Supplementary Material for

Predicting stick-slips in sheared granular fault using machine learning optimized dense fault dynamics data

Weihan Huang[1], Ke Gao[1,2*], Yu Feng[3]

[1] Department of Earth and Space Sciences, Southern University of Science and Technology, Shenzhen, Guangdong, China

[2] Guangdong Provincial Key Laboratory of Geophysical High-resolution Imaging Technology, Southern University of Science and Technology, Shenzhen, Guangdong, China

[3] School of Civil Engineering, Sun Yat-sen University

Contents of this file

Texts

Figures

Figure S1. Numerical representation of the granular fault system using DEM and FDEM [13]. (**a**) In DEM, the plates are simplified by a set of bonded particles, and the gouge layer is composed of a series of rigid particles. Therefore, both the gouge particles and the shearing plates cannot deform. (**b**) In FDEM, the plates are explicitly represented, and both plates and particles are further discretized into finite elements to capture their detailed deformation and movement.

Figure S2. Probability density distribution of the seismic moment of all slip events, where the detailed calculation of the moment is explained by Gao, et al. [13]. The results agree with the physical

* Corresponding to: K. Gao (gaok@sustech.edu.cn)

experiment data collected in Geller, et al. [14]. The probability density distribution is consistent with the Gutenberg-Richter distribution [19] and is predicted to scale as $M^{-3/2}$ (the power -3/2 is within the range of -1.4 to -1.8 observed in natural earthquakes [14].

Figure S3. Performance of the two LightGBM models trained with different datasets. (**a**) Training with optimized sensor data. (b) Training with optimized sensor data and their statistics.

Figure S4. The SHAP value versus the feature value at each instant of time for the top 200 input features with relatively larger cumulative magnitudes of the yielded SHAP values. The results are based on the last LightGBM model trained using the optimized dataset and their statistics (*see the separate pdf file at the end*).

Figure S5. Evolutions of D_x for several sensors and their comparison with the normalized shear stress. (**a**) Normalized shear stress. (**b**) D_x of sensor No. 95. (c) D_x of sensor No. 295. (d) D_x of sensor No. 1832.

Tables

Table S1. Material and numerical simulation parameters [13].

Table S2. The optimal hyperparameters of the LightGBM model trained using 8812 features.

Text S1. Brief introduction of FDEM

The FDEM was originally developed by Munjiza in the early 1990s to simulate the material transition from continuum to discontinuum [1]. The essence of this method is to merge the algorithmic advantages of the discrete element method (DEM) with those of the finite element method (FEM). The main theory of FDEM involves the algorithms of governing equations, deformation description, contact detection, and contact interaction [2, 3].

The general governing equation of FDEM is [4]

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} = \mathbf{f} \,, \tag{1}$$

where **M** is the lumped mass matrix, **C** is the damping matrix, **x** is the displacement vector, and **f** is the equivalent force vector acting on each FEM node. This equation solves the dynamic response of a solid material subjected to external forces and automatically satisfies the mass and momentum conservation. An explicit time integration scheme based on a central difference method is employed to solve Eq. (1) with respect to time to obtain the transient evolution of the system. The deformation of finite elements can be described by a multiplicative decomposition-based formulation, which allows for a detailed analysis of material deformation [5]. The contact detection between discrete elements is conducted using the NBS (none binary search) algorithm [6], which determines whether any two given elements, one called the contactor and the other one the target, share at least one search cell. After processing the contact detection, a list that contains all the element pairs potentially in contact is established and sent for contact interaction analysis. A penalty function based contact interaction algorithm is used to calculate the contact forces between contacting elements [4, 7].

It is beyond the scope of this Supplementary Material to provide a complete description of the above principles; however, details of these can be found in FDEM monographs [4, 5, 7]. FDEM allows the explicit geometric and mechanical realization of systems involving both continua and discontinua, making it superior to FEM and DEM. Since its inception [1], FDEM has proven its computational efficiency and reliability, and has been extensively used in a wide range of endeavors in both industry and academia [8-13].

Text S2. Supplementary explanation of FDEM model setup

The geometry of the FDEM model is built based on a two-dimensional laboratory experiment conducted by Geller, et al. [14]. Here, for simplicity, two-dimensional plane stress conditions are assumed. A total of 2,817 non-broken circular particles (arranged in 9 rows and 313 columns before consolidation) are randomly generated and placed in space between the two identical plates. According to the existing literature on the simulation of stick-slip problems [15, 16], the number of particles adopted in this work is appropriate. To avoid significant model distortion caused by shearing, two stiff bars are attached to the top and bottom plates.

The material parameters are similar to the laboratory experiment [14]. The particle-particle and particle-plate friction coefficients are set to 0.15 based on previous studies showing that a smaller friction coefficient allows for larger slip events and reduces the frequency of small fluctuations in the macroscopic friction signal [15]. In terms of the penalty parameter between contacts (e.g., particle-particle and particle-plate contact), it theoretically should be infinity to avoid penetration between contacting elements; however, a large penalty parameter will yield a significantly small time step. A recent study has shown that a penalty parameter about 1-2 orders of Young's modulus generally can ensure computational correctness [17]. By compromising between achieving the correct elastic response between contact elements and maximizing the time step size to reduce the overall computational expense, a penalty parameter of 4 GPa is used.

In the first phase of the simulation, the granular fault gouge is consolidated by moving the top and bottom stiff bars towards each other. The top stiff bar is then fixed in space, and the normal load acting on the bottom stiff bar is gradually increased until it reaches the prescribed value P. The entire model then undergoes a dynamic relaxation phase where the system's kinetic energy is gradually reduced while the system settles. The model is considered to have reached equilibrium when the kinetic energy has stabilized. To alleviate model distortion and to assure the effectiveness of the normal load and shear motion on the granular fault gouge, during the shearing stage, the top stiff bar is allowed to move only in the x direction and the bottom stiff bar is allowed to move only in the y direction.

In the FDEM model, three-node constant-strain triangular elements are used for the particles and the plates. A total of ~90,000 triangle elements are generated in the simulation domain. Each particle is

composed of 24 approximately equal size elements. This number of elements allows each particle to maintain a nearly circular shape after meshing, and is sufficient to precisely capture the particle deformation while assuring the model is not too computationally expensive. The finite elements used in the present study are assumed to follow an isotropic elastic relationship, and the objects such as particles and plates behave as elastic materials.

The inertial number of the system can be calculated using

$$I = \frac{\dot{\gamma} d}{\sqrt{P/\rho}} \tag{2}$$

to quantify the significance of dynamic effects in a granular material [18]. Here $\dot{\gamma}$ is the shear rate, i.e.,

$$\dot{\gamma} = \frac{V}{H_{\mathrm{m}}}, \tag{3}$$

where V is the shear velocity, and for the soft system used here H_{m} is the height of the model which is approximately 520 mm, d is the average particle diameter, P is the normal load, and ρ is the density. The parameters of the fault gouge system give an inertial number ~9.0E-6, which guarantees that the simulation falls into the quasi-static flow regime.

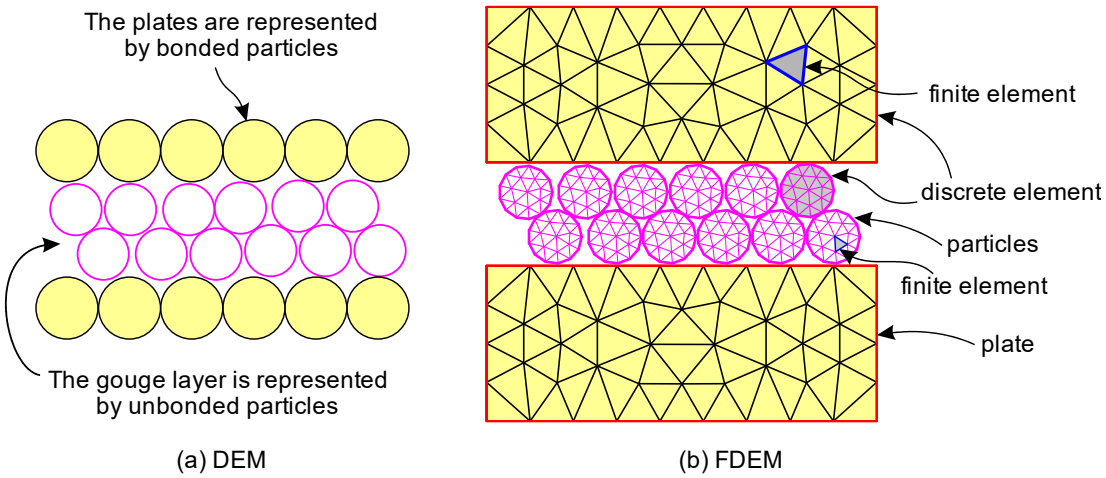

Figure S1. Numerical representation of the granular fault system using DEM and FDEM [13]. (**a**) In DEM, the plates are simplified by a set of bonded particles, and the gouge layer is composed of a series of rigid particles. Therefore, both the gouge particles and the shearing plates cannot deform. (**b**) In FDEM, the plates are explicitly represented, and both plates and particles are further discretized into finite elements to capture their detailed deformation and movement.

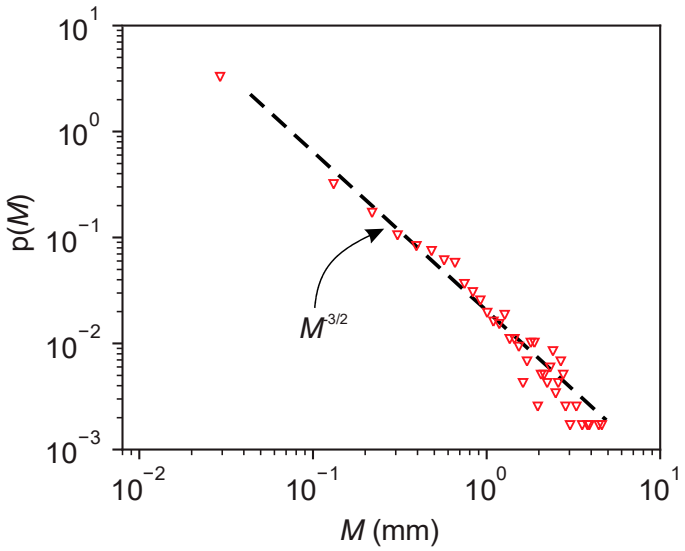

Figure S2. Probability density distribution of the seismic moment of all slip events, where the detailed calculation of the moment is explained by Gao, et al. [13]. The results agree with the physical experiment data collected in Geller, et al. [14]. The probability density distribution is consistent with the Gutenberg-Richter distribution [19] and is predicted to scale as $M^{-3/2}$ (the power -3/2 is within the range of -1.4 to -1.8 observed in natural earthquakes [14].

Table S1. Material and numerical simulation parameters [13].

Parameter	Value	Parameter	Value
Particle diameter	1.2 or 1.6 mm	Stiff bar density	2800 kg/m^3
Particle density	1150 kg/m^3	Stiff bar Young's modulus	30 GPa
Particle Young's modulus	0.4 GPa	Stiff bar Poisson's ratio	0.33
Particle Poisson's ratio	0.4	Foam density	1150 kg/m^3
Particle-particle friction coefficient	0.15	Foam Young's modulus	1 MPa
Number of Particles	2817	Foam Poisson's ratio	0.4
Main plate density	1150 kg/m^3	Contact penalty	4 GPa
Main plate Young's modulus	2.5 MPa	Timestep	10^{-7} s
Main plate Poisson's ratio	0.49	Normal load P	28 kPa
Particle-plate friction coefficient	0.15	Shear velocity V	0.5 mm/s

Text S3. LightGBM (Light Gradient Boosting Machine)

The LightGBM [20] is an ensemble learning model, a fast, distributed, high-performance gradient boosting framework based on the Gradient Boosting Decision Tree (GBDT) algorithm. LightGBM splits the tree leaf-wise with the best fit, unlike other boosting algorithms in which depth-wise or level-wise approaches are used [20]. The leaf-wise algorithm can produce much more complex trees, reduce more loss than the level-wise algorithm, and hence result in much better accuracy that is not easily achievable by other boosting algorithms, such as the XGBoost [21]. Besides, LightGBM uses a histogram-based

algorithm that buckets continuous feature values into discrete bins to speed up the training procedure. The possibility of overfitting can be reduced by setting the max depth parameter. Compared to XGBoost, especially for large datasets, LightGBM performs equally well but with a significant reduction in training time. Additionally, LightGBM uses a novel Gradient-based One-Side Sampling (GOSS) technique to filter out the data instances for finding appropriate split values. In contrast, XGBoost uses a pre-sorted algorithm for computing the best split. In general, the histogram-based split algorithm, where the bins of all data points are used to find the split values of the histogram, is more efficient than the pre-sorted algorithm where all possible split points on the pre-sorted feature values are enumerated.

In each iteration of LightGBM, the negative gradients of the loss function concerning the output of the model are denoted as g_i, which can be illustrated as

$$\begin{aligned} g_i &= \partial_{\hat{y}^{(t-1)}} \operatorname{loss}\left(y_i, \hat{y}^{(t-1)}\right) \\ h_i &= \partial^2_{\hat{y}^{(t-1)}} \operatorname{loss}\left(y_i, \hat{y}^{(t-1)}\right) \end{aligned}. \tag{4}$$

The decision tree model splits each node with the largest information gain for each feature. With the GOSS method, the information gain measured by the variance of splitting features j at point d for node n can be defined as

$$V_{j|O}(d) = \frac{1}{n}\left(\frac{\left(\sum_{\{x_i \in A_l; x_{ij} \le d\}} g_i + \frac{1-a}{b} \sum_{xi \in B_l} g_i \right)^2}{n_l^j(d)} + \frac{\left(\sum_{\{x_i \in A_r; x_{ij} \le d\}} g_i + \frac{1-a}{b} \sum_{xi \in B_r} g_i \right)^2}{n_r^j(d)} \right). \tag{5}$$

The GOSS method keeps the top $a \times 100\%$ instances with the larger gradients and yields an instance subset A. After that, the remaining set A^c consisting of $(1\text{-}a) \times 100\%$ instances have smaller gradients. GOSS method randomly samples a subset B with size $b \times A^c$. Thus, the method estimates information gain over a much smaller sample subset, in contrast to other methods such as pre-sort sampling used in XGBoost [21] that calculate accurate information gain through all samples. As a result, the computation cost of LightGBM can be significantly reduced.

Table S2. The optimal hyperparameters of the LightGBM model trained using 8812 features.

Parameter	Value
Learning rate	0.008811
Max depth	15
Number of leaves	10
Max bin	199
Feature fraction	0.226693
Bagging fraction	0.087749
Lambda L1	4.25191
Lambda L2	0.378177
Min sum hessian in leaf	0.096908
Min data in leaf	351
Bagging frequency	5

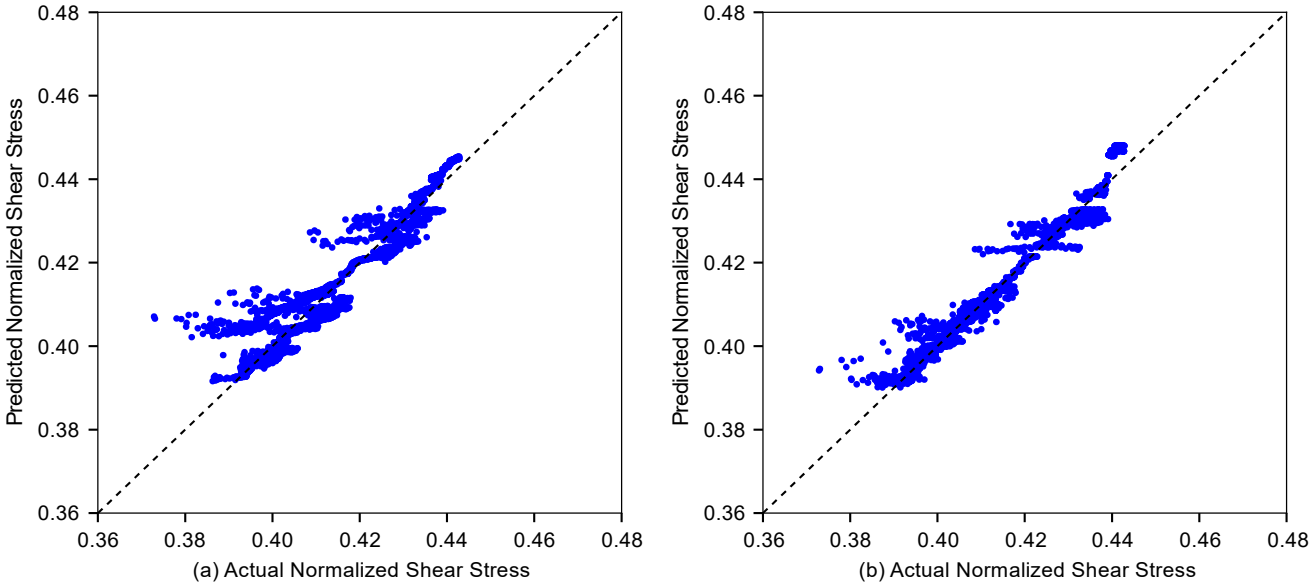

Figure S3. Performance of the two LightGBM models trained with different datasets. (**a**) Training with optimized sensor data. (b) Training with optimized sensor data and their statistics.

Figure S4. The SHAP value versus the feature value at each instant of time for the top 200 input features with relatively larger cumulative magnitudes of the yielded SHAP values. The results are based on the last LightGBM model trained using the optimized dataset and their statistics (*see the separate pdf file at the end*).

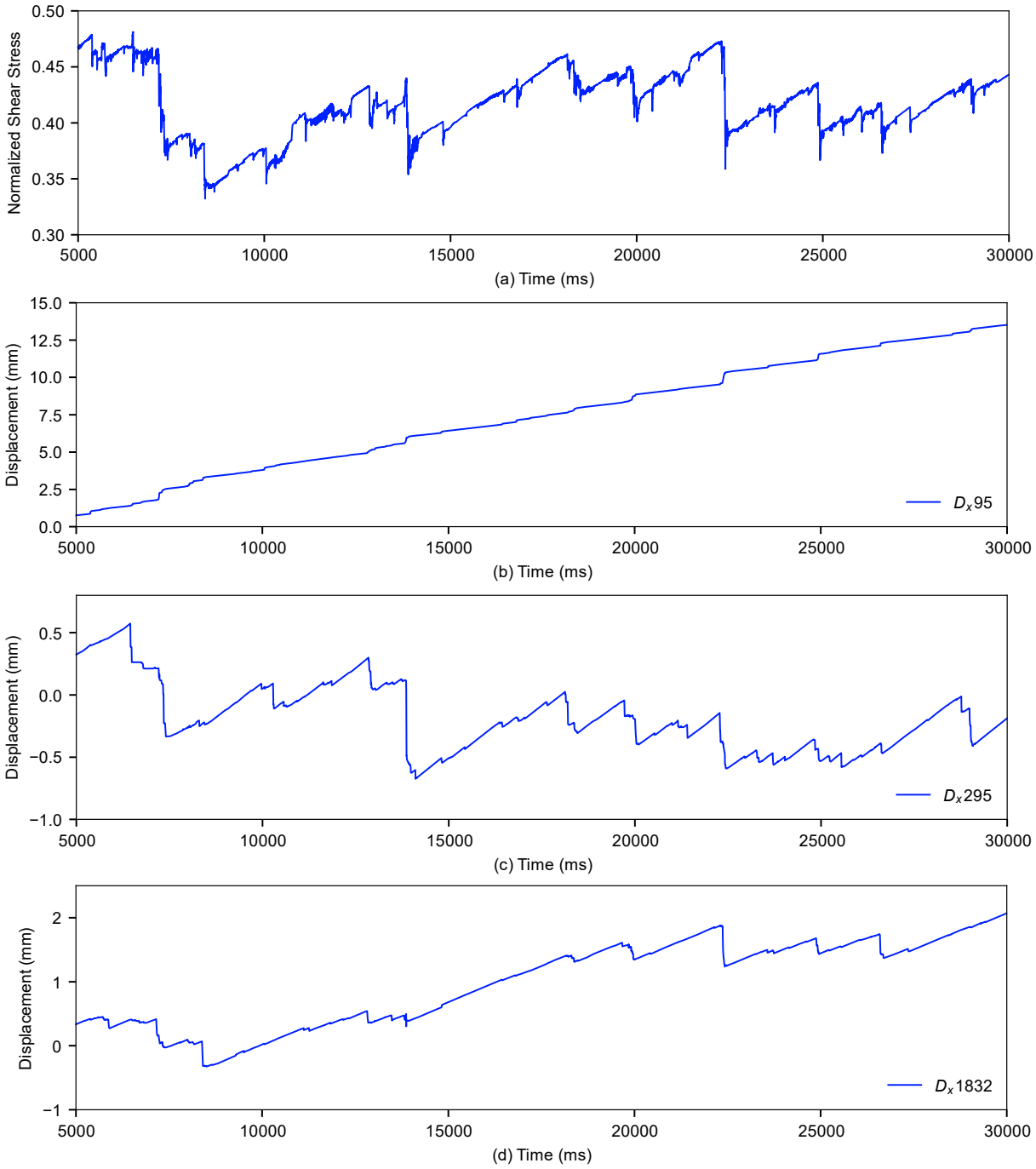

Figure S5. Evolutions of D_x for several sensors and their comparison with the normalized shear stress. (**a**) Normalized shear stress. (**b**) D_x of sensor No. 95. (c) D_x of sensor No. 295. (d) D_x of sensor No. 1832.

References

1. Munjiza, Antonio A. Discrete Elements in Transient Dynamics of Fractured Media. PhD Thesis, Swansea University, 1992.
2. Lei, Z., Esteban Rougier, E. E. Knight, Antonio A. Munjiza, and H. Viswanathan. "A Generalized Anisotropic Deformation Formulation for Geomaterials." *Computational Particle Mechanics* 3, no. 2 (2016): 215-28.
3. Munjiza, A., E. Rougier, and N. W. M. John. "Mr Linear Contact Detection Algorithm." *International Journal For Numerical Methods In Engineering* 66, no. 1 (2006): 46-71.
4. Munjiza, Antonio A. *The Combined Finite-Discrete Element Method*: John Wiley & Sons, 2004.
5. Munjiza, Antonio A., Esteban Rougier, and Earl E Knight. *Large Strain Finite Element Method: A Practical Course*: John Wiley & Sons, 2014.
6. Munjiza, A, and KRF Andrews. "Nbs Contact Detection Algorithm for Bodies of Similar Size." *International Journal For Numerical Methods In Engineering* 43, no. 1 (1998): 131-49.
7. Munjiza, Antonio A., Earl E. Knight, and Esteban Rougier. *Computational Mechanics of Discontinua*: John Wiley & Sons, 2011.
8. Lei, Zhou, Esteban Rougier, Antonio Munjiza, Hari Viswanathan, and Earl E. Knight. "Simulation of Discrete Cracks Driven by Nearly Incompressible Fluid Via 2d Combined Finite-Discrete Element Method." *International Journal For Numerical And Analytical Methods In Geomechanics* 43, no. 9 (2019): 1724-43.
9. Lei, Qinghua, and Ke Gao. "Correlation between Fracture Network Properties and Stress Variability in Geological Media." *Geophysical Research Letters* 45 (2018): 3994– 4006.
10. Rougier, Esteban, Antonio Munjiza, Zhou Lei, Viet Tuan Chau, Earl Eugene Knight, Abigail Hunter, and Gowri Srinivasan. "The Combined Plastic and Discrete Fracture Deformation Framework for Fdem." *International Journal For Numerical Methods In Engineering* (2019).
11. Euser, Bryan, E. Rougier, Z. Lei, E. E. Knight, L. P. Frash, J. W. Carey, H. Viswanathan, and A. Munjiza. "Simulation of Fracture Coalescence in Granite Via the Combined Finite–Discrete Element Method." *Rock Mechanics and Rock Engineering* 52, no. 9 (2019): 3213-27.
12. Okubo, Kurama, Harsha S. Bhat, Esteban Rougier, Samson Marty, Alexandre Schubnel, Zhou Lei, Earl E. Knight, and Yann Klinger. "Dynamics, Radiation and Overall Energy Budget of Earthquake Rupture with Coseismic Off-Fault Damage." *Journal of Geophysical Research: Solid Earth* (2019).
13. Gao, Ke, Bryan J Euser, Esteban Rougier, Robert A Guyer, Zhou Lei, Earl E Knight, Jan Carmeliet, and Paul A Johnson. "Modeling of Stick-Slip Behavior in Sheared Granular Fault Gouge Using the

Combined Finite-Discrete Element Method." *Journal of Geophysical Research: Solid Earth* 123 (2018): 5774– 92.

14. Geller, Drew A., Robert E. Ecke, Karin A. Dahmen, and Scott Backhaus. "Stick-Slip Behavior in a Continuum-Granular Experiment." *Physical Review E* 92, no. 6 (2015): 060201.
15. Ferdowsi, B. Discrete Element Modeling of Triggered Slip in Faults with Granular Gouge. Application to Dynamic Earthquake Triggering. Ph.D Thesis, ETH-Zürich, Switzerland, 2014.
16. Dorostkar, Omid, Robert A. Guyer, Paul A. Johnson, Chris Marone, and Jan Carmeliet. "On the Role of Fluids in Stick-Slip Dynamics of Saturated Granular Fault Gouge Using a Coupled Computational Fluid Dynamics-Discrete Element Approach." *Journal of Geophysical Research: Solid Earth* 122, no. 5 (2017): 3689-700.
17. Tatone, B. S. A., and G. Grasselli. "A Calibration Procedure for Two-Dimensional Laboratory-Scale Hybrid Finite–Discrete Element Simulations." *International Journal of Rock Mechanics and Mining Sciences* 75, no. Supplement C (2015): 56-72.
18. MiDi, GDR. "On Dense Granular Flows." *The European Physical Journal E* 14, no. 4 (2004): 341-65.
19. Gutenberg, B., and C. F. Richter. "Magnitude and Energy of Earthquakes." *Nature* 176 (1955): 795.
20. Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. "Lightgbm: A Highly Efficient Gradient Boosting Decision Tree." In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 3149–57. Long Beach, California, USA: Curran Associates Inc., 2017.
21. Chen, Tianqi, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, and Kailong Chen. "Xgboost: Extreme Gradient Boosting." *R package version 0.4-2* 1, no. 4 (2015): 1-4.