Understanding an irregular pattern of fracture slip from laboratory earthquake using machine learning

Cheng Mei, Wenzhao Meng, Wei Wu
School of Civil and Environmental Engineering, Nanyang Technological University, Singapore

Corresponding author’s e-mail address: wu.wei@ntu.edu.sg.

Abstract. A series of direct-shear experiments on a sawcut fracture under a multi-stage reduction in normal stress shows slip transition behaviors from stick-slip to stable sliding. The irregular slip pattern near the stability transition is traditionally considered unexplainable and unpredictable. This study utilizes a support vector machine (SVM) to classify the fracture slip behaviors and to predict new, unseen slip events. The results exhibit that the SVM trained by the acoustic emission amplitude well predicts the variation of shear stress near the stability transition, and the prediction accuracy slightly decreases with more irregular slip events occurred in the transition stage. The prediction for a fast rupture is more accurate than that for a slow rupture as relatively weak signals from a slow rupture are more sensitive at the same attenuation level. Understanding the irregular slip pattern allows us to better explain the occurrence of Parkfield tremors in California and provides more insights into the mechanisms of repeating earthquakes.

1. Introduction
Laboratory earthquake is considered as an analogue of natural earthquake, in terms of energy release and recurrence time, and as a robust approach to reveal the physics of natural earthquakes [1-4]. Although the laboratory simulation cannot capture all the physical processes on rock fractures and faults, this approach is believed to uncover the physics of friction and to implicate unexplained phenomena from the nature. In recent years, the combination of laboratory earthquake and machine learning has become a promising approach for earthquake prediction [5]. The laboratory earthquake can isolate key factors and use the corresponding results to improve the prediction of machine learning on specific characteristics of natural earthquakes. Particularly, in natural settings, an insufficient contrast between the seismic background noise and low rupture signals is commonly overlooked. However, reliable identification of low rupture signals based on machine learning may provide new clues for earthquake prediction.

Machine learning models have been widely applied to learn complex seismic signals and maximize useful features in the signals. The seismic signals from natural earthquakes have demonstrated successfully to detect small earthquakes [6] and to estimate earthquake magnitudes [7] using neural network models. The acoustic emission (AE) signals from laboratory earthquakes have been learnt to estimate fracture friction using a random forest model [8] and to classify tectonic earthquakes using a support vector machine (SVM) [9]. Other machine learning models, such as the gradient boosted
model [10] and binary classification model [11], are also used to study the physics of earthquakes. Among these models, the neural network cannot address how the independent variables influence the dependent variables [12], and the random forest is sensitive to small changes in the training data [13]. The SVM model can overcome these shortcomings, transform low-dimensional data into a high-dimensional feature space, and solve nonlinear problems efficiently [14]. Hence, the SVM is proposed to investigate complex slip behaviors in this study.

Our study aimed to understand an irregular pattern of fracture slip behaviors occurred during a multi-stage reduction in normal stress. These irregular patterns were poorly known, but the SVM was expected to be trained using previous slip events and to predict subsequent new, unseen events. Fast and slow ruptures inferred from shear stress drops in these irregular patterns were also predicted by the SVM. Finally, the irregular patterns of fracture slip behaviors were discussed and compared with unusual fault slip behaviors in nature.

2. Methods

2.1. Experimental study

We carried out a series of direct-shear experiments on a sawcut fracture in Makrolon polycarbonate. The bulk density, shear modulus, and Poisson’s ratio of the polycarbonate are 1190 kg/m³, 0.91 GPa, and 0.38, respectively. Figure 1 shows the direct-shear setup with fixed and driving plates, and the fracture is simulated by the interface between the two plates. The normal and shear loads were applied on the two plates via horizontal and vertical servo-controlled hydraulic rams, respectively, and measured using two load cells positioned on the ram noses. The load point displacement was measured by a linear variable differential transducer installed on the top of the vertical ram. The loads and displacement were recorded by a National Instrumental data acquisition system at a sampling rate of 10 kHz. A SAMOS AE monitoring system was used to record strain energy released from the fracture via four piezoelectric sensors attached on the driving plate at a sampling rate of 3 MHz.

![Figure 1. Photo of direct-shear setup. The lengths of the left and right polycarbonate plates are 180 and 150 mm, respectively. The width and thickness of two plates are 100 and 9.3 mm, respectively.](image)

The direct-shear experiments were conducted at a load point velocity of 50 μm/s under a normal stress decreasing from 2.0 to 0.8 MPa in seven steps, each with a normal stress reduction of 0.2 MPa and a load point displacement of 3 mm. We expected the occurrence of a full spectrum of fracture slip behaviors, including stick-slip, transition, and stable sliding, under various normal stresses, and used shear stress and AE amplitude obtained from the stick-slip and transition behaviors in the machine learning and predictive analysis.
2.2 Machine learning model

A support vector machine (SVM) is proposed for classification and regression problems. The SVM transforms the input data \( X = \{x_1, x_2, \ldots, x_i, \ldots, x_n\} \) (i.e., AE amplitude) into a high-dimensional feature space via a non-linear function \( \phi \) [14]:

\[
 f(X) = \omega^T \phi(X) + b, \quad \omega \in R^n, X \in R^n, b \in R
\]

(1)

where \( f(X) \) is the output data (i.e., shear stress), \( \omega \) is the adjustable weight vector, \( b \) is a bias constant, and \( R \) and \( R^n \) are the one-dimensional and \( n \)-dimensional vector spaces, respectively. The SVM requires a small \( \omega \) to minimize the error between the predicted and measured values using the \( \varepsilon \)-loss function:

\[
\text{Minimize } \frac{1}{2} \omega^T \omega
\]

(2)

Subject to

\[
\begin{align*}
 y_i - (\omega^T \phi(X) + b) & \leq \varepsilon \\
(\omega^T \phi(X) + b) - y_i & \leq \varepsilon
\end{align*}
\]

(3)

where \( y_i \) is the shear stress for the \( i \)-th AE amplitude, and \( \varepsilon \) is the the margin of tolerance. A smaller \( \varepsilon \) means a lower tolerance for error, and more input data fall within the \( \varepsilon \)-intensive band.

To increase the robustness of this model, as shown in Figure 2, we used the relaxation variables (\( \xi_i \) and \( \xi_i^* \)) to penalize misclassified data points:

\[
\text{Minimize } \frac{1}{2} \omega^T \omega + C \sum_{i=1}^{n} (\xi_i + \xi_i^*)
\]

(4)

Subject to

\[
\begin{align*}
 y_i - (\omega^T \phi(X) + b) & \leq \varepsilon + \xi_i \\
(\omega^T \phi(X) + b) - y_i & \leq \varepsilon + \xi_i^*
\end{align*}
\]

(5)

where \( C \) is the regularization parameter.

By introducing the Lagrange multipliers (\( \alpha \) and \( \alpha^* \)), the dual formulation can be expressed as:

\[
\text{Minimize } \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(X_i, X_j) - \sum_{i=1}^{n} y_i (\alpha_i - \alpha_i^*) + \varepsilon \sum_{i=1}^{n} (\alpha_i + \alpha_i^*)
\]

(6)

Subject to \( \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0 \), \( 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, \ldots, n \).

(7)

where \( K(X_i, X_j) \) is a Kernel function that can transfer the high-dimensional feature space into the low-dimensional input space [15]. Therefore, Equation (1) can be written as:

\[
 f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*)K(X_i, X_j) + b
\]

(8)

Here we adopted the radial basis function due to its efficiency in addressing nonlinear problems.
The performance of the SVM is evaluated through the mean absolute percentage error ($M$):

$$M = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - f_i}{y_i} \right| \cdot 100\%$$  \hspace{1cm} (9)

where $f_i$ is the predicted value.

The establishment of the SVM consisted of three main steps: (1) the training dataset included 1360 slip events with AE amplitudes obtained from the stick-slip and transition behaviors; (2) the parameters $\varepsilon$, $C$, and $\xi$ were optimized using the 10-fold cross-validation until the cross-validation mean square error reached $1 \times 10^{-6}$; and (3) the testing dataset involving 45 slip events was used to evaluate the performance of the SVM.

3. Results and discussion

A full spectrum of fracture slip behaviors was observed under the seven unloading stages, including the stick-slip behaviors under 2.0 and 1.8 MPa normal stresses, transition behaviors under 1.6, 1.4, 1.2, and 1.0 MPa normal stresses, and stable sliding under 0.8 MPa normal stress (Figure 3). We selected three typical steps with repeatable slip behaviors, one with the stick-slip behaviors under 2.0 MPa normal stress (stage 1) and two with the transition behaviors under 1.6 and 1.4 normal stresses (stages 3 and 4), in the machine learning and predictive analysis. The three stages contained 1360 slip events, which were used as the training data to build the SVM. The testing data containing 45 slip events were produced at the same load point velocity under the three normal stresses.

The results show that the stick-slip behaviors are regular and characterized by similar shear stress drops. The transition behaviors are irregular, but a sequence of repeating ruptures can still be observed, such as a fast rupture followed by one or more slow ruptures. Here fast and slow ruptures are inferred from large and small shear stress drops, respectively. The transition behaviors become more irregular under a lower normal stress. The AE signals released from the stick-slip behaviors are stronger than those from the transition behaviors, and the AE signals from a fast rupture are greater than those from

Figure 2. A nonlinear supportive vector regression with the $\varepsilon$-intensive band.
a slow rupture. The AE signals almost disappear during the stable sliding and thus not considered in the machine learning and predictive analysis.

The AE signals contain the information of failure characteristics of the fracture and are thus used to train the SVM. Figure 4 presents the predicted shear stress during periodic fast ruptures in the stick-slip stage and aperiodic slow and fast ruptures in the transition stages. The predicted shear stress fits well with the measured shear stress. The $M$ value for the stick-slip behaviors is as low as 7.61%. Moreover, the SVM satisfactorily predicts irregular shear stresses in the transition stages with the $M$ value in a range of 8.15% - 10.55%.

The measured and predicted shear stress drops are derived from the shear stresses in Figure 4 and used to demonstrate the predictions of fast and slow ruptures. As shown in Figure 5, the measured shear stress drop in the stick-slip behaviors under 2.0 MPa normal stress (stage 1) keeps roughly a constant value of 0.4 MPa. In the transition behaviors, the shear stress drop oscillates in a range of 0.15 - 0.35 MPa under 1.6 MPa normal stress (stage 3) and 0.15 - 0.32 MPa under 1.4 MPa normal stress (stage 4). The $M$ value for periodic fast ruptures is 5.03%, which is lower than those in the transition behaviors, 9.38% and 10.14% for aperiodic fast ruptures under lower normal stresses, indicating that the SVM has a better performance on the prediction of shear stress drop for regular stick-slip behaviors. The $M$ value increases in the prediction of shear stress drop for aperiodic fast ruptures can be attributed to the irregular changes in input data. The AE signals with different amplitudes travel through the same distance from the failure point to the measurement point. The attenuations of these signals generated from fast and slow ruptures can be similar as both the ruptures share the common mechanism [16]. Hence, for the same attenuation level, the mean absolute percentage error for a slow rupture with a low AE amplitude is larger than that for a fast rupture with a high AE amplitude.
The irregular patterns of fracture slip have been frequently observed in nature but not well explained in the literature. Fault instabilities occur in different modes, such as stick-slip and stable sliding. Depending on different stress variations, the stabilities can be modified between stick-slip and stable sliding, and the stability transition may occur before the fault stabilizes in a new mode. A striking example of the transition behaviors is a sequence of seismic tremors at the Parkfield section of the San Andreas Fault [17-19]. The recurrence interval of these tremors varies between three and six days (i.e., period-two cycles), and the recurrence pattern is interrupted by the 2004 $M_w$ 6.0 Parkfield earthquake, after which the recurrence interval becomes three days. The recurrence interval finally returns to the period-two cycles over the next two years. Our laboratory earthquake reproduces the complex slip behaviors with period-multiplying cycles and indicates that the irregular patterns are understandable and predictable. A further investigation on the irregular fracture slip behaviors in natural settings could lead us to predict the timing and size of repeating earthquakes.
4. Conclusions
Our results showed that an irregular pattern of fracture slip behaviors between the stick-slip and stable sliding. The SVM trained by the AE amplitude well performed in predicting the irregular pattern, especially fast and slow ruptures. The results revealed that the complex slip behaviors are understandable and predictable. The irregular pattern consists of a fast rupture followed by one or more slow ruptures. The predication accuracy slightly reduces with more slow ruptures occurred in the transition stage, and the predication accuracy for a fast rupture is higher than that for a slow ruptures. The results further explained the occurrence of Parkfield tremors in California. Laboratory earthquake certainly cannot reproduce all physical processes on rock fractures and faults. However, the combination of laboratory earthquake and machine learning can look into some key factors and to implicate unexplained faulting phenomena in nature.
References

[1] Leeman J R, Saffer D M, Scuderi M M and Marone C 2016 Laboratory observations of slow earthquakes and the spectrum of tectonic fault slip modes. *Nat. Commun.* **7** 11104

[2] Xia K, Rosakis A J and Kanamori H 2004 Laboratory earthquakes: the sub-Rayleigh-to-supershear rupture transition. *Science* **303** 1859-1861

[3] Wu W 2015 Slip initiation of granular gouge friction in a rock discontinuity induced static and dynamic loads. *Int. J. Rock Mech. Min. Sci.* **80** 196-201

[4] Wu W, Zhao J 2013 A dynamic-induced direct-shear model for dynamic triggering of frictional slip on simulated granular gouges. *Exp. Mech.* **54** 605-613

[5] Johnson P A, Rouet-Leduc B, Pyrak-Nolte L J, Beroza G C, Marone C J, Hulbert C, Howard A, Singer P, Gordeev D, Karaflos D, Levinso C J, Pfeiffer P, Puk K M and Reade W 2021 Laboratory earthquake forecasting: A machine learning competition. *Proc. Natl. Acad. Sci. U.S.A.* **118** e2011362118

[6] Mousavi S M, Ellsworth W L, Zhu W, Chuang L Y and Beroza G C 2020 Earthquake transformer – an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* **11** 3952

[7] Mousavi S M and Beroza G C 2020 A Machine-learning approach for earthquake magnitude estimation. *Geophys. Res. Lett.* **47** e2019GL085976

[8] Rouet-Leduc B, Hulbert C, Lubbers N, Barros K, Humphreys C J and Johnson P A 2017 Machine learning predicts laboratory earthquakes. *Geophys. Res. Lett.* **44** 9276-9282

[9] Tang L, Zhang M and Wen L 2020 Support vector machine classification of seismic events in the Tianshan orogenic belt. *J. Geophys. Res. Solid Earth* **125** e2019JB018132

[10] Rouet-Leduc B, Hulbert C, Bolton D C, Ren C X, Riviere J, Marone C, Guyer, R A and Johnson P A 2018 Estimating fault friction from seismic signals in the laboratory. *Geophys. Res. Lett.* **45** 1321-1329

[11] Corbi F, Bedford J, Sandri L, Funiciello F, Gualandi A and Rosenau M 2020 Predicting imminence of analog megathrust earthquakes with machine learning: Implications for monitoring subduction zones. *Geophys. Res. Lett.* **47** e2019GL086615

[12] Panakkat A and Adeli H 2007 Neural network models for earthquake magnitude prediction using multiple seismicity indicators. *Int. J. Neural Syst.* **17** 13-33

[13] Provost F, Hibert C and Malet J P 2017 Automatic classification of endogenous landslide seismicity using the Random Forest supervised classifier. *Geophys. Res. Lett.* **44** 113-120

[14] Smola A J and Schölkopf B 2004 A tutorial on support vector regression. *Stat. Comput.* **14** 199-222

[15] Samui P 2008 Support vector machine applied to settlement of shallow foundations on cohesionless soils. *Comput. Geotech.* **35** 419–427

[16] Hulbert C, Rouet-Leduc B, Johnson P A, Ren C X, Rivière J, Bolton D C and Marone C 2019 Similarity of fast and slow earthquakes illuminated by machine learning. *Nat. Geosci.* **12** 69–74

[17] Shelly D R 2010 Migrating tremors illuminate complex deformation beneath the seismogenic San Andreas fault. *Nature* **463** 648–652

[18] Mei C, Barbot S and Wu W 2021 Period-multiplying cycles at the transition between stick-slip and stable sliding and implications for the Parkfield period-doubling tremors. *Geophys. Res. Lett.* **48** e2020GL091807

[19] Mei C and Wu W 2021 Fracture asperity evolution during the transition from stick slip to stable sliding. *Philos. Trans. R. Soc. A* **379** 20200133