issue, we designed a procedure based on feature engineering and outlying aspects mining to systematically and objectively identify the optimal feature subspace (OFS). One particular feature—the third quartile in the displacement within a small area—characterized the dynamics in the impending failure region the best. This suggests a potential utility of the identified OFS in standard operational monitoring and early warning of landslides based on the Sentinel-1 data. To demonstrate this, a prediction algorithm was proposed to forecast the failure location by identifying the temporally persistent outlying cell across a range of spatial scales during the monitoring period. For Xinmo, the proposed method accurately finds the location of the 2017 landslide as early as 1 year in advance of the event. For Stromboli, the area subject to rockfall, excluding the areas that are instead subject to thermal contraction and movement, are not generally identifiable from satellite data using traditional methods. However, we have shown that the proposed method is able to detect the area despite the fact that there are many moving regions and that the one affected by rockfall is not the area characterized by the highest displacements at most timestamps during the monitoring period. In addition, even though it is difficult to detect each single collapse given the spatial and temporal resolution of Sentinel-1 data, our method is able to identify the periods of highest frequency of rockfalls.

With the anticipated improvements in the spatial and temporal resolution of spaceborne InSAR sensors in this new decade, our approach has the potential to deliver timely and accurate intelligence in aid of efforts toward near-real-time early warning and risk mitigation of landslide hazards.

Appendix A: Permutation Test

The purpose is to test the null hypothesis that the OFS, f_9, distributes homogeneously over all cells in the monitoring domain. Let $\mathcal{R}$ be the monitoring domain with N locations, and $x(s), s \in \mathcal{R}$ be the displacement of the sth location. We apply an approach of spatial random permutation test (Good, 2005; Haining, 2001) to investigate this homogeneity issue, and our spatial random permutation test procedure is described as following.

1. Partition $\mathcal{R}$ into an $I \times J$ matrix grid $\mathcal{G}_{I\times J}$. Calculate f_9 for each grid cell from all $x(s)$ observations in that cell. Denote the results as $\{f_9(i,j),(i,j) \in \mathcal{G}_{I\times J}\}$. Compute $\bar{f}_9 = (IJ)^{-1}\sum_{ij} f_9(i,j)$ and the sample variance in f_9 among all cells

$$X_0^2 = \frac{1}{IJ-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(f_9(i,j)-\bar{f}_9\right)^2$$

2 Draw $K = 200$ permutations from $\{x(s), s \in \mathcal{R}\}$ as following. For each $k = 1, \ldots, K$

(a) Randomly choose a proportion p, e.g., $p = 10\%$, of all N locations in $\mathcal{R}$. Denote the selected locations as $\{s_1, s_2, \ldots, s_{[pN]}\}$

(b) Generate a permutation of $\{s_1, s_2, \ldots, s_{[pN]}\}$. Denote the result as $\left\{s_1^{(k)}, s_2^{(k)}, \ldots, s_{[pN]}^{(k)}\right\}$

(c) Construct the kth permutation sample $\left\{x^{(k)}(s), s \in \mathcal{R}\right\}$ as following: Set $x^{(k)}(s) = x(s)$ if $s \notin \{s_1, s_2, \ldots, s_{[pN]}\}$; set $x^{(k)}(s_\ell) = x(s_\ell^{(k)})$ for $\ell = 1, \ldots, [pN]$

Note if the permutation proportion p is chosen larger, the spatial dependence pattern in the displacement $x(s)$ observations tends to be more severely corrupted through permutation, resulting in more spurious evidences toward rejecting the null hypothesis. In practice, often a small p is chosen, say $p = 10\%$, to ease the effect of spatial dependence distortion, but to still have enough randomness to prevent generating a degenerate permutation distribution (Haining, 2001).

3 Compute $f_9^{(k)}(i,j)$ based on the kth permutation sample for cell-(i,j) of $\mathcal{G}_{I\times J}$. Then compute the associated sample variance statistic

$$X_k^2 = \frac{1}{IJ-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\left(f_9^{(k)}(i,j)-\bar{f}_9\right)^2$$

4 Compare X_0^2 with the empirical distribution (e.g., histogram) of $\left\{X_k^2, k = 1, \cdots K\right\}$ by computing the p-value as the proportion of permutations whose variance statistic are large or equal to X_0^2

If H_0 is true, the observed variance statistic X_0^2 value would be typical of the X_k^2 values, meaning the p-value would be large. In other words, if the p-value is small, we would have strong statistical evidence to reject H_0.

Data Availability Statement

The InSAR data are freely available from Copernicus Open Access Hub (https://scihub.copernicus.eu/dhus). Software used in the analysis of this study and to produce figures includes the following Python packages: scikit-learn (Pedregosa et al., 2011), SciPy (Virtanen et al., 2020), and matplotlib (Hunter, 2007).

Acknowledgements
A.T., S.Z. acknowledge support from the U.S. Army International Technology Center Pacific (ITC-PAC) and US DoD High Performance Computing Modernization Program (HPCMP) under Contract FA5209-18-C-0002. Open access publishing facilitated by Victoria University of Wellington, as part of the Wiley - Victoria University of Wellington agreement via the Council of Australian University Librarians.

References

Aggarwal, C. C., Han, J., Wang, J., & Yu, P. S. (2004). A framework for projected clustering of high dimensional data streams. In *Proceedings of the Thirtieth International Conference on Very Large Data Bases - Volume 30* (pp. 852–863). https://doi.org/10.1016/b978-012088469-8.50075-9

Allasia, P., Manconi, A., Giordan, D., Baldo, M., & Lollino, G. (2013). Advice: A new approach for near-real-time monitoring of surface displacements in landslide hazard scenarios. *Sensors*, *13*(7), 8285–8302. https://doi.org/10.3390/s130708285

Anantrasirichai, N., Biggs, J., Albino, F., Hill, P., & Bull, D. (2018). Application of machine learning to classification of volcanic deformation in routinely generated InSAR data. *Journal of Geophysical Research: Solid Earth*, *123*(8), 6592–6606. https://doi.org/10.1029/2018JB015911

Barla, M., & Antolini, F. (2016). An integrated methodology for landslides' early warning systems. *Landslides*, *13*(2), 215–228. https://doi.org/10.1007/s10346-015-0563-8

Calvari, S., Intrieri, E., Di Traglia, F., Bonaccorso, A., Casagli, N., & Cristaldi, A. (2016). Monitoring crater-wall collapse at active volcanoes: A study of the 12 January 2013 event at Stromboli. *Bulletin of Volcanology*, *78*(5), 39. https://doi.org/10.1007/s00445-016-1033-4

Carlà, T., Intrieri, E., Raspini, F., Bardi, F., Farina, P., Ferretti, A., & Casagli, N. (2019). Perspectives on the prediction of catastrophic slope failures from satellite InSAR. *Scientific Reports*, *9*(1), 14137. https://doi.org/10.1038/s41598-019-50792-y

Cigna, F., Bianchini, S., & Casagli, N. (2013). How to assess landslide activity and intensity with persistent scatterer interferometry (psi): The psi-based matrix approach. *Landslides*, *10*(3), 267–283. https://doi.org/10.1007/s10346-012-0335-7

Cigna, F., & Tapete, D. (2021). Sentinel-1 big data processing with P-SBAS InSAR in the geohazards exploitation platform: An experiment on coastal land subsidence and landslides in Italy. *Remote Sensing*, *13*(5), 885. https://doi.org/10.3390/rs13050885

Collins, B. D., & Stock, G. M. (2016). Rockfall triggering by cyclic thermal stressing of exfoliation fractures. *Nature Geoscience*, *9*(5), 395–400. https://doi.org/10.1038/ngeo2686

Crosta, G., & Agliardi, F. (2002). How to obtain alert velocity thresholds for large rockslides. *Physics and Chemistry of the Earth, Parts A/B/C*, *27*(36), 1557–1565. https://doi.org/10.1016/s1474-7065(02)00177-8

Das, S., & Tordesillas, A. (2019). Near real-time characterization of spatio-temporal precursory evolution of a rockslide from radar data: Integrating statistical and machine learning with dynamics of granular failure. *Remote Sensing*, *11*(23), 2777. https://doi.org/10.3390/rs11232777

Di Traglia, F., Bartolini, S., Artesi, E., Nolesini, T., Ciampalini, A., Lagomarsino, D., et al. (2018). Susceptibility of intrusion-related landslides at volcanic islands: The Stromboli case study. *Landslides*, *15*(1), 21–29. https://doi.org/10.1007/s10346-017-0866-z

Di Traglia, F., Fornaciai, A., Favalli, M., Nolesini, T., & Casagli, N. (2020). Catching geomorphological response to volcanic activity on steep slope volcanoes using multi-platform remote sensing. *Remote Sensing*, *12*(3), 438. https://doi.org/10.3390/rs12030438

Di Traglia, F., Nolesini, T., Ciampalini, A., Solari, L., Frodella, W., Bellotti, F., et al. (2018). Tracking morphological changes and slope instability using spaceborne and ground-based SAR data. *Geomorphology*, *300*, 95–112. https://doi.org/10.1016/j.geomorph.2017.10.023

Di Traglia, F., Nolesini, T., Solari, L., Ciampalini, A., Frodella, W., Steri, D., et al. (2018). Lava delta deformation as a proxy for submarine slope instability. *Earth and Planetary Science Letters*, *488*, 46–58. https://doi.org/10.1016/j.epsl.2018.01.038

Dick, G. J., Eberhardt, E., Cabrejo-Liévano, A. G., Stead, D., & Rose, N. D. (2015). Development of an early-warning time-of-failure analysis methodology for open-pit mine slopes utilizing ground-based slope stability radar monitoring data. *Canadian Geotechnical Journal*, *52*(4), 515–529. https://doi.org/10.1139/cgj-2014-0028

Duan, L., Tang, G., Pei, J., Bailey, J., Campbell, A., & Tang, C. (2015). Mining outlying aspects on numeric data. *Data Mining and Knowledge Discovery*, *29*(5), 1116–1151. https://doi.org/10.1007/s10618-014-0398-2

Fan, X., Xu, Q., Scaringi, G., Dai, L., Li, W., Dong, X., et al. (2017). Failure mechanism and kinematics of the deadly June 24th 2017 Xinmo landslide, Maoxian, Sichuan, China. *Landslides*, *14*(6), 2129–2146. https://doi.org/10.1007/s10346-017-0907-7

Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., & Rucci, A. (2011). A new algorithm for processing interferometric data-stacks: SqueeSAR. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(9), 3460–3470. https://doi.org/10.1109/tgrs.2011.2124465

Fukuzono, T. (1985). A new method for predicting the failure time of a slope. In *Proceedings of 4th International Conference and Field Workshop on Landslide* (pp. 145–150).

Gaida, M., Cambio, D., Robotham, M., & Pere, V. (2021). Development and application of a reliability-based approach to slope design acceptance criteria at Bingham Canyon mine. In *SSIM 2021: Second International Slope Stability in Mining* (pp. 83–94). Australian Centre for Geomechanics. Retrieved from https://papers.acg.uwa.edu.au/p/2135_02_Robotham, https://doi.org/10.36487/ACG_repo/2135_02

Gao, K., Guyer, R., Rougier, E., Ren, C. X., & Johnson, P. A. (2019). From stress chains to acoustic emission. *Physical Review Letters*, *123*(4), 048003. https://doi.org/10.1103/physrevlett.123.048003

Good, P. (2005). *Permutation, parametric and bootstrap tests of hypotheses* (3rd ed.). Springer.

Haining, R. (2001). Spatial sampling. In N. J. Smelser, & P. B. Baltes (Eds.), *International encyclopaedia of the social & behavioral sciences* (pp. 14822–14827). Oxford: Pergamon. https://doi.org/10.1016/b0-08-043076-7/02510-9

Hungr, O., Leroueil, S., & Picarelli, L. (2014). The Varnes classification of landslide types, an update. *Landslides*, *11*(2), 167–194. https://doi.org/10.1007/s10346-013-0436-y

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Intrieri, E., Bardi, F., Fanti, R., Gigli, G., Fidolini, F., Casagli, N., et al. (2017). Big data managing in a landslide early warning system: Experience from a ground-based interferometric radar application. *Natural Hazards and Earth System Sciences*, *17*(10), 1713–1723. https://doi.org/10.5194/nhess-17-1713-2017

Intrieri, E., Carlà, T., & Gigli, G. (2019). Forecasting the time of failure of landslides at slope-scale: A literature review. *Earth-Science Reviews*, *193*, 333–349. https://doi.org/10.1016/j.earscirev.2019.03.019

Intrieri, E., & Gigli, G. (2016). Landslide forecasting and factors influencing predictability. *Natural Hazards and Earth System Sciences*, *16*(12), 2501–2510. https://doi.org/10.5194/nhess-16-2501-2016

Intrieri, E., Gigli, G., Mugnai, F., Fanti, R., & Casagli, N. (2012). Design and implementation of a landslide early warning system. *Engineering Geology*, *147–148*, 124–136. https://doi.org/10.1016/j.enggeo.2012.07.017

Intrieri, E., Raspini, F., Fumagalli, A., Lu, P., Del Conte, S., Farina, P., et al. (2018). The Maoxian landslide as seen from space: Detecting precursors of failure with sentinel-1 data. *Landslides*, *15*(1), 123–133. https://doi.org/10.1007/s10346-017-0915-7

Johnston, E. C., Davenport, F. V., Wang, L., Caers, J. K., Muthukrishnan, S., Burke, M., & Diffenbaugh, N. S. (2021). Quantifying the effect of precipitation on landslide hazard in urbanized and non-urbanized areas. *Geophysical Research Letters*, *48*, e2021GL094038. https://doi.org/10.1029/2021GL094038

Justice, S. R. (2020). *A slippery slope: A hedonic property value study of landslide risk and economic costs in Watauga county*. (Unpublished doctoral dissertation). Appalachian State University.

Kothari, U. C., & Momayez, M. (2018). New approaches to monitoring, analyzing and predicting slope instabilities. *Journal of Geology and Mining Research*, *10*(1), 1–14. https://doi.org/10.5897/JGMR2017.0272

Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. *Paper presented at 2008 Eighth IEEE International Conference on Data Mining* (pp. 413–422). https://doi.org/10.1109/icdm.2008.17

Lloyd, S. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129–137. https://doi.org/10.1109/tit.1982.1056489

Macciotta, R., Hendry, M., & Martin, C. D. (2016). Developing an early warning system for a very slow landslide based on displacement monitoring. *Natural Hazards*, *81*(2), 887–907. https://doi.org/10.1007/s11069-015-2110-2

Masih, A. (2018). An enhanced seismic activity observed due to climate change: Preliminary results from Alaska. In *IOP Conference Series: Earth and Environmental Science* (Vol. 167, p. 012018). https://doi.org/10.1088/1755-1315/167/1/012018

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, *2*(1), 37–63.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sansosti, E., Berardino, P., Bonano, M., Calò, F., Castaldo, R., Casu, F., et al. (2014). How second generation SAR systems are impacting the analysis of ground deformation. *International Journal of Applied Earth Observation and Geoinformation*, *28*, 1–11. https://doi.org/10.1016/j.jag.2013.10.007

Schaefer, L. N., Di Traglia, F., Chaussard, E., Lu, Z., Nolesini, T., & Casagli, N. (2019). Monitoring volcano slope instability with synthetic aperture radar: A review and new data from Pacaya (Guatemala) and Stromboli (Italy) volcanoes. *Earth-Science Reviews*, *192*, 236–257. https://doi.org/10.1016/j.earscirev.2019.03.009

Singh, K., & Tordesillas, A. (2020). Spatiotemporal evolution of a landslide: A transition to explosive percolation. *Entropy*, *22*, 67. https://doi.org/10.3390/e22010067

Stähli, M., Sättele, M., Huggel, C., McArdell, B. W., Lehmann, P., Van Herwijnen, A., et al. (2015). Monitoring and prediction in early warning systems for rapid mass movements. *Natural Hazards and Earth System Sciences*, *15*(4), 905–917. https://doi.org/10.5194/nhess-15-905-2015

Tordesillas, A., Kahagalage, S., Campbell, L., Bellett, P., Intrieri, E., & Batterham, R. (2021). Spatiotemporal slope stability analytics for failure estimation (SSSAFE): Linking radar data to the fundamental dynamics of granular failure. *Scientific Reports*, *11*, 9729. https://doi.org/10.1038/s41598-021-88836-x

Tordesillas, A., Kahagalage, S., Ras, C., Nitka, M., & Tejchman, J. (2020). Early prediction of macrocrack location in concrete, rocks and other granular composite materials. *Scientific Reports*, *10*(1), 20268. https://doi.org/10.1038/s41598-020-76616-y

Tordesillas, A., Kahagalage, S., Ras, C., Nitka, M., & Tejchman, J. (2020). Coupled evolution of preferential paths for force and damage in the pre-failure regime in disordered and heterogeneous, quasi-brittle granular materials. *Frontiers in Materials*, *7*, 1–20. https://doi.org/10.3389/fmats.2020.00079

Tordesillas, A., Zhou, S., Di Traglia, F., & Intrieri, E. (2020). New insights into the spatiotemporal precursory failure dynamics of the 2017 Xinmo landslide and its surrounds. In *Workshop on world landslide forum* (pp. 331–338). https://doi.org/10.1007/978-3-030-60311-3_39

Tordesillas, A., Zhou, Z., & Batterham, R. (2018). A data-driven complex systems approach to early prediction of landslides. *Mechanics Research Communications*, *92*, 137–141. https://doi.org/10.1016/j.mechrescom.2018.08.008

Varnes, D. J., & Cruden, D. (1996). *Landslide types and processes. Landslides: Investigation and mitigation* (p. 247). Transportation Research Board Special Report

Vinh, N. X., Chan, J., Romano, S., Bailey, J., Leckie, C., Ramamohanarao, K., & Pei, J. (2016). Discovering outlying aspects in large datasets. *Data Mining and Knowledge Discovery*, *30*(6), 1520–1555. https://doi.org/10.1007/s10618-016-0453-2

Vinh, N. X., Epps, J., & Bailey, J. (2010). Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research*, *11*, 2837–2854. http://jmlr.org/papers/v11/vinh10a.html

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Wang, H., Qian, G., & Tordesillas, A. (2020). Modeling big spatio-temporal geo-hazards data for forecasting by error-correction cointegration and dimension-reduction. *Spatial Statistics*, *36*, 100432. https://doi.org/10.1016/j.spasta.2020.100432

Wang, X., Nie, G., & Wang, D. (2010). Relationships between ground motion parameters and landslides induced by Wenchuan earthquake. *Earthquake Science*, *23*(3), 233–242. https://doi.org/10.1007/s11589-010-0719-5

Yang, C., Yu, M., Li, Y., Hu, F., Jiang, Y., Liu, Q., et al. (2019). Big earth data analytics: A survey. *Big Earth Data*, *3*(2), 83–107. https://doi.org/10.1080/20964471.2019.1611175

Zhou, S., Bondell, H., Tordesillas, A., Rubinstein, B. I. P., & Bailey, J. (2020). Early identification of an impending rockslide location via a spatially-aided Gaussian mixture model. *Annals of Applied Statistics*, *14*(2), 977–992. https://doi.org/10.1214/20-aoas1326

Zinno, I., Mossucca, L., Elefante, S., De Luca, C., Casola, V., Terzo, O., et al. (2016). Cloud computing for earth surface deformation analysis via spaceborne radar imaging: A case study. *IEEE Transactions on Cloud Computing*, *4*(1), 104–118. https://doi.org/10.1109/TCC.2015.2440267