Abstract  Founded on understanding of a slope’s likely failure mechanism, an early warning system for instability should alert users of accelerating slope deformation behaviour to enable safety-critical decisions to be made. Acoustic emission (AE) monitoring of active waveguides (i.e. a steel tube with granular internal/external backfill installed through a slope) is becoming an accepted monitoring technology for soil slope stability applications; however, challenges still exist to develop widely applicable AE interpretation strategies. The objective of this study was to develop and demonstrate the use of machine learning (ML) approaches to automatically classify landslide kinematics using AE measurements, based on the standard landslide velocity scale. Datasets from large-scale slope failure simulation experiments were used to train and test the ML models. In addition, an example field application using data from a reactivated landslide at Hollin Hill, North Yorkshire, UK, is presented. The results show that ML can automatically classify landslide kinematics using AE measurements with the accuracy of more than 90%. The combination of two AE features, AE rate and AE rate gradient, enable both velocity and acceleration classifications. A conceptual framework is presented for how this automatic approach would be used for landslide early warning in the field, with considerations given to potentially limited site-specific training data.

Keywords  Slope stability · Landslide movement · Acoustic emission monitoring · Machine learning · Data driven

Introduction  Rainfall-induced landslides cause significant damage to infrastructure and kill thousands of people each year (Petley 2012). All landslide studies must be founded on an understanding of the critical failure mechanism, including type of failure (e.g. sliding, fall, flow) and extent (e.g. depth to shear surface). This paper focusses on sliding mechanisms in soils and the kinematics (i.e. motion) of the failure event. It is established practice to monitor slopes to alert users of accelerating slope deformation behaviour, enable evacuation of vulnerable people, and conduct timely repair and maintenance of critical infrastructure; these are termed early warning systems (EWS). EWS can be classified as alarm, warning, and forecasting systems (Stähli et al. 2015). Alarm systems provide a timely alert to people in the immediate vicinity of the landslide. Warning systems are preferred where progressive stages of failures can be identified, and an alert can be provided to experts who are responsible to analyse the situation and manage risk by implementing appropriate interventions. Forecasting systems commonly produce data that are interpreted by experts on a regular basis, often for a regional scale, with a typical output being danger levels that are communicated to the public with a bulletin.

Shear zones develop in strain-softening soils when the shear stress exceeds the peak shear strength locally within the slope, causing post-peak reductions in strength to occur. These shear zones propagate through the slope, ultimately developing a continuous shear surface and leading to first-time slope failure. The boundary stresses remain unchanged (i.e. self-weight), and hence the reduction in strength in first-time slides causes progressively accelerating movements that can reach high velocities and large displacements. In contrast, reactivated landslides have a shear surface already at, or near, residual shear strength, and hence no further strain-softening can occur. Reactivated landslides move with comparatively low velocities and over small displacements, and their behaviour is controlled by transient elevations and dissipations of pore-water pressures (Chandler 1984; Cooper et al. 1998; Leroueil 2001; Skempton 1964; Skempton and Petley 1967; Smith et al. 2017).

The standard landslide velocity scale (Cruden and Varnes 1996; Hungr et al. 2014) comprises a series of classifications that progress from ‘extremely slow’ (a few millimetres per year) to ‘extremely rapid’ (metres per second), and each velocity classification is separated by two orders of magnitude (Table 1). Acceleration quantifies how rapidly a slope progresses through velocity classifications, and whether the slope is slowing down (decelerating), and hence acceleration provides critical information for use in early warning and risk management.

Acoustic emission (AE) is becoming an accepted monitoring technology for geotechnical applications (Berg et al. 2018; Dixon et al. 2015; Dixon et al. 2015; Lin et al. 2020; Mao et al. 2020; Smith et al. 2014; Smith et al. 2019); however, challenges still exist to develop widely applicable AE interpretation strategies. AE is relatively high-frequency (i.e. typically non-audible) elastic stress waves that propagate through materials surrounding the generation source. AE is generated when a material undergoes irreversible changes in its internal structure—for example, crack formation or plastic deformation due to ageing, temperature gradients, or external mechanical forces. In soil, AE is generated by inter-particle friction and hence the detection of AE is an indication of deformation.

Figure 1 shows an active waveguide system that is used for soil slope stability monitoring. The active waveguide is installed inside a borehole or driven into the soil, intersecting existing or anticipated shear surfaces, and comprises a steel tube with internal/external backfill, which is typically coarse-grained, ‘noisy’ soil (i.e. when deformed). As the slope moves, it causes deformations in the waveguide and backfill, which generates AE that propagate as guided waves up the steel waveguide to the sensor at ground level. Extensive field trials and large-scale laboratory experiments of AE monitoring using active waveguides installed in slopes have produced a significant body of evidence showing that generated AE rates (i.e. the number of times in each period the collected signal exceeds a pre-defined voltage threshold) are proportional to the rate of slope movement (Berg et al. 2018; Dixon et al. 2018; Dixon et al. 2015; Dixon et al. 2015; Smith and Dixon 2015; Smith et al. 2017; Smith et al. 2014). Moreover, Smith et al. (2017) demonstrated that the AE approach can detect the development of new shear surfaces. A significant benefit of the AE monitoring approach is...
that it provides continuous operation, high sensitivity able to detect very slow displacement rates, remote, automated and robust operation, and production of (near) real-time warnings at lower costs than current subsurface in-place deformation instrumentation for continuous measurements (e.g. in-place inclinometers and ShapeAccelArrays) (Dixon et al. 2018).

The magnitude of measured AE rates in response to an applied slope displacement rate is influenced by a series of parameters, including sensor sensitivity and configuration, and the depth to the shear surface (i.e. the magnitude of attenuation AE experiences as it propagates along the waveguide to the sensor). Smith (2015) investigated and quantified the influence of these parameters and developed a framework to determine initial AE rate-displacement rate calibration relationships for any AE system installation. This calibration approach was shown to achieve accuracy to within an order of magnitude, that is, an order of magnitude better than the standard landslide velocity scale. However, site-specific data are required to achieve greater accuracy of slope displacement rates interpreted from AE measurements, and new techniques are required for automatic classification of landslide kinematics.

Machine learning (ML) can provide solutions for automatically accomplishing tasks such as regression, classification, and clustering (Alpaydin 2020; Pedregosa et al. 2011). Ensemble learning methods are a branch of ML that combines multiple data-driven learning algorithms into one predictive model to achieve superior performance (Polikar 2012). Ensemble learning comprises three methods: bagging (also known as bootstrap aggregation), boosting, and stacking. Bagging and boosting are two of the most commonly used techniques (Oza 2005), which can reduce variance and bias, respectively, and thus improve the predictive accuracy. Bagging methods aggregate the predictions of multiple models to produce a generalised result, and boosting methods combine multiple weak learners to form a single strong learner (Dietterich 2000). Bagging and boosting are both typically applied to tree models (i.e. ML model with tree structure). Random forest (RF) and XGBoost (eXtreme Gradient Boosting) are examples of widely used bagging and tree-boosting algorithms, respectively (Breiman 2001a; Chen and Guestrin 2016). Ensemble learning is efficient for automatic classification and has the potential to obtain classification results with higher accuracy than traditional ML approaches (Gomes et al. 2017).

The objective of this study was to develop and demonstrate the use of ML approaches to automatically classify landslide kinematics, based on the standard landslide velocity scale, using two AE features: AE rate and AE rate gradient. It is intended the approach complements and extends the tools currently available to landslide experts for classifying slope kinematics and delivering EWS. AE rate gradient is the time derivative of AE rate and acceleration is the time derivative of velocity. Therefore, there is potential to use AE rate gradient to interpret acceleration behaviour of a landslide (Deng et al. 2019). Support vector machine (SVM), RF, and XGBoost were selected as three representative algorithms for automatic classification. Datasets from large-scale slope failure simulation experiments performed by Smith et al. (2017) were used to train and test the ML models. In addition, an example field application using data from a reactivated landslide at Hollin Hill, North Yorkshire, UK, is presented. The paper introduces the ML approaches investigated, describes the AE datasets employed, and systematically assesses the performance of ensemble learning approaches to automatically perform classification of landslide kinematics. The use and benefits of the proposed approach for routine landslide monitoring and generation of early warning are considered.

Methodology

The data-driven approach adopted in this study for automatic classification comprises the components shown in Fig. 2. First, using AE measurements from a series of displacement controlled
large-scale shear tests (detailed in the ‘Experimental investigation’ section), data pre-processing was conducted, including smoothing and feature scaling. Second, category labels were generated based on velocity and acceleration magnitudes. Third, classification was performed using three different ML models, which were then compared. Two operations were used in the ML classification process. The first operation was classification for each dataset and called ‘typical operation’ (detailed in the ‘Model performance assessment approach’ section). The second operation was classification for independent datasets and called ‘cross-check operation’ (detailed in the ‘Model performance assessment approach’ section). Finally, model performance was assessed for each ML method using deformation measurements made concurrent with the AE, focused on classification accuracy and generalisation ability (i.e. adaptability of ML algorithm to successfully predict classification for a new dataset).

Data pre-processing and label generation
AE rate and velocity measurements with high sampling frequency (e.g. of the order of seconds) exhibit variability due to stick-slip behaviour of the active waveguide backfill and resolution limitations inherent in the deformation instrument operation. Smoothing is required to reduce this variability and ensure the measurements are compatible with measurements in the field, where significantly larger measurement intervals would be typical (e.g. 1-h or 1-day intervals). Smoothing was performed by applying moving averages to both AE rate and velocity measurements.

Two AE features, AE rate and AE rate gradient, were selected to investigate their potential when used with ML to automatically classify landslide kinematics, based on the standard landslide velocity scale (Table 1). Smith et al. (2017) plotted AE rate-velocity relationships on logarithmic scales because the landslide velocity descriptors are separated by two orders of magnitude. Moreover, a benefit of using logarithmic scales is that it emphasises data at low velocities, whereas the data at high velocities dominate on linear scales. A similar approach was taken in this study, whereby the logarithm (base 10) of each AE rate and AE rate gradient measurement was computed and used in subsequent stages. Figure 3 describes the process of base 10 logarithm scaling.

Measured AE in response to applied displacement rates can differ between active waveguide installations, as described in the ‘Introduction’ section. To overcome this in data pre-processing, scaling transformations were used to ensure data from different installations were comparable. Maximum absolute (max-abs) scaling was applied to normalise the AE measurements, whereby each series was scaled relative to its maximum absolute value (Eq. 1). Max-abs normalisation was beneficial for maintaining the sparsity (i.e. significant proportion of zero readings) of AE data. Each scaled value $x'$ was dimensionless and in a range between $-1$ and $1$, and all zeros in AE data were preserved.

$$x' = \frac{x}{\max(x)}$$

Classification labels are required as targets for supervised learning. Labels are usually generated from measurement data or manual operations. In this study, labels for the landslide kinematics were produced automatically based on deformation measurements (velocity and acceleration) and a set of criteria. Figure 4 shows the process of label generation. The initial step (label 0) establishes whether the slope velocity and acceleration are zero, and hence the slope is stable (not moving). Following this, a series of labels are generated based on the kinematics. In the standard landslide velocity scale, the recommended response changes (see Table 1) at the transitions from ‘extremely slow’ to ‘very slow’ (0.002 mm/h, maintenance) and from ‘slow’ to ‘moderate’ (20 mm/h, evacuation); hence, these two velocity criteria were used to generate labels in this study. It is notable that the same approach could be adopted for a range of applications, with additional and/or modified labels as required.

In addition to velocity classifications, information on whether a slope is accelerating or decelerating provides critical information to decision-makers, as this information can be used to interpret slope behaviour and the likelihood of incipient damaging failure. The acceleration of reactivated slides and pre-failure movements in first-time slides typically oscillate around 0 (Xu et al. 2011).
User-definable acceleration criteria, \( a_u \) (positive) and \( a_l \) (negative), are proposed, where values for acceleration thresholds are selected by a landslide expert considering the site-specific landslide mechanisms, deformation history, and likely failure kinematics (i.e. rate of movement of failed body). Early warning of incipient failure is an important task for landslide experts. Landslide acceleration is typically insignificant if the velocity remains extremely slow; thus, acceleration threshold \( a_u \) is not included for label 1.

### Machine learning models for classification

Ensemble learning reflects the state-of-the-art ML technology for classification tasks (Huang et al. 2020). RF and XGBoost are two of the most successful and widely used algorithms (Krkač et al. 2020; Krkač et al. 2017; Li et al. 2018), remote sensing (Gibson et al. 2020), statistics (Ishwaran and Lu 2019), and disease diagnosis (Lan et al. 2019). SVM is a traditional classification algorithm (Cortes and Vapnik 1995; Gola et al. 2019; Maldonado et al. 2020), which has been used for text classification (Tong and Koller 2002), image recognition (Raghout 2015), and fault diagnosis (Widodo and Yang 2007). These three data-driven models were investigated for use in landslide kinematic classification in this study: RF, XGBoost, and SVM. These algorithms were selected because the monitoring datasets were relatively small and only two features (AE rate and AE rate gradient) were used for prediction of multiple output classes. Extreme learning machine (ELM) and neural networks (NN) were also investigated initially, but they produced low accuracies and required higher run times, so they have not been considered further in this paper. SVM was selected as an established benchmark for comparison with the nascent techniques (RF and XGBoost). Model training was undertaken using the automatic tuning method of cross-validation (i.e. the dataset is split into multiple smaller subsets, and the model is trained by one subset, while the other subsets are used for subsequent verification) to determine the optimal parameters for each model (Ito and Nakano 2003; Probst et al. 2019).

RF is a bagging method for ensemble learning, which combines multiple independent decision trees in a random configuration (Breiman 2001a). RF first randomly selects features and samples to create subsets. Then, a series of decision trees are generated based on different subsets of input data. Finally, the individual predictions of these decision trees are aggregated to produce the ultimate prediction. Figure 5(a) illustrates how the RF model functions: majority voting selects the most common predicted output from all decision trees (e.g. class-A in Fig. 5(a)). The RF algorithm is robust for unbalanced datasets with missing data (Breiman 2001b) and can improve prediction accuracies without significantly increasing computational work.

Overfitting occurs when a model learns noise (e.g. corrections) in datasets to such an extent that it impairs its performance (Hawkins 2004). RF is not prone to overfitting during random operations of sampling and feature selection (Breiman 2001a). Thus, RF is expected to achieve low generalisation errors and high accuracy when making predictions using new data. Automatic tuning of 5-fold cross-validation was used to determine the optimal model parameters, such as the number of trees and the maximum depth (i.e. the longest path to traverse the tree from top to bottom).

XGBoost has received widespread attention in the literature because of its excellent performance in ML competitions (Chen and Guestrin 2016). The base learner of XGBoost is a decision tree, and multiple weak classifiers are combined to form a strong classifier through a significant number of iterations (Fig. 5(b)). Each variable is assigned a weight before progressing through the decision tree. The variables’ weights are subsequently updated before progressing through the next decision tree, and this sequence is continued to the n-th decision tree. These sequential classifiers are combined to form an additive model expressed in Eq. (2) and provide the final prediction. XGBoost has fast calculation rates and high accuracy and can be simplified to prevent overfitting (Chen and He 2020). The number of trees, maximum depth of each tree, and learning rate of XGBoost are automatically determined using 5-fold cross-validation method.

\[
\hat{y}_i^{(n)} = \frac{1}{n} \sum_{k=1}^{n} f_k(x_i) = \hat{y}_i^{(n-1)} + f_n(x_i)
\]

where \(\hat{y}_i^{(n)}\) denotes the prediction result of sample \(i\) after the \(n\)-th iteration, \(\hat{y}_i^{(n-1)}\) represents the prediction result of the previous \(n-1\) trees, and \(f_n(x_i)\) is the function of the \(n\)-th tree.

SVM is a linear classifier that constructs a set of hyperplanes (i.e. subplanes whose dimensions are one less than its host plane).

---

**Fig. 4 Flow diagram of label generation with smoothed velocity and acceleration**

![Flow diagram of label generation with smoothed velocity and acceleration](image-url)
To minimise generalisation errors, good separation should be achieved through large functional margins (i.e. large distances between hyperplanes and training data-points in the feature space) (Suykens and Vandewalle 1999). Figure 5(c) shows the architecture of SVM, in which the input layer contains the vector $x$ of dimension $n$, the hidden layer comprises product operations with the input vector $x$ for each of the $N$ support vectors using a kernel function, and the decision function $y$ is output as a combination of the $N$ kernel inner products (Ruiz-Gonzalez et al. 2014) (Eq. 3).

$$y = \sum_{i=1}^{N} a_i K(x, x_i) + b$$  \hspace{3cm} (5)

where $y$ is the decision function, $a_i$ is a coefficient, $K(x, x_i)$ represents the $i$-dimensional kernel inner product of the input vector $x$ with the support vector $x_i$, and $b$ denotes the bias.

Radial basis function (RBF) is the most commonly used kernel to modify SVM for nonlinear classification (Chang et al. 2010), which is expressed in Eq. (4).

$$K(x, x') = \exp\left(-\frac{||x-x'||^2}{2\sigma^2}\right)$$  \hspace{3cm} (4)

where $x$ and $x'$ are vectors of the two input samples and $\sigma$ represents standard deviation of the input data. Modified SVM can generalise for multi-class classification tasks (i.e. classifying instances into one of three or more classes) (Guo and Wang 2015; Mayoraz and Alpaydin 1999). In the training process of SVM classification, the optimal parameters are determined using 5-fold cross-validation method.

Model performance assessment approach

Accuracy is an indicator used to assess the performance of a classifier and refers to the proportion of instances correctly predicted by the classifier, which is expressed in Eq. (5).

$$A = \frac{N}{T} \times 100\%$$  \hspace{3cm} (5)

where $A$ denotes accuracy, $N$ is the number of correct predictions by the classifier, and $T$ is the total number of predictions.

A second important evaluation criterion when considering usefulness of an approach is the generalisation ability of the classification model. This can be assessed through calculating errors when applying the trained approach to new input datasets. Generalisation ability is demonstrated if the trained model avoids both underfitting and overfitting when applied to a range of diverse datasets. In order to assess the generalisation ability of the three ML models selected in this study to classify landslide kinematics, two testing operations were used to explore the classification accuracy of the trained models. The first operation called ‘typical operation’ (detailed in the ‘Classification result for datasets split from each test (typical operation)’ section) splits all collected observations from a dataset (described in the ‘Experimental investigation’ section) into a training set and a testing set. Each model is then trained with the training set based on each observation and associated label (Rippengill et al. 2003; Yella et al. 2007). The classification accuracy of the trained classifier is then calculated using testing data by comparing the consistency between actual labels and predicted labels. The second operation called ‘cross-check operation’ (detailed in the ‘Classification result using Test 3 as training set (cross-check operation)’ section) uses a dataset from one experiment/event to train the classifier, and then uses the trained classifier to perform classification for a fresh dataset from different experiments/events. Hence, the generalisation ability of the three ML models for classification is evaluated both using a subset of the same dataset and several independent datasets, the latter having ranges of different conditions.

Experimental investigation

Overview of large-scale physical test

Figure 6 shows the large-scale first-time slope failure simulation apparatus developed by Smith et al. (2017), which allowed active waveguides to be subjected to first-time slope failure dynamics (i.e. development of new shear surfaces and accelerating deformation behaviour). The apparatus was a bespoke large shear box, which comprised two concrete blocks, each with external dimensions $1.0 \times 0.7 \times 0.7$ m. The bottom box was fixed to a reinforced concrete floor to prevent movement. Each box had an open column ($0.3 \times 0.3$ m) forming a continuous opening though the two halves, which was filled with compacted clay to represent an element of the slope. A full-scale active waveguide and ShapeAccelArray (SAA, in-place inclinometer) was installed to the base of the clay column. A pulley system connected the top box to a hydraulically controlled loading ram, which was used to apply load and displacement to the shear box. The loading ram moved upwards, which pulled the wire rope around the shear block, and displaced the top box horizontally to induce shearing in the clay column. Figure 6(b–d) shows the active waveguide backfill materials employed: limestone gravel (LSG) in Tests 1 to 3, Leighton Buzzard sand (LBS) in Test 4, and granite gravel (GG) in Test 5.

Figure 7 shows the SAA-measured shear surface displacements Smith et al. (2017) reported for five tests. The duration of the first two loading stages was progressively increased from Test 1 to 3, and the displacement-time behaviour remained the same for Tests 3 to 5. Continuous AE and deformation measurements were collected. Each test had two AE system settings: one with a voltage threshold set to 0.25 V and the other at 0.1 V. The AE measurements were ring-down counts (RDC), which are the number of times the waveform crosses the pre-set voltage threshold (i.e. the threshold is set to exclude system electronic noise). For any given event, higher RDC would be recorded by the system with a lower voltage threshold. The measurements obtained at 0.25 V were primarily used in the research reported here. Comparisons between Tests 1, 2, and 3 focused on different movement time series, and comparisons between Tests 3, 4, and 5 focused on different backfill materials. The additional 0.1 V threshold dataset from Test 3 was also used (i.e. called Test 6 hereafter) to investigate the influence of AE system settings.

Data processing

SAA-measured velocity was smoothed using 2-min moving average (2-min MA) values, calculated over the 1 min preceding and the 1 min following each measurement. Derived acceleration data
exhibited more significant oscillations and hence was smoothed using 10-min moving average (10-min MA) values. It should be noted that this temporal resolution and smoothing is used in this study as an example; the same ML methodologies could be used for any temporal resolution and smoothing approaches relevant for different applications. Figure 8(a) shows the time series of smoothed velocity and acceleration and generated movement labels for Test 3. The majority of the smoothed velocity values are above 0.5 mm/h, with only few data in the range below 0.2 mm/h (Fig. 8(a)). The slowest velocity possible with the loading machine was 3.6 mm/h, and therefore, ‘stable’ and ‘extremely slow’ classifications are excluded in the subsequent analysis. Smoothed acceleration values oscillate between ± 3 mm/h² for most of the test duration and only exceed 3 mm/h² in the final stage. Negative values of acceleration are attributed to the servo-controlled loading process and transitions between velocity stages. Thus, deceleration is not considered in the subsequent analysis. In summary, the following labels and associated behaviours are considered in the analysis of landslide kinematics: very slow/slow moving (label 2), very slow/slow moving and accelerating (label 2A), moderate/rapid moving (label 3), and moderate/rapid moving and accelerating (label 3A).

Consistent with the smoothing processes used for velocity and acceleration, 2-min MA smoothing was applied to AE rate and 10-min MA smoothing to AE rate gradient. Figure 8(b) shows the time series of smoothed AE rate and AE rate gradient in Test 3. The labels generated by the velocity and acceleration data, as previously shown in Fig. 8(a), are also plotted. AE generation began during shear surface formation (i.e. transition from elastic to plastic load-displacement behaviour) at approximately \( t = 60 \) mins and a total displacement of 1.9 mm, and hence, AE measurements before this are 0, while the corresponding SAA-derived label is 2 (very slow/slow moving). In the final stage, AE rate rises sharply to its maximum and AE rate gradient fluctuates while most labels are 3A (moderate/rapid moving and accelerating). The evolution of AE rate and AE rate gradient is consistent with the kinematic
labels obtained using the SAA deformation measurements, indicating slope behaviour can be interpreted using AE parameters.

**Classification using machine learning**

Based on the labels (i.e. generated from the SAA measurements) and two AE features (i.e. AE rate and AE rate gradient), ML models were trained to derive predicted classifications for landslide kinematics with only AE measurements. The aim is to provide continuous quantitative information on slope behaviour when direct deformation measurements are not available. Two types of classification operations for landslide kinematics were performed in this study: typical operation and cross-check operation ('Model performance assessment approach' section). For typical operation, classification was conducted for a whole dataset of each Test. The original dataset was randomly split into a training set (70%) and a testing set (30%). A classifier for learning multiple classes was fitted using the training set, and subsequently the
performance of the trained classifier for landslide kinematics was assessed using the testing set. An additional cross-check operation used different datasets (i.e. from separate tests; see the 'Overview of large-scale physical test' section) to further explore the generalisation ability of the trained classifier for landslide kinematics. In cross-check operation, Test 3 was used as the training set and datasets from the other five tests were employed as testing sets. The purpose of the cross-check was to evaluate the generalisation ability of the ML model when encountering datasets from Tests with different conditions (e.g. movements time series, backfill materials and voltage threshold settings, as detailed in the 'Overview of large-scale physical test' section 3.1).

All three ML methods were applied using the two classification operations. Due to space limitations, the following analysis uses RF as an example to demonstrate the classification results for landslide kinematics. However, classification accuracy of all three ML methods is compared in Tables 2 and 3.

Classification result for datasets split from each test (typical operation) 

Figure 9 shows label information (obtained from SAA measurements) and AE data for training (red, 70%) and testing (green, 30%) data from Test 3. In training, both the two AE features and the SAA-derived labels were input into the model and used to obtain a trained classifier for landslide kinematics. In testing, only the two AE features were input into the trained classifier and the AE-derived labels were predicted, which were then compared with the SAA-derived target labels to assess the classification accuracy for landslide kinematics.

Figure 10 shows the classification results of all testing samples in each test derived from the prediction of RF model. The large majority of data points have been classified correctly as shown by the red crosses (AE-derived output labels) coinciding with the blue dots (SAA-derived measured labels). The classification accuracy for landslide kinematics is greater than 90% for all tests.

Consideration of feature (i.e. AE rate and AE rate gradient) importance, which is an output from the RF model, can be used to provide additional insight into the classification logic and thus can be helpful for guiding model improvements. The sum of importance of each AE feature used in the RF classifier is 1, and the larger the feature importance, the greater its contribution in determining the classification for landslide kinematics. Figure 11 shows the output feature importance for each test. AE rate was consistently more important than AE rate gradient in producing the classification for landslide kinematics. This difference is in part due to there being 3 labels (i.e. extremely slow, very slow/slow, or moderate/rapid) governed by AE rate and only 2 labels (i.e. accelerating or not) governed by AE rate gradient.
Table 2 Classification accuracy of testing set split from each test

| Method  | Test 1 | Test 2 | Test 3 | Test 4 | Test 5 | Test 6 |
|---------|--------|--------|--------|--------|--------|--------|
| SVM     | 90.6%  | 94.3%  | 93.9%  | 100.0% | 98.5%  | 98.5%  |
| XGBoost | 93.8%  | 90.6%  | 92.4%  | 95.5%  | 98.5%  | 96.9%  |
| RF      | 93.8%  | 92.5%  | 93.9%  | 95.5%  | 98.5%  | 98.5%  |

In addition to application of RF, this study also investigated the performance of SVM and XGBoost methods (‘Machine learning models for classification’ section) to automatically generate classification labels for landslide kinematics using AE measurements. Table 2 shows the classification result for each method. No one method consistently out-performed the others, but overall, the best results were obtained by RF. Of note is that SVM delivered 100% classification accuracy for Test 4.

Classification result using Test 3 as training set (cross-check operation)

Figure 12 shows scaled AE rate-AE rate gradient relationships for all six tests, which are colour- and symbol-coded based on their associated SAA-derived target labels. The labels progress through 2, 3, and 3A as the AE rate-AE rate gradient relationship increases (i.e., from the bottom left to the top right) in each plot. There is some overlap in data points (i.e., mixing between two labels) at the boundaries between the classifications; this highlights the importance of using order of magnitude differences to make the classifications for landslide kinematics. In terms of data points with the same SAA-derived labels (e.g., green dots), the distribution of Test 5 points shift slightly to the upper right in Fig. 12(e) when compared to Test 3 in Fig. 12(c). AE rates generated from the GG backfill in Test 5 were two orders of magnitude higher than comparable data in Test 3 when they were at the same velocity. Although base 10 logarithm and max-abs scaling were applied to reduce the large difference of AE rates between Test 3 and Test 5, AE-derived labels using Test 3 as training set are still not always consistent with the SAA-derived labels of Test 5.

The full dataset from Test 3 was used to train a classifier using SVM, XGBoost, and RF, respectively. The three trained classifiers were then tested using data from the other five tests. Table 3 shows the classification results for the five tests produced by each ML method. The classification accuracy of each independent test is about 90% except Test 5, which demonstrates these ML models have promising generalisation ability when applied to fresh datasets from experiments with different conditions.

Example field application using data from a reactivated landslide

To demonstrate use of the ML approach for field monitoring, an example is presented for a shallow (1.5-m-deep shear surface) reactivated translational landslide in weathered Whitby Mudstone Formation at Hollin Hill, North Yorkshire, UK [SE 68122 68852 (UK system); latitude, 54.11014; longitude, −0.95948786], which experiences periods of slow movement triggered by intensive periods of rainfall (Dixon et al. 2015). Continuous AE monitoring was conducted from 2010 to 2016, and for part of this time, a SAA was installed next to one of the AE waveguides in order to provide high temporal resolution subsurface deformation measurements (Smith et al. 2014). Landslide velocity has been derived from the measured displacement, with the large majority of velocity values being between 0.002 and 0.2 mm/h (i.e., measured over extended time periods). This indicates that the periodic slope movements at Hollin Hill are typically classified as very slow. These reactivated movements are characterised by accelerating movement from an initial stable condition followed by decelerating movement returning to stable. Figure 13 shows the 8-label classification framework designed for Hollin Hill to incorporate observed velocity and acceleration behaviour (i.e., very slow and slow velocities, both with acceleration and deceleration phases). Figure 14(a) shows the measured velocity, acceleration, and SAA-derived labels. Figure 14(b) shows the AE rate and AE rate gradient, with the SAA-derived labels from Fig. 14(a) included for comparison. All data is from the example period March 2016 to April 2016. The labels shown in Fig. 14 were obtained from the smoothed velocity (10-h MA) and acceleration (5-h average of velocity difference) SAA measurements by employing the flow diagram shown in Fig. 13.

The AE rate and AE rate gradient values presented in Fig. 14(b) are smoothed values obtained using the same method as for velocity and acceleration. After training the RF classifier with 70% data from Hollin Hill, Fig. 15 demonstrates that the classification accuracy using the testing set (30% data from the same period) is more than 90%. It is also shown that the model performs equally well across the range of classifications. This example validates ML as an approach to classify slope movement behaviour automatically based on AE data obtained through field monitoring.

Discussion

There are many studies on approaches and systems for landslide early warning using displacement measurements (Corsini and Mulas 2017; Intrieri et al. 2012; Pecoraro et al. 2019; Thiebes et al. 2014). All require knowledge of the site-specific conditions controlling mechanisms and rates of behaviour. Research has demonstrated that AE technology can complement and extend direct deformation measurement approaches and quantify the

Table 3 Classification accuracy for each test employing Test 3 as training set

| Method  | Test 1 | Test 2 | Test 4 | Test 5 | Test 6 |
|---------|--------|--------|--------|--------|--------|
| SVM     | 87.5%  | 87.9%  | 97.3%  | 71.1%  | 92.2%  |
| XGBoost | 81.7%  | 89.1%  | 96.3%  | 70.4%  | 90.4%  |
| RF      | 86.5%  | 86.8%  | 97.7%  | 74.8%  | 93.6%  |
The findings of this study clearly show that ML can automatically generate accurate classification of slope behaviour based on AE measurements. In addition, it has been shown that ML can be achieved using one dataset to train the model and then successfully applied to data obtained using other similar AE monitoring systems and slope failure conditions. This indicates that the method developed in this study has potential to be established as a general landslide early warning approach. These findings agree with the framework of AE early warning system proposed by Smith et al. (2017), but employ automatic classification of slope velocity by ML instead of a calibrated, empirical relationship converting from AE rate to velocity. The ML approach also advances the framework by incorporating quantified slope acceleration behaviour extracted from AE rate gradient.

For landslide early warning, accuracy of the predicted output class is the most important criteria: to minimise false alarms and ensure warnings are delivered when needed. Hence, accuracy was the focus for comparing the ML techniques in this study. Classification speed was also investigated by including a timer in the Python code; for example, for the Hollin Hill dataset (610 data points), the run times for RF, XGBoost, and SVM were 85 s, 57 s, and 13 s, respectively. The Hollin Hill field measurements included typical levels of extraneous noise, and so the accuracies obtained are indicative of those obtained in the field environment. Further work is required to evaluate the approaches for irrelevant/redundant attributes and missing data. RF achieved marginally greater accuracies in this study than XGBoost and SVM, which is likely to be because RF is robust for unbalanced datasets (Breiman 2001b). Moreover, it is expected that RF will perform better in cases with irrelevant/redundant attributes and missing data.

The next step is to apply the automated classification approach to a range of field conditions (i.e. landslide mechanisms and triggers) and to examine the challenges of implementing a site-based real-time early warning using AE monitoring data. Ideally, slope displacement-time information will be available for a slope during a period of AE monitoring. This can then be used to scale and calibrate the AE. A cost-effective approach using AE monitoring is to employ one inclinometer installation adjacent to an AE waveguide to provide this information, with multiple AE installations distributed across the slope to provide high-quality spatial and temporal information on movements (i.e. AE monitoring has been shown to be lower cost than use of in-place inclinometer systems (Dixon et al. 2018)).

An important step is to consider landslide types (e.g. soil properties, mechanisms) according to anticipated movement patterns and rates. For reactivated landslides with slow-moving episodes triggered by rainfall, ML training can be achieved using measured AE from periods of movement during the initial phase of monitoring. Subsequent movement rates can then be automatically classified in real-time using the trained model. However, for first-time slope movements, for ML to produce automatic classification, it must be trained using relevant laboratory tests and/or data from comparable field sites (i.e. employing similar AE waveguides and with analogous failure mechanism); this is because AE time series associated with the first-time failure will only be known after the event has occurred. For extremely slow/very slow landslides, such as exhibiting creep movements, it is a challenge using any currently available approach to differentiate movement rates unless the monitoring period is several days. In these types of applications, the classification system shown in Fig. 4 could be expanded to add additional classes, thus responding to site-specific conditions and monitoring requirements.

As noted above, when applying ML in field monitoring, difficulties may arise due to limitations of available direct displacement measurements and/or preliminary AE measurements (i.e.
Fig. 10 Label classification accuracy of testing samples from all the six tests using RF model. (a) Test 1, accuracy = 93.8%. (b) Test 2, accuracy = 92.5%. (c) Test 3, accuracy = 93.9%. (d) Test 4, accuracy = 95.5%. (e) Test 5, accuracy = 98.5%. (f) Test 6, accuracy = 98.5%
the AE range is unknown). To address these challenges, a conceptual framework is presented as follows to allow ML methods to be applied to AE data collected at a new field site.

In circumstances where direct deformation measurements are unavailable, it is recommended to use an existing approach based on laboratory tests or similar field sites to train the model and establish the calibrated AE-displacement relationship (Smith 2015). The site-specific measured AE data can then be used to produce classification labels describing the kinematics of the slope behaviour. These labels inform exceedance of predetermined thresholds defined by a landslide expert and hence generation of slope failure warnings. For sites where direct deformation measurements are available to train the ML model, the AE rate-AE rate gradient training data can be produced directly. If displacement monitoring continues contemporaneous with the AE measurements, then the ML model can also be updated during monitoring to improve the classification accuracy. Measurement of displacement at one location in the slope can be used to train multiple AE installations at the site.

The second challenge is when there is no history of AE monitoring at a site that can be used to define maximum expected AE...

![Fig. 11](Image) Feature importance produced by RF for each test

![Fig. 12](Image) AE rate-AE rate gradient relationships for Tests 1 to 6. SAA-derived labels 2, 3, and 3A are coloured as green circle, blue star, and red cross, respectively.
rate values (AE rate max), for use in scaling AE measurements (i.e. max-abs scaling). However, existing extensive data from laboratory experiments on common types of waveguides (i.e. backfill materials) subjected to ranges of deformation rates experienced by landslides, can be used to estimate AE rate max for anticipated slope behaviour. Dixon and Spriggs (2007) demonstrated that the Cruden and Varnes (1996) velocity descriptors are differentiated by two orders of magnitude changes in AE rate, and hence, errors in estimating AE rate max in order to scale AE parameters are unlikely to lead to significant error. The AE rate max used can be reviewed during monitoring, and updated if warranted, based on measured values.

The focus of ML applications in landslides research has to date been in susceptibility assessment (Chen et al. 2017; Dang et al. 2020; Dao et al. 2020; Moayedi et al. 2019; Nguyen et al. 2019; Park et al. 2019; Pourghasemi and Kerle 2016). However, recent studies
have demonstrated that ML approaches can be used to predict (i.e. derive) landslide displacement using time series measurements of groundwater levels, precipitation, and ground surface deformation (Krkač et al. 2020; Krkač et al. 2017; Li et al. 2018). The performance of these approaches has typically been assessed through quantification of prediction errors.

This study is the first evaluation of ML approaches for classifying landslide kinematics using AE measurements. Classification accuracy was the criterion used to assess model performance, and two types of operations were employed to test the generalisation ability of the trained models. The procedures used for training and testing were consistent with those used by other researchers, such as automatic tuning of model parameter via 5-fold cross-validation, model validation via comparison of predicted and measured data, and evaluation of feature importance (Krkač et al. 2020; Krkač et al. 2017; Li et al. 2018). The high prediction accuracy obtained by RF was also consistent with previous data-driven studies for both regression and classification tasks (Ge et al. 2021; Ishwaran and Lu 2019; Krkač et al. 2020; Li et al. 2018; Maxwell et al. 2020; Provost et al. 2017).

Ongoing research by the authors is investigating a generic regression method for landslide displacement prediction using AE, which uses existing laboratory and field datasets to pre-train the classification model. This would allow slope movement classification to be conducted using AE data obtained during initial monitoring in the field. The ML model can subsequently be updated with continuous real-time monitoring data to improve the prediction accuracy. This will provide an automatic and reliable basis for developing a landslide early warning system.

Conclusions
This study has developed and demonstrated the use of machine learning (ML) approaches to automatically classify landslide kinematics, based on the standard landslide velocity scale, using acoustic emission (AE) measurements. It is not a blackbox approach; in line with all comparable monitoring early warning systems, it is founded on understanding of the site-specific slope mechanism, soil conditions, and triggers. Three ML methods have been employed in this novel application to automate classification of slope behaviour: support vector machine (SVM), random forest (RF), and XGBoost (eXtreme Gradient Boosting). A programme of training and testing of the ML approaches using datasets from both large-scale slope failure simulation experiments and field measurements from a reactivated landslide led to the following principal findings.

a) The ML approaches can automatically classify landslide kinematics using AE measurements. All three data-driven models were demonstrated to be effective, with classification accuracies exceeding 90% when trained using a randomly selected subset of the data series. Accuracies greater than 90% were achieved using AE data from both large-scale slope failure simulation experiments and field measurements from a reactivated landslide.

b) Near 90% classification accuracy was achieved when one dataset was used to train the models, which were then applied to AE data measured in similar but independent tests with different conditions (e.g. deformation time series, backfill material, and sensor settings), demonstrating high generalisation ability, and hence, the ML techniques have the potential to be established as a general landslide early warning approach. It is notable that 90% accuracy (i.e. 1 in 10 errors) is acceptable because it is good practice for decision-makers to use trends in behaviour with time to trigger warnings; that is, they would not typically act on any individual measurement made over a short time interval to generate an early warning of slope failure.

c) The combination of two AE features, AE rate and AE rate gradient, enabled both velocity and acceleration kinematic classifications. Acceleration quantifies how rapidly a slope progresses through velocity classifications, and whether the slope is slowing down (decelerating), and hence, acceleration provides critical information for use in early warning and risk management.

A conceptual framework has been developed for how this approach would be used for landslide early warning in the field, with considerations given to potentially limited site-specific training data (deformation and/or AE measurements).
Acknowledgements
Alister Smith gratefully acknowledges the support of an EPSRC Fellowship (Listening to Infrastructure, EP/P012493/1). This paper is the outcome of the visit of Lizheng Deng to Loughborough University sponsored by Tsinghua Scholarship for Overseas Graduate Studies (Contract Number: 2019273).

Funding
This research was supported by the National Key Research and Development Program of China (No. 2018YFC0806900, No. 2018YFC0807000, and No. 2018YFC0810205) and the Key Research and Development Program of Anhui Province (Grant No. S20104B1020044).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References
Alpaydin E (2020) Introduction to machine learning. MIT press
Barghout L (2015) Spatial-taxon information granules as used in iterative fuzzy-decision-making for image segmentation. In: Pedrycz W, Chen S-M (eds) Granular computing and decision-making: interactive and iterative approaches. Springer International Publishing, Cham, pp 285–318
Barghout L (2015) Spatial-taxon information granules as used in iterative fuzzy-decision-making for image segmentation. In: Pedrycz W, Chen S-M (eds) Granular computing and decision-making: interactive and iterative approaches. Springer International Publishing, Cham, pp 285–318
Berg N, Smith A, Russell S, Dixon N, Proudfoot D, Take WA (2018) Correlation of acoustic emissions with patterns of movement in an extremely slow-moving landslide at Peace River, Alberta, Canada. Can Geotech J 55:1475–1488. https://doi.org/10.1139/cgj-2016-0668
Breiman L (2001a) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/a:1007450703504
Breiman L (2001b) Statistical modelling: the two cultures. Stat Sci 16:199–215. https://doi.org/10.1214/ss/1009212372
Chandler RJ (1984) Recent European experience of landslides in over-consolidated clays and soft rocks. Proc 4th International Symposium on Slopesides, Toronto, 61-81.
Chang Y-W, Hsieh C-J, Chang K-W, Ringgaard M, Lin C-J (2010) Training and testing low-degree polynomial data mappings via linear SVM. J Mach Learn Res 11:1471–1490
Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. Proceedings of the 32nd International Conference on Machine Learning, 274–282.
Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. Proceedings of the 32nd International Conference on Machine Learning, 274–282.
Cheng H, Lu M (2019) Standard errors and confidence intervals for variable importance in random forest. Statistical Analysis and Data Mining 12(4):503–516.
Cheng H, Lu M (2019) Standard errors and confidence intervals for variable importance in random forest. Statistical Analysis and Data Mining 12(4):503–516.
Cheng H, Lu M (2019) Standard errors and confidence intervals for variable importance in random forest. Statistical Analysis and Data Mining 12(4):503–516.
Cheng H, Lu M (2019) Standard errors and confidence intervals for variable importance in random forest. Statistical Analysis and Data Mining 12(4):503–516.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Cheng H, Lu M, Zhao J, Li J (2018) Variable selection procedures by using conditional variable important measures. Computational Statistics and Data Analysis 122:137–157.
Leroueil S (1982) Natural slopes and cuts: movement and failure mechanisms. Geotechnique 51:197–243. https://doi.org/10.1680/geot.1982.51.3.197
Li H, Xu Q, He Y, Deng J (2018) Prediction of landslide displacement with an ensemble-based extreme learning machine and copula models. Landslides 1-13
Lin W, Mao W, Liu A, Koseki J (2020) Application of an acoustic emission source-tracing method to visualise shear banding in granular materials. Geotechnique 0:1–12. https://doi.org/10.1680/jgeot.19.P.260
Maldonado S, López J, Jiménez-Molina A, Lira H (2020) Simultaneous feature selection and heterogeneity control for SVM classification: an application to mental workload assessment. Expert Syst Appl 143:112988. https://doi.org/10.1016/j.eswa.2019.112988
Mao W, Aoyama S, Towhata I (2020) A study on particle breakage behavior during pile penetration process using acoustic emission source location. Geosci Front 11:413–427. https://doi.org/10.1016/j.geofr.2019.04.006
Maxwell AE, Sharma M, Kite JS, Donaldson KA, Thompson JA, Bell ML, Maynard SM (2020) Slope failure prediction using random forest machine learning and LiDAR in an eroded folded mountain belt. Remote Sens 12. https://doi.org/10.3390/rs12030486
Mayoraz E, Alpaydin E (1999) Support vector machines for multi-class classification.
Nguyen VV, Pham BT, Vu BT, Park S, Hamm SY, Kim J (2019) Performance evaluation of the GIS-based data-mining techniques decision tree, random forest, and rotation forest for landslide susceptibility modeling. Sustainability 11. https://doi.org/10.3390/su11205659
Pecoraro G, Calvello M, Piculllo L (2019) Monitoring strategies for local landslide early warning systems. Landslides 16:213–231. https://doi.org/10.1007/s10346-018-01066-z
Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in python. J Mach Learn Res 12:2825–2830
Petley D (2012) Global patterns of loss of life from landslides. Geology 40: 927–930. https://doi.org/10.1130/g31232.1
Polikar R (2012) Ensemble Learning. In: Zhang C, Ma Y (eds) Ensemble Machine Learning: Methods and Applications. Springer US, Boston, pp 1–34
Pourghasemi HR, Kerle N (2016) Random forests and evidential belief function-based landslide susceptibility assessment in Western Mazandaran Province, Iran. Environ Earth Sci 75:185. https://doi.org/10.1007/s12665-015-4950-1
Probst P, Wright MN, Boulesteix A-L (2019) Hyperparameters and tuning strategies for random forest. WIREs Data Min Knowl Discov 9:e1301. https://doi.org/10.1002/widm.1301
Provost F, Hibert C, Malet J-P (2017) Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier. Geophys Res Lett 44:113–120. https://doi.org/10.1002/2016GL070709
Rippengill S, Worden K, Holford KM, Pullin R (2003) Automatic classification of acoustic emission patterns. Strain 39:39–41. https://doi.org/10.1016/j.strain.2003.05.001
Ruiz-Gonzalez R, Gomez-Gil J, Gomez-Gil FJ, Martinez-Martinez V (2014) An SVM-based classifier for estimating the state of various rotating components in agro-industrial machinery with a vibration signal acquired from a single point on the machine chassis. Sensors (Basel, Switzerland) 14:20713–20735. https://doi.org/10.3390/s141120713
Skempton AW (1964) Long-Term Stability of Clay Slopes. Geotechnique 14:77–102. https://doi.org/10.1680/geot.1964.14.2.77

Technical Notes

L. Deng · H. Yuan
Institute of Public Safety Research, Department of Engineering Physics, Tsinghua University, Beijing, 100084, China
L. Deng · H. Yuan
Beijing Key Laboratory of City Integrated Emergency Response Science, Tsinghua University, Beijing, 100084, China
L. Deng · H. Yuan
Anhui Province Key Laboratory of Human Safety , Tsinghua University , Hefei Anhui, 230601, China
A. Smith · N. Dixon
School of Architecture, Building and Civil Engineering, Loughborough University, Leicestershire, UK
Email: A.Smith10@lboro.ac.uk

2974 | Landslides 18 • (2021)