Machine Learning Predicts the Timing and Shear Stress Evolution of Lab Earthquakes Using Active Seismic Monitoring of Fault Zone Processes

Srisharan Shreedharan[1*], David Chas Bolton[1], Jacques Rivière[2], Chris Marone[1,3]

[1]Dept. of Geosciences, Pennsylvania State University, University Park, PA 16802 USA

[2]Dept. of Engineering Science and Mechanics, Pennsylvania State University, University Park, PA 16802 USA

[3]Dipartimento di Scienze della Terra, La Sapienza Università di Roma, Italy

***Corresponding author: srisharan@psu.edu**

Abstract:

Machine learning (ML) techniques have become increasingly important in seismology and earthquake science. Lab-based studies have used acoustic emission data to predict time-to-failure and stress state, and in a few cases the same approach has been used for field data. However, the underlying physical mechanisms that allow lab earthquake prediction and seismic forecasting remain poorly resolved. Here, we address this knowledge gap by coupling active-source seismic data, which probe asperity-scale processes, with ML methods. We show that elastic waves passing through the lab fault zone contain information that can predict the full spectrum of labquakes from slow slip instabilities to highly aperiodic events. The ML methods utilize systematic changes in p-wave amplitude and velocity to accurately predict the timing and shear stress during labquakes. The ML predictions improve in accuracy closer to fault failure, demonstrating that the predictive power of the ultrasonic signals improves as the fault approaches failure. Our results demonstrate that the relationship between the ultrasonic parameters and fault slip rate, and in turn, the systematically evolving real area of contact and asperity stiffness allow the gradient boosting algorithm to 'learn' about the state of the fault and its proximity to failure. Broadly, our results demonstrate the utility of physics-informed machine learning in forecasting the imminence of fault slip at the laboratory scale, which may have important implications for earthquake mechanics in nature.

Keywords:

Stick-slips; slow earthquakes; Friction; Machine Learning; Gradient boosted trees; XGBoost

Key points:

1. Machine learning can be used on p-wave amplitude and velocity to predict the timing and shear stress evolution of laboratory seismicity.
2. The ability of the ML algorithm to predict time-to-failure improves as the fault approaches failure.
3. Predictions rely on the systematic reduction in elastic properties prior to failure, which is linked to a reduction in real area of contact.

1. Introduction:

Machine learning (ML) methods have followed improvements in geophysical techniques, instrumentation and data availability over the past decade to rapidly emerge as indispensable toolkits for the geophysical community (*Bergen et al., 2019; Kong et al., 2019*). For instance, significant effort has been devoted to using ML to improve event detection, arrival-time picking, phase association, and earthquake location (*Yoon et al., 2015; Zhu and Beroza, 2018; Ross et al., 2018; Perol et al., 2018; Wu et al., 2018; McBrearty et al, 2019; Trugman and Ross, 2019; Mousavi et al., 2020*). Importantly, this has seen a revival in earthquake forecasting research, particularly focused on applying ML techniques to lab data on fault friction (*Rouet-Leduc et al., 2017, 2018*) and rock damage (*McBeck et al., 2020*) to infer failure modes and predictability (*Corbi et al.,* 2019). Notably, recent studies have successfully demonstrated that both cataloged (*Lubbers et al., 2018*) and continuous (*Rouet-Leduc et al., 2017; Hulbert et al., 2019*) acoustic emission (AE) data can be used to infer fault friction and predict the timing, shear stress, and in some cases the size of labquakes. Moreover, ML has been used to demonstrate that slow and fast earthquakes share similar physics (*Hulbert et al., 2019*). With varying degrees of success, these techniques can predict field observations of volcanic eruption (*Ren et al., 2020*) and subduction zone fault slip (*Rouet-Leduc et al., 2019; Corbi et al., 2020; Hulbert et al., 2020*).

Broadly, ML techniques can be grouped as supervised or unsupervised. The former involves predetermined features that are mapped to labeled datasets in order to construct a regression model that typically involves highly non-linear functions. Unsupervised learning models are used where such labeled datasets may not be available or the focus is on trying to identify patterns embedded within the data (e.g., Tan et al., 2006; Oliver et al., 2018; Bolton et al., 2019). Especially as applied to laboratory data, supervised ML techniques have relied on

systematic evolution of 'features' such as the AE energy (e.g., *Rouet-Leduc et al., 2017, 2018*). Indeed, based on the existing works, one could conclude that systematic changes in microseismicity (AE) indicative of fault zone criticality are required to successfully predict labquakes or tectonic fault slip. However, while precursors are routinely documented in laboratory studies of fault failure and brittle fracturing (*Brace et al., 1966; Scholz, 1968a; Sammonds et al., 1992; Hedayat et al., 2014; Kaproth and Marone, 2013; Scuderi et al., 2016; Shreedharan et al., 2020*), observations of systematic precursors or foreshocks prior to earthquakes in nature are not routinely documented (*Bakun et al., 2005; Niu et al., 2008; Main et al., 2012*). Additionally, even when precursory slip is present, it may be masked by changes in elastic properties of the wallrock (*Chiarabba et al., 2020; Shreedharan et al., 2021*). Moreover, the underlying physics of ML prediction of lab earthquakes is poorly understood. One recent study by *Bolton et al. (2020)* demonstrated that the precursory increase in AE energy prior to fault failure is likely linked to preseismic fault slip. Thus, the question arises of whether detection of preseismic fault slip and lab foreshocks are necessary conditions for ML-based prediction of lab earthquakes, and ultimately, whether the same is true for field observations.

Here, we use high-resolution time-lapse active seismic monitoring to document the evolution of p-wave amplitude and velocity during the lab seismic cycle of highly variable and aperiodic labquakes. We find that the gradient boosting ML algorithm (*Friedman, 2001*) can accurately predict both the timing and shear stress state of labquakes using a subset of the amplitude and velocity features. Interestingly, our results show that the ML predictive power improves as the fault approaches failure. Our previous works and those of others have established the physical links between systematic variations in ultrasonic attributes (p-wave amplitudes and velocities), fault zone preslip and wallrock stiffening throughout the laboratory

seismic cycle (*Shreedharan et al., 2020, 2021*). Further, these ultrasonic attributes can be linked to asperity scale mechanical parameters such as the real area of contact and asperity stiffness (*Kendall and Tabor, 1971; Pyrak-Nolte et al., 1990; Hedayat et al., 2014; Kilgore et al., 2017; Shreedharan et al., 2019*). Thus, by successfully training our ML algorithm on these ultrasonic attributes, we assign a physical underpinning for the predictive power of the ML approach.

2. Methods:

2.1.Friction experiments

We performed frictional shear experiments in a servo-controlled biaxial testing apparatus using a double-direct shear (DDS) configuration (*Karner and Marone*, 1998). We sheared two frictional interfaces created by mating three blocks of Westerly granite, with a fine layer (<200 µm thickness and ~0.25 g/layer by mass) of quartz powder between the interfaces to simulate frictional wear material (Figure 1; Also see *Shreedharan et al., 2020*). The granite surfaces were roughened with #60 grit silicon carbide thus producing a mean roughness of ~20 µm. The surface roughness and the dusting of quartz powder (median particle size of 10.5 µm) have comparable dimensions; thus frictional processes include direct contact of the wall rock, wear, and internal deformation in the gouge. The fault normal stress was held constant via a fast-acting servocontroller. The biaxial testing apparatus is fully servo-controlled with independent hydraulic pistons supplying normal and shear loads. The experiments were instrumented with calibrated strain-gauge load cells to measure normal and shear stresses, and direct current differential transformers (DCDTs) to measure fault normal and shear displacement. An additional DCDT was attached to the center block close to one of the frictional interfaces, to measure fault slip (Figure 1). All load cells used in this study have a resolution of ±5 N and the

DCDTs have a displacement resolution of ±0.1 μm. We acquired mechanical data at 10 kHz and averaged in real-time to 100 or 1000 Hz prior to recording. A constant shear displacement rate was prescribed for the longer, central block of the DDS configuration. In all experiments reported here, the samples had a constant nominal frictional contact area of 25 cm^2 and were subjected to a 10 MPa normal stress. The prescribed background shear rate was set at 11 μm/s. All experiments were performed at a nominal room temperature range of 22 – 24 ^oC and the sample humidity was maintained at 100% to ensure reproducibility. We produced a spectrum of slip modes from slow to fast (Inset to Figure 1; *Leeman et al., 2016*) by varying the stiffness of the shear loading system (*Shreedharan et al., 2020*).

2.2.Ultrasonic monitoring

We conducted continuous seismic monitoring via ultrasonic pulses transmitted through the fault interfaces. We used 500 kHz broadband, P-polarized lead-zirconate-titanate (PZT) crystals (Boston Piezo-Optics Inc. PZT-5A 0.5" diameter) and transmitted half-sine pulses every 1 ms continuously throughout frictional shear. The PZT sensors were epoxied in blind holes within steel platens and positioned adjacent to the granite side blocks of the DDS configuration (See inset to Figure 1). Each pulse was sampled at 25 MHz, thus ensuring high temporal resolution in sampling. We use the first P-wave arrival to calculate travel-times and velocities, and take the largest peak-to-peak amplitude within the first 5 μs for acoustic transmissivity (Inset to Figure 1). Following *Nagata et al. (2014),* we report transmissivity, |T|, as

$$|T| = \sqrt{\frac{A_{DDS}}{A_{Intact}}} \quad (1)$$

In Eq. (1), the square root term accounts for the two frictional interfaces that the ultrasonic pulses traverse (*Shreedharan et al., 2019*). We calculate the velocity as the ratio of the

distance traversed by the wavelet to the travel time through the DDS blocks, after accounting for the time spent in steel and at interfaces. We refer readers to *Shreedharan et al. (2021)* for a detailed description of our velocity and transmissivity measurements.

2.3.ML model – training and testing

We analyze the acoustic and mechanical data using ML to predict the temporal evolution of shear stress and the time remaining prior to failure for multiple laboratory seismic cycles (Figure 2). Because we use a supervised learning approach, our first step involves data preparation and selection of features and labelled datasets. Previous studies have demonstrated that the P-wave amplitudes and velocities evolve systematically throughout the lab seismic cycle. This, in addition to their direct proxy relationship to asperity deformation mechanics, makes them good candidate features for our model. The labelled data, i.e., the 'unknown' datasets that we want to predict are the fault shear stress, time since the previous failure event and, importantly, the time remaining until failure. Figures 2a-b show the labelled data and Figures 2c-d show the features used in this study. Note that while the shear stress is readily determined as part of regular data collection during the experiment, the time-since-failure and time-to-failure labels are measured from the shear stress data. The time-since-failure for an event is computed from the shear stress minimum of the previous event, which has a time-since-failure label of 0 s, to the shear stress maximum for the next event. Conversely, the time-to-failure for each stick-slip cycle is counted from the shear stress minimum at the end of the previous cycle, up to the shear stress maximum corresponding to the current cycle, which represents a time to failure of 0 s (Figure 2b). During the coseismic portion of the cycle, i.e., the stress drop, the time-since and time-to-failure are set to zero and these data are not used for the ML regression.

As part of feature preparation, we smooth the features using a backward-looking 10-point moving average to reduce feature-side noise in the ML regression. We then normalize the features for each labquake against the minimum value of the feature at the end of the previous cycle (Figure 2e-f). This is done to ensure that long-term trends in the features due to layer-thinning, shear localization, wear product formation and smoothening of the granite surfaces do not overprint on their shorter-term evolution during the slip instabilities. Further, this ensures that the features in any slip instability are independent and have no 'memory' of the fault's previous state in an earlier cycle. As a final feature engineering step, we estimate a time-advanced version of the feature (amplitude or velocity). This is done by shifting the feature vector forward by five time steps (0.05 s). We use the original feature corresponding to time t and a time-advanced feature corresponding to time $t - 0.05\ s$ in our regression. The temporal evolution of the amplitude and velocity features includes an increasing and comparably decreasing gradient. When regressing these features against a monotonically varying label such as shear stress or time to failure, there ceases to exist a one-to-one functional mapping, i.e., non-unique solutions. This results in poor performance by regression-based supervised learning techniques, as noted by previous works (*Rouet-Leduc et al.*, 2017) because, for a given stick-slip, two data labels can correspond to the same feature value. We conducted a series of benchmarking tests and found that this 'offsetting' procedure solves the problem by informing the ML models whether the feature corresponding to a data label comes from the increasing or decreasing gradient space. Additionally, because we use a time-advanced offset, the new offset feature is backward-looking in time, thus eliminating any potential data leakage issues. This method is similar to and was derived from the sub-windowing procedure utilized by *Hulbert et al. (2019)* to solve a similar functional mapping problem in their ML regression. We note here

that the offsetting is a mathematical transformation rather than a physical one. Hence, the choice of an *n*-point offset is arbitrary and this parameter, in general, can and should be optimized to ensure good fits to the labelled data. We selected a 5-point offset for this study since it provided excellent fits to the test data without overfitting the training data.

Subsequently, we split our labeled dataset and features into training and testing sets (Figure 1). In this study, we report on the results from a 70-30 split in a contiguous fashion, i.e., the initial 70% of the data were designated as the training set and the remaining 30% as the test dataset. Note that, unlike previous ML studies on experimental faults, which make use of windowing procedures to estimate features (*Rouet-Leduc et al., 2017; Hulbert et al., 2019; Bolton et al., 2019; Corbi et al, 2019*), our mechanical and ultrasonic data are synchronized in time. In other words, for each temporally evolving value of the labels (shear stress, time-since-failure, time-to-failure), there exists a corresponding data-point in the feature-space (amplitudes, velocities).

We use the gradient boosting ML algorithm based on decision trees (*Friedman, 2001; Hulbert et al., 2019*), to jointly analyze our acoustic and mechanical datasets. Specifically, we utilize an open-source implementation of this algorithm named XGBoost (*Chen and Guestrin, 2016*). Hyperparameter tuning is the first step in implementing this model. Hyperparameters define and determine the parameter-space of functions that can serve as potential models, and the model performance is significantly sensitive to the assigned hyperparameters. The XGBoost implementation has a suite of hyperparameters which must be optimized prior to training. In this study, we determine the optimal hyperparameters by implementing an Efficient Global Optimization (EGO) function to minimize a misfit function (*Jones et al., 1998*). The optimal hyperparameters are determined via a five-fold cross-validation. Here, a subset of the training

data is modeled using a given set of hyperparameters and validated against the remaining training data. Once the hyperparameter 'tuning' step is complete, we train the ML model by performing regressions on the feature-label relationship. This step involves iteratively developing the structure of the gradient boosted decision trees. For a detailed description of this method, we refer the readers to the XGBoost documentation (xgboost.readthedocs.io) and *Hulbert et al. (2019)*. Finally, the optimal model is tested and its performance is compared against the labels, i.e., the true experimental values of shear stress and time to failure. We evaluate the model performance in two ways: (1) a qualitative benchmark against a naïve constant-recurrence interval model for the time-to-failure and (2) a quantitative estimation of the model performance using the coefficient of determination (R^2) and the root mean squared error (RMSE) for all labelled data.

3. Results:

3.1.Friction data and stick-slip periodicity

Our lab results include the full spectrum of frictional slip modes from slow to fast labquakes. In particular, the experimental data are characterized by a range of aperiodic frictional slip instabilities (Figure 3) at the friction stability boundary (e.g., *Gu et al.*, 1984; *Leeman et al., 2016)*. Figure 3a quantifies this frictional chaos by plotting the inter-event times as a cross-plot of time since the last instability T_{prev} and time to the next instability T_{next} (See inset to Figure 1). Data that fall on the 1:1 line represent perfect periodicity, i.e., they are time and slip predictable events (eg. *Shimazaki and Nakata, 1980*). The 1:2 and 2:1 lines form an envelope representing period doubling (*Veedu et al., 2020*). Further, data points are colored by the magnitude of stress drop for these events. Broadly, while our data have periodicities between perfectly periodic and

doublets, the higher stress drop events are generally represented by more aperiodic behavior while the smaller events cluster tightly around the 1:1 line.

We observe that stick-slip instability behavior and frictional chaos evolves with shear displacement, represented by increasing event number (Figure 3). Event numbers are calculated from the ML train-test catalog (Figure 1), i.e., between load point displacements of 18 – 21 mm, which contains 220 events. While the peak stress remains relatively constant throughout shear, the stress minimum during the coseismic phase increases with shear (Figure 3b). In other words, the fault starts off in a quasi-stable condition and becomes increasingly unstable. This is a common observation related to shear driven reduction in the friction critical slip distance (Dc) and increase in the friction rate parameter, *(b-a)*, within the rate-state frictional framework (e.g., *Marone, 1998*b). The inter-event ratio from Figure 3a is plotted in Figure 3c as a function of event number, which is used here as a proxy for shear displacement. For the first ~50 events, the slip behavior is relatively periodic. As the stress drop increases with shear, the slip behavior becomes increasingly complex which makes this an ideal dataset to challenge ML approaches for prediction.

3.2.Co-evolution of friction and elastic properties

Fault zone elastic properties and frictional strength co-evolve in a systematic manner during the lab seismic cycle (Figure 4). The fault zone elastic wave amplitude and velocity co-evolve with shear stress and fault slip rate during both slow and fast slip lab earthquakes (Figure 4a). The corresponding fault displacement, obtained from an onboard displacement sensor (Inset to Figure 1) and slip velocity, measured from the time derivative of this fault slip, also increase systematically as the fault approaches failure (Figure 4b). For our loading rate of 11 μm/s the

fault experiences peak slip rates of ~300 μm/s, and interseismic locking rates of under 0.1 μm/s for the largest, fastest events. The corresponding transmissivity and p-wave velocities are shown in Figures 4c-d. Note the strong inverse correlation between fault slip rate and the ultrasonic attributes, which is consistent with observations from previous ultrasonic studies of laboratory stick-slips (*Nagata et al., 2014; Kilgore et al., 2017; Shreedharan et al., 2020, 2021*).

3.3.Machine learning models of shear stress, time-since and time-to-failure

We document the training, testing and performance metrics of the XGBoost models for shear stress, time-since-failure and time-to-failure labels in Figures 5-7. For each label, we train and test three ML models: one for transmitted wave amplitude, one for velocity, and one for the combined features of amplitude and velocity. We do not use a separate validation set; rather we perform a five-fold cross-validation on the training datasets.

We report the ML results of shear stress prediction using transmissivity (and its offset) as the sole feature (Figure 5a-b) and velocity (and its offset) as the sole feature (Figure 5d-e). Here, we quantify the model performance using a standard R^2 metric (Figure 5). Models trained with both amplitude and velocity have higher R^2 metrics for training compared to the test set (Figure 5), which is nominally expected since the models are bound to perform better on datasets they have previously 'seen.' Regardless, the test set has a reasonable model performance of $R^2 = 0.80$ for both amplitude and velocity-based ML models. Cross-plots of ML model estimates and experimentally measured values of shear stress show the model performance (Figures 5c,f), with the solid (1:1) line representing perfect predictions. We observe that the models tend to deviate from the experimental data early in the seismic cycle, i.e., the shear stress minima, and close to the middle of the cycle, approximately where the amplitudes and velocity features switch from

an increasing to decreasing gradient. However, the ML model performance improves as time to failure decreases (Figure 5g/h). For both amplitude and velocity-based ML models the root mean squared error (RMSE) calculated over a 10-point moving window shows somewhat poor performance early in the lab seismic cycle (i.e., at ~ 3555.4 s) but improved performance later in the cycle, i.e., between 3556 – 3557 s (Figure 5g/h).

Figure 6 shows a snapshot of the results from a training and testing exercise on the time-since-failure data label. The ML performance using amplitudes and velocities as features are shown in Figures 6a-c and 6d-f respectively. The ML models perform better during the early parts of the seismic cycle (Figure 6c/f). In other words, the time-since-failure label is best predicted by the early portions of the amplitude and velocity evolution, immediately following a seismic event. Zooms of representative time-since-failure evolution over one seismic cycle, show the corresponding ML model and 10-point windowed RMSE evolution (Figures 6g-h). These plots show the superior fits to the experimental data early in the seismic cycle. Specifically, the RMSE is lowest (or nearly 0) in the initial ~1 s following a stick-slip event and then continually increases until the next stick-slip event.

ML models focused on the time-to-failure label are the most pertinent for earthquake forecasting. Figures 7a-c show the results from a ML model created using amplitudes as the primary feature, and Figures 7d-f show the results form a model using velocities. Generally, we document poor performance early in the cycle (see Figures 7c,f between 2-4 s) when the time-to-failure is highest. The model fits improve as the fault approaches failure. Figures 7g-h show representative stick-slip cycles with their associated RMSE for the amplitude- and velocity-trained ML models. Again, we document a reducing RMSE as the fault approaches failure, i.e.,

as the time-to-failure approaches zero, particularly in the final ~0.8 s or final ~22% of the interseismic period.

3.4.Benchmarking performance against a naïve model

As a final performance evaluation, we benchmark our ML results against a simple model of recurrence times constructed from an averaged recurrence interval for all events in a given experiment (Figure 8). Because it assumes a statistical average of all recurrence intervals, this model makes no implicit assumptions about and has no knowledge of the experiment or fault behavior (Figure 8). With an R^2 = 0.49, the performance of this naïve model is relatively unreliable. When compared with our ML models (Figures 5-7) which have R^2 metrics of 0.8 – 0.9 over the entire slip cycle, it is clear that the XGBoost models deliver superior performance.

4. Discussion:

4.1.A physical basis for the evolution of fault zone elastic properties in the seismic cycle

The evolution of elastic wave properties around fault zones has been extensively studied in the laboratory (eg. *Stanchits et al., 2003; Paterson and Wong, 2005; Passelègue et al., 2018*) and, to a smaller degree, on crustal faults (eg. *Niu et al., 2008; Brenguier et al., 2008; Malagnini et al., 2019; Chiarabba et al., 2020*). Specifically, systematic variations in the p-wave velocity field prior to fault failure in the laboratory over multiple slow and fast cycles has been documented (*Kaproth and Marone, 2013; Tinti et al., 2016*) and attributed to preseismic creep (*Scuderi et al., 2016*). Similarly, variations in p-wave amplitudes (or transmissivity or transmission coefficient) have been documented as arising from preseismic creep (*Hedayat et al., 2014; Shreedharan et al., 2020*). In crustal faults, *Malagnini et al. (2019)* observed a preseismic

attenuation signal in the 20-40 Hz frequency range leading up to the 2004 M6 Parkfield earthquake, which they attributed to fluctuations in the fault zone crack density. However, the ultrasonic amplitudes can be connected to the microscopic asperity stiffness and during stable sliding, the real area of contact (*Kendall and Tabor, 1971; Kilgore et al., 2017; Shreedharan et al., 2019*). Formalized mathematically, when the ultrasonic pulse wavelength is significantly larger than the fault zone width as is the case here, the specific interface stiffness, k_{sp}, is related to the transmissivity, |T|, as (*Pyrak-Nolte et al., 1990; Kilgore et al., 2017*)

$$k_{sp} = \frac{\omega \rho v_p}{\sqrt{\frac{1}{|T|^2} - 1}} \quad (2)$$

Here, ρ is the density of the surrounding medium (Westerly granite) and v_p is the p-wave velocity through this medium. More recently, *Shreedharan et al. (2021)* demonstrated that while the ultrasonic amplitudes track fault creep, seismic velocity contains information about fault creep as well as shear stiffening of the wallrock. This 'duality' of information contained in the seismic velocities has been documented in crustal faults as well (*Chiarabba et al., 2020*).

Based on Eq. (2), the strong inverse correlation between wave amplitude and the fault zone slip rate (Figure 4) can be interpreted as the interseismic and coseismic evolution of fault zone asperity stiffness. More specifically, immediately following a stick-slip event, the fault locks up and heals interseismically (*Dieterich*, 1972; *Marone*, *1998a; Kaproth and Marone*, *2014; McLaskey et al., 2012*). During this period, the increasing amplitudes and velocities can be interpreted as an increase in asperity stiffness due to reduced slip rate. Similarly, the reduction of precursory amplitude and velocity at and after the onset of preslip can be interpreted as a reduction in asperity stiffness (or, perhaps, asperity destruction) due to welded contact junctions being broken by fault slip. While not a prominent feature of this dataset (Figure 4), the amplitudes and velocities are nominally distinctly out of phase due to the additional information

pertaining to shear stiffening contained in the velocity (See Figure 2 of *Shreedharan et al., 2021*). In a physical model, this may be translated as an interseismic and preseismic reduction in wallrock crack density due to increased deviatoric stresses experienced by the wallrock during elastic strain energy build-up. Thus, by training our regression-based models on transmissivity and velocity, we are, by proxy, training our models on the systematic temporal evolution of the interface stiffness and bulk shear stiffening over multiple seismic cycles.

4.2.Connecting ML model response to the physics of frictional sliding

The supervised ML approach used in this study 'learns' from a regression between the features and labels; thus, no temporal information about a prior state of the fault is explicitly transferred from the experimental data to the ML model. At any point in time, the coefficient of friction, μ, of the fault can be expressed within the rate and state frictional framework as a function of the fault slip velocity, V, and a state evolution term, θ, as

$$\mu = \mu_0 + a \ln\left(\frac{V}{V_0}\right) + b \ln\left(\frac{\theta}{\theta_0}\right) \quad (3)$$

Here, a and b are rate-state constants and the subscript '$_0$' denotes a reference variable. The frictional state term, at steady state is generally thought of as an average asperity lifetime and is described as the ratio of a characteristic slip distance, D_c, and the fault slip rate, V. The term θ can only be determined via an iterative inversion procedure by solving Eq. (3) simultaneously with a state-evolution law and an elastic coupling equation. However, Eq. (3) can be rearranged (*Nagata et al., 2012*) to redefine frictional state as φ, where

$$\varphi = b \ln\left(\frac{\theta}{\theta_0}\right) = \mu - \mu_0 - a \ln\left(\frac{V}{V_0}\right) \quad (4)$$

Phase plane cross-plots between the mechanical attributes of the fault (Figure 9a,b), i.e., between the friction coefficient, slip rate and frictional state (estimated using Eq. 4 with a =

0.0091), show the inter-relationships between the rate and state parameters described in Eq. (3) (*Marone, 1998b*). Similarly, cross-plots of the friction coefficient and the ultrasonic attributes (Figure 9c,d) are the relationships studied by the XGBoost models to optimize regressions for shear stress prediction (Figure 5). The phase plane plots in Figure 9 contain hysteretic loops for ~10 slow and fast seismic cycles. Note the strong inverse correlation between the slip rate (Figure 9a) and the ultrasonic attributes (Figure 9c-d). Likewise, the frictional state evolution (Figure 9b) is more directly correlated with these attributes. Qualitatively, our data illustrate the relationships between ultrasonic amplitudes and state suggested by *Nagata et al. (2014)*. This implies a mechanical relationship between the physically determined asperity stiffness and the more empirical frictional state, thought to be an average asperity or contact junction lifetime. However, because our stick-slip cycles represent non-steady state, unstable slip behavior of the fault, no quantitative relationships between these parameters can be derived (*Kame et al., 2014; Shreedharan et al., 2019*).

Here, we consider the potential relationships between the ML model results (Figures 5-7) and the physical basis for the variations in ultrasonic attributes throughout the seismic cycle (Figure 9e). For a multi-contact interface composed of numerous contact junctions interseismic healing phase is marked by increasing contact junction size and number of contacts. During this period the ultrasonic amplitudes and velocities increase (*Ryan et al., 2018*). Subsequently, as the fault creeps prior to failure, some of these contact junctions are destroyed and shrink in size or cease to exist. This is marked by a reduction in the transmissivity and fault zone velocity (*Shreedharan et al., 2021*). Finally, during the coseismic slip phase, a number of contact junctions are broken and the fault slips, releasing the stored strain energy in these asperities.

Recall that, for shear stress prediction (Figure 5) and time-to-failure prediction (Figure 7), the models perform somewhat poorly during the early portions of the seismic cycle, whereas the model performance is remarkably accurate as the fault approaches failure. In the case of time-since-failure prediction (Figure 6), this trend is reversed and the ML model predictions are excellent post-seismically, immediately following a stick-slip event. This is consistent with the observations of *Lubbers et al. (2018)* who documented that the cataloged AEs early in the seismic cycle were better predictors of time-since-failure than time-to-failure. Surprisingly, our observations are in contrast with the model of *Hulbert et al. (2019).* They document an inverse relationship between the duration of the next slip event and the acoustic energy early in the slip cycle of the current event. We report data at the same normal stress and, in turn, the same fractional asperity contact area (*Shreedharan et al., 2019*) for different stick-slip magnitudes whereas *Hulbert et al. (2019)* observe this inverse relationship on data collected at different normal stresses (and thus, different real contact areas). Because stick-slip magnitudes (*Leeman et al., 2016*) and AE amplitudes (and energy) scale with normal stress (*Rivière et al., 2018*), the different trends documented here and by *Hulbert et al. (2019)* could, in part, be explained by normal stress dependence (or a lack thereof) in the experimental designs. Moreover, the relationship between dynamics of AEs and active seismic data are unknown, which makes it harder to compare the results directly. However, we note that *Hulbert et al. (2019)* document better model fits to their shear stress data closer to failure rather than early in the seismic cycle. Thus, our results are qualitatively consistent in this regard. Based on our observations of the accurate model fits early in the cycle for time-since-failure and later in the cycle for time-to-failure, we posit that the early, interseismic ultrasonic attributes have significant predictive power and may contain information about the past state of the fault (Interseismic healing in

411 Figure 9e). This is qualitatively similar to the idea that aftershock duration may be linked to
412 mainshock size (*Dieterich, 1994; Lubbers et al., 2018*) although no evidence for this assertion
413 appears to exist in crustal faults (*Ziv, 2006*). Likewise, the ultrasonic attributes after the onset of
414 preslip likely contain predictive power about the future state of the fault (Preseismic creep phase
415 in Figure 9e), i.e., timing and size of the imminent failure (*Rouet-Leduc et al., 2017;*
416 *Shreedharan et al., 2020*).

417 For each of the ML feature sets we studied, the model performance suffers during the
418 middle of the seismic cycle, as the fault begins to unlock. We propose that this occurs because
419 the frictional asperities experience a competition between healing, which strengthens contact
420 junctions, and preslip, which rejuvenates contacts. These competing mechanisms translate into
421 nearly zero rate-of-change for the features (Figure 4) at the onset of preslip, which in turn, results
422 in poor regression fits because the monotonically varying data labels are associated with nearly
423 constant features. For instance, in the limited number of stick-slip instabilities plotted in Figure
424 9c,d, notice that the p-wave amplitudes and velocities are relatively constant at ~71 a.u. and 5719
425 m/s respectively, for friction coefficients in the range of 0.575 – 0.59, i.e., during the transition
426 from interseismic healing to preseismic creep.

427 Taken together, our ML model predictions (Figures 5-7) and the physical mechanisms
428 behind the evolution of the ultrasonic parameters (Figure 4, 9) indicate that even a relatively
429 simple regression, incorporating no explicit temporal information, can provide insights into the
430 mechanics of fault stability and earthquake nucleation, at least on the laboratory scale. More
431 specifically, because the ultrasonic (amplitude/velocity) evolution is physically related to healing
432 and preslip-driven interfacial asperity stiffness variations, and shear stiffening of the wallrock,

433 the ML model is able to infer the state of the fault and proximity to failure from these
434 microphysical variations.

435 These results have important implications for faulting and earthquake forecasting on the
436 crustal scale, particularly near shallow subduction zones (*Reasenberg, 1999*), and in regions
437 where elastodynamic foreshocks (*Dodge et al., 1996; Ellsworth and Bulut, 2018*) and slow
438 aseismic creep fronts (*Melbourne and Webb, 2002*) have been detected prior to a mainshock.
439 Preseismic crustal velocity anomalies have been observed prior to a limited number of
440 earthquakes (eg. *Niu et al., 2008; Chiarabba et al., 2020*). Thus, these ML methods could
441 potentially be applied to such regions and, particularly around repeating earthquakes or shallow
442 slow earthquakes, if velocity or attenuation trends are available over multiple seismic cycles.
443

444 **5. Conclusions:**

445 We study the feasibility of predicting the timing and shear stress of laboratory
446 earthquakes using high resolution measurements of transmitted wave amplitude and velocity.
447 Our data provide the first test of using active source seismic data to predict labquakes. We find
448 that supervised machine learning is capable of predicting the timing and shear stress state of
449 labquakes with reasonable accuracy. Moreover, our results indicate that post-seismic increases in
450 the ultrasonic amplitudes and velocities, often associated with fault and frictional healing, may
451 contain memory of the past state of the fault. Importantly, our predictions of fault time-to-failure
452 improve in accuracy prior to failure, indicating that fault preslip, which reduces ultrasonic
453 amplitude and velocity, has significant predictive power in the context of imminent failure.
454 Finally, the physical underpinning of the systematic changes in ultrasonic attributes is grounded
455 in the deformation mechanics and the evolution of stiffness of microscopic load-bearing

456 asperities. Hence, we are able to assign a physical model for the inner workings of predictive
457 'black-box' ML models. Overall, our study demonstrates the utility of ML techniques in the
458 study of fault mechanics at the laboratory scale and serves to motivate future pursuits in the quest
459 to improve earthquake forecasting and hazard preparedness in crustal faults.
460
461

Acknowledgements:

We thank Steven Swavely for technical assistance. This study was supported by European Research Council Advance Grant 835012 (TECTONIC), US Department of Energy grants DE-SC0020512 and DE-EE0008763, and US National Science Foundation grants EAR 1520760, EAR1547441, and EAR1763305 to CM. Data are available from PSU Scholarsphere (https://www.doi.org/10.26207/c4as-9s14) or by contacting the corresponding author.

Figures

Figure 1. Data for one complete experiment showing the shear stress evolution as a function of load point displacement. Unload-reload cycles at ~2 and 3 mm promote shear localization and steady state shear. Training and test data come from 18-20 mm and 20-21 mm of shear displacement, respectively. Insets show: schematic of the experiment setup, zoom of slow and fast labquakes with definitions of time-since-failure (T_{prev}) and time-to-failure (T_{next}), and a sample elastic wave with the wavelet (blue) used to measure p-wave travel time and amplitude.

Figure 2. Data labels (a,b) and feature engineering (c-f) for the ML models. (a) Shear stress over multiple lab seismic cycles showing a complex range of stress drops and recurrence intervals; (b) Time-since- (dashed) and time-to-failure (solid) for the events in Panel a; (c and d) raw (green) and smoothed (black) p-wave amplitude (c) and velocity (d); (e and f) relative changes in p-wave amplitude (e) and velocity (f) normalized against the start of each stick-slip cycle (black) and an offset vector (blue) of the normalized amplitudes.

484 **Figure 3.** Complexity and periodicity of the labquake recurrence interval. (a) Cross-plot of time
485 elapsed since the previous event and time remaining to the next event colored by labquake stress
486 drop. Fiducial lines for 1:2 and 2:1 represent period doubling and 1:1 represents perfectly
487 periodic events. (b) Evolution of peak and minimum shear stress for multiple events in the
488 training and testing datasets, plotted by event number and thus increasing shear displacement.
489 Note that stress drop increases slightly with shear (c) Ratio of time since previous and time to
490 next event plotted by event number. The dashed lines at 0.5 and 2 represent the envelopes for
491 period doubling. A ratio of 1 represents perfectly periodic events.
492
493 **Figure 4.** Evolution of the mechanical and ultrasonic data over multiple laboratory seismic
494 cycles. Temporal evolution of (a) Shear stress (b) slip (black) and slip rate (blue) estimated from
495 the onboard slip sensor. Note that the fault is nearly locked for a big fraction of the seismic cycle
496 (c) Transmissivity and (d) p-wave velocity. Note that both transmitted wave amplitude (c) and
497 velocity show clear precursory changes prior to failure.
498
499 **Figure 5.** Shear stress prediction using ML. (a-c) Prediction using amplitudes as features
500 showing (a) Training (b) Testing dataset and (c) model performance expressed as a cross-plot of
501 experimental shear stress versus ML model results with 1:1 line indicating perfect model
502 accuracy. (d-f) Prediction using velocities as features showing (d) Training (e) testing dataset and
503 (c) model performance. Detailed comparison for a representative seismic cycle showing data
504 (black) and model (blue) trained on amplitude (g) and velocity(h). In both panels the lower plot
505 (grey) shows the RMSE misfit.
506

Figure 6. Time-since-failure prediction. (a-c) Prediction using amplitudes as features showing (a) Training (b) Testing dataset and (c) model performance expressed as a cross-plot of experimental shear stress versus ML model results with 1:1 line indicating perfect model accuracy. (d-f) Prediction using velocities as features showing (d) Training (e) testing dataset and (c) model performance. Detailed comparison for a representative seismic cycle showing data (black) and model (blue) trained on amplitude (g) and velocity(h). In both panels the lower plot (grey) shows the RMSE misfit.

Figure 7. Time-to-failure prediction. (a-c) Prediction using amplitudes as features showing (a) Training (b) Testing dataset and (c) model performance expressed as a cross-plot of experimental shear stress versus ML model results with 1:1 line indicating perfect model accuracy. (d-f) Prediction using velocities as features showing (d) Training (e) testing dataset and (c) model performance Detailed comparison for a representative seismic cycle showing data (black) and model (blue) trained on amplitude (g) and velocity(h). In both panels the lower plot (grey) shows the RMSE misfit.

Figure 8. Naïve model for benchmarking performance. Time to failure for the experimental data (black) and a model (green) based on constant recurrence interval. The low R^2 value shows that the naïve model predicts earthquake failure times poorly.

Figure 9. Phase plane cross-plots of (a) friction and slip velocity expressed as a ratio of the slip velocity to loading rate on a logarithmic scale (b) friction and frictional state expressed on a logarithmic scale (c) friction and p-wave amplitudes (d) p-wave velocity. The interseismic

530 locking (green), preseismic creep (blue) and coseismic slip (orange) phases are annotated in each
531 plot.

532

533 **References**

534 Bakun, W. H., Aagaard, B., Dost, B., Ellsworth, W. L., Hardebeck, J. L., Harris, R. A., ... &
535 Michael, A. J. (2005). Implications for prediction and hazard assessment from the 2004
536 Parkfield earthquake. Nature, 437(7061), 969-974.
537 Bergen, K. J., Johnson, P. A., Maarten, V., & Beroza, G. C. (2019). Machine learning for data-
538 driven discovery in solid Earth geoscience. *Science*, *363*(6433).
539 Bolton, D. C., Shokouhi, P., Rouet-Leduc, B., Hulbert, C., Rivière, J., Marone, C., & Johnson, P.
540 A. (2019). Characterizing acoustic signals and searching for precursors during the laboratory
541 seismic cycle using unsupervised machine learning. *Seismological Research Letters*, *90*(3),
542 1088-1098.
543 Bolton, D. C., Shreedharan, S., Rivière, J., & Marone, C. (2020). Acoustic Energy Release
544 During the Laboratory Seismic Cycle: Insights on Laboratory Earthquake Precursors and
545 Prediction. Journal of Geophysical Research: Solid Earth, 125(8), e2019JB018975.
546 Brace, W. F., Paulding Jr, B. W., & Scholz, C. H. (1966). Dilatancy in the fracture of crystalline
547 rocks. *Journal of Geophysical Research*, *71*(16), 3939-3953.
548 Brenguier, F., Campillo, M., Hadziioannou, C., Shapiro, N. M., Nadeau, R. M., & Larose, E.
549 (2008). Postseismic relaxation along the San Andreas fault at Parkfield from continuous
550 seismological observations. science, 321(5895), 1478-1481.
551 Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system.
552 In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and*
553 *data mining* (pp. 785-794).
554 Chiarabba, C., De Gori, P., Segou, M., & Cattaneo, M. (2020). Seismic velocity precursors to the
555 2016 Mw 6.5 Norcia (Italy) earthquake. *Geology*.
556 Corbi, F., Sandri, L., Bedford, J., Funiciello, F., Brizzi, S., Rosenau, M., & Lallemand, S. (2019).
557 Machine learning can predict the timing and size of analog earthquakes. *Geophysical*
558 *Research Letters*, *46*. https://doi.org/ 10.1029/2018GL081251
559 Corbi, F., Bedford, J., Sandri, L., Funiciello, F., Gualandi, A., & Rosenau, M. (2020). Predicting
560 imminence of analog megathrust earthquakes with Machine Learning: Implications for
561 monitoring subduction zones. *Geophysical Research Letters*, *47*(7), e2019GL086615.
562 Dieterich, J. H. (1972). Time-dependent friction in rocks. *Journal of Geophysical*
563 *Research*, *77*(20), 3690-3697.
564 Dieterich, J. (1994). A constitutive law for rate of earthquake production and its application to
565 earthquake clustering. *Journal of Geophysical Research: Solid Earth*, *99*(B2), 2601-2618.
566 Dodge, D. A., Beroza, G. C., & Ellsworth, W. L. (1996). Detailed observations of California
567 foreshock sequences: Implications for the earthquake initiation process. Journal of
568 Geophysical Research: Solid Earth, 101(B10), 22371-22392.
569 Ellsworth, W. L., & Bulut, F. (2018). Nucleation of the 1999 Izmit earthquake by a triggered
570 cascade of foreshocks. *Nature Geoscience*, *11*(7), 531-535.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

Gu, J. C., Rice, J. R., Ruina, A. L., & Simon, T. T. (1984). Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction. Journal of the Mechanics and Physics of Solids, 32(3), 167-196.

Hedayat, A., Pyrak-Nolte, L. J., & Bobet, A. (2014). Precursors to the shear failure of rock discontinuities. *Geophysical Research Letters*, *41*(15), 5467-5475.

Hulbert, C., Rouet-Leduc, B., Johnson, P. A., Ren, C. X., Rivière, J., Bolton, D. C., & Marone, C. (2019). Similarity of fast and slow earthquakes illuminated by machine learning. *Nature Geoscience*, *12*(1), 69-74.

Hulbert, C., Rouet-Leduc, B., Jolivet, R., & Johnson, P. A. (2020). An exponential build-up in seismic energy suggests a months-long nucleation of slow slip in Cascadia. *Nature communications*, *11*(1), 1-8.

Jones, D. R., Schonlau, M., & Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, *13*(4), 455-492.

Kame, N., Nagata, K., Nakatani, M., & Kusakabe, T. (2014). Feasibility of acoustic monitoring of strength drop precursory to earthquake occurrence. *Earth, Planets and Space*, *66*(1), 41.

Kaproth, B. M., & Marone, C. (2013). Slow earthquakes, preseismic velocity changes, and the origin of slow frictional stick-slip. *Science*, *341*(6151), 1229-1232.

Kaproth, B. M., and C. Marone, Evolution of elastic wave speed during shear-induced damage and healing within laboratory fault zones, J. Geophys. Res. Solid Earth, 119, 10.1002/2014JB011051, 2014.

Karner, S. L., & Marone, C. (1998). The effect of shear load on frictional healing in simulated fault gouge. *Geophysical research letters*, *25*(24), 4561-4564.

Kendall, K., & Tabor, D. (1971). An utrasonic study of the area of contact between stationary and sliding surfaces. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, *323*(1554), 321-340.

Kilgore, B., Beeler, N. M., Lozos, J., & Oglesby, D. (2017). Rock friction under variable normal stress. *Journal of Geophysical Research: Solid Earth*, *122*(9), 7042-7075.

Kong, Q., Trugman, D. T., Ross, Z. E., Bianco, M. J., Meade, B. J., & Gerstoft, P. (2019). Machine learning in seismology: Turning data into insights. *Seismological Research Letters*, *90*(1), 3-14.

Leeman, J. R., Saffer, D. M., Scuderi, M. M., & Marone, C. (2016). Laboratory observations of slow earthquakes and the spectrum of tectonic fault slip modes. *Nature communications*, *7*(1), 1-6.

Lubbers, N., Bolton, D. C., Mohd-Yusof, J., Marone, C., Barros, K., & Johnson, P. A. (2018). Earthquake catalog-based machine learning identification of laboratory fault states and the effects of magnitude of completeness. *Geophysical Research Letters*, *45*(24), 13-269.

Main, I. G., Bell, A. F., Meredith, P. G., Geiger, S., & Touati, S. (2012). The dilatancy–diffusion hypothesis and earthquake predictability. *Geological Society, London, Special Publications*, *367*(1), 215-230.

Malagnini, L., Dreger, D. S., Bürgmann, R., Munafò, I., & Sebastiani, G. (2019). Modulation of seismic attenuation at Parkfield, before and after the 2004 M6 earthquake. *Journal of Geophysical Research: Solid Earth*, *124*(6), 5836-5853.

Marone, C. (1998a). The effect of loading rate on static friction and the rate of fault healing during the earthquake cycle. *Nature*, *391*(6662), 69-72.

Marone, C. (1998b). Laboratory-derived friction laws and their application to seismic faulting. *Annual Review of Earth and Planetary Sciences*, *26*(1), 643-696.

McBeck, J., Aiken, J. M., Ben-Zion, Y., & Renard, F. (2020). Predicting the proximity to macroscopic failure using local strain populations from dynamic in situ X-ray tomography triaxial compression experiments on rocks. *Earth and Planetary Science Letters*, *543*, 116344.

McBrearty, I. W., Delorey, A. A., & Johnson, P. A. (2019). Pairwise association of seismic arrivals with convolutional neural networks. Seismological Research Letters, 90(2A), 503-509.

McLaskey, G. C., Thomas, A. M., Glaser, S. D., & Nadeau, R. M. (2012). Fault healing promotes high-frequency earthquakes in laboratory experiments and on natural faults. *Nature*, *491*(7422), 101-104.

Melbourne, T. I., & Webb, F. H. (2002). Precursory transient slip during the 2001 Mw= 8.4 Peru earthquake sequence from continuous GPS. *Geophysical Research Letters*, *29*(21), 28-1.

Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature communications, 11(1), 1-12.

Nagata, K., Nakatani, M., & Yoshida, S. (2012). A revised rate-and state-dependent friction law obtained by constraining constitutive and evolution laws separately with laboratory data. Journal of Geophysical Research: Solid Earth, 117(B2).

Nagata, K., Kilgore, B., Beeler, N., & Nakatani, M. (2014). High-frequency imaging of elastic contrast and contact area with implications for naturally observed changes in fault properties. *Journal of Geophysical Research: Solid Earth*, *119*(7), 5855-5875.

Niu, F., Silver, P. G., Daley, T. M., Cheng, X., & Majer, E. L. (2008). Preseismic velocity changes observed from active source monitoring at the Parkfield SAFOD drill site. *Nature*, *454*(7201), 204-208.

Olivier, Gerrit, Julien Chaput, and Brian Borchers. "Using supervised machine learning to improve active source signal retrieval." *Seismological Research Letters* 89.3 (2018): 1023-1029.

Passelègue, F. X., Pimienta, L., Faulkner, D., Schubnel, A., Fortin, J., & Guéguen, Y. (2018). Development and Recovery of Stress-Induced Elastic Anisotropy During Cyclic Loading Experiment on Westerly Granite. Geophysical Research Letters, 45(16), 8156-8166.

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. Science Advances, 4(2), e1700578.

Pyrak-Nolte, L. J., Myer, L. R., & Cook, N. G. (1990). Transmission of seismic waves across single natural fractures. Journal of Geophysical Research: Solid Earth, 95(B6), 8617-8638.

Reasenberg, P. A. (1999). Foreshock occurrence before large earthquakes. *Journal of Geophysical Research: Solid Earth*, *104*(B3), 4755-4768.

Ren, C. X., Peltier, A., Ferrazzini, V., Rouet-Leduc, B., Johnson, P. A., & Brenguier, F. (2020). Machine learning reveals the seismic signature of eruptive behavior at piton de la fournaise volcano. *Geophysical Research Letters*, *47*(3), e2019GL085523.

Rivière, J., Lv, Z., Johnson, P. A., & Marone, C. (2018). Evolution of b-value during the seismic cycle: Insights from laboratory experiments on simulated faults. *Earth and Planetary Science Letters*, *482*, 407-413.

Ross, Z. E., Meier, M. A., Hauksson, E., & Heaton, T. H. (2018). Generalized seismic phase detection with deep learning. Bulletin of the Seismological Society of America, 108(5A), 2894-2901.

Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C. J., & Johnson, P. A. (2017). Machine learning predicts laboratory earthquakes. *Geophysical Research Letters*, *44*(18), 9276-9282.

Rouet-Leduc, B., Hulbert, C., Bolton, D. C., Ren, C. X., Riviere, J., Marone, C., ... & Johnson, P. A. (2018). Estimating fault friction from seismic signals in the laboratory. *Geophysical Research Letters*, *45*(3), 1321-1329.

Rouet-Leduc, B., Hulbert, C., & Johnson, P. A. (2019). Continuous chatter of the Cascadia subduction zone revealed by machine learning. *Nature Geoscience*, *12*(1), 75-79.

Ryan, K. L., Rivière, J., & Marone, C. (2018). The role of shear stress in fault healing and frictional aging. Journal of Geophysical Research: Solid Earth, 123(12), 10-479.

Sammonds, P. R., Meredith, P. G., & Main, I. G. (1992). Role of pore fluids in the generation of seismic precursors to shear fracture. *Nature*, *359*(6392), 228-230.

Scuderi, M. M., Marone, C., Tinti, E., Di Stefano, G., & Collettini, C. (2016). Precursory changes in seismic velocity for the spectrum of earthquake failure modes. *Nature geoscience*, *9*(9), 695-700.

Scholz, C. H. (1968). Microfracturing and the inelastic deformation of rock in compression. Journal of Geophysical Research, 73(4), 1417-1432.

Shimazaki, K., & Nakata, T. (1980). Time-predictable recurrence model for large earthquakes. Geophysical Research Letters, 7(4), 279-282.

Shreedharan, S., Rivière, J., Bhattacharya, P., & Marone, C. (2019). Frictional state evolution during normal stress perturbations probed with ultrasonic waves. *Journal of Geophysical Research: Solid Earth*, *124*(6), 5469-5491.

Shreedharan, S., Bolton, D. C., Rivière, J., & Marone, C. (2020). Preseismic fault creep and elastic wave amplitude precursors scale with lab earthquake magnitude for the continuum of tectonic failure modes. *Geophysical Research Letters*, *47*(8), e2020GL086986.

Shreedharan, S., Bolton, D. C., Rivière, J., & Marone, C. (2021). Competition between preslip and deviatoric stress modulates precursors for laboratory earthquakes. *Earth and Planetary Science Letters.*

Stanchits, S. A., Lockner, D. A., & Ponomarev, A. V. (2003). Anisotropic changes in P-wave velocity and attenuation during deformation and fluid infiltration of granite. Bulletin of the Seismological Society of America, 93(4), 1803-1822.

Tan, P. N., M. Steinbach, and V. Kumar (2006). Cluster analysis: basic concepts and algorithms, in Introduction to Data Mining, Vol. 8, Addision-Wesley, Boston, Massachusetts, 487–568.

Tinti, E., Scuderi, M., Scognamiglio, L., Di Stefano, G., Marone, C., & Collettini, C. (2016). On the evolution of elastic properties during laboratory stick-slip experiments spanning the transition from slow slip to dynamic rupture. Journal of Geophysical Research: Solid Earth, 121(12), 8569-8594.

Trugman, D. T., & Ross, Z. E. (2019). Pervasive foreshock activity across southern California. Geophysical Research Letters, 46(15), 8772-8781.

Veedu, D.M., Giorgetti, C., Scuderi, M., Barbot, S., Marone, C., & Collettini, C. (2020). Bifurcations at the Stability Transition of Earthquake Faulting. *Geophysical Research Letters*, *47*(19), e2020GL087985.

Wu, Y., Lin, Y., Zhou, Z., Bolton, D. C., Liu, J., & Johnson, P. (2018). DeepDetect: A cascaded region-based densely connected network for seismic event detection. IEEE Transactions on Geoscience and Remote Sensing, 57(1), 62-75.

Yoon, C. E., O'Reilly, O., Bergen, K. J., & Beroza, G. C. (2015). Earthquake detection through computationally efficient similarity search. Science advances, 1(11), e1501057.

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. Geophysical Journal International, 216(1), 261-273.

Ziv, A. (2006). Does aftershock duration scale with mainshock size?. *Geophysical research letters*, *33*(17).

Figure 1.

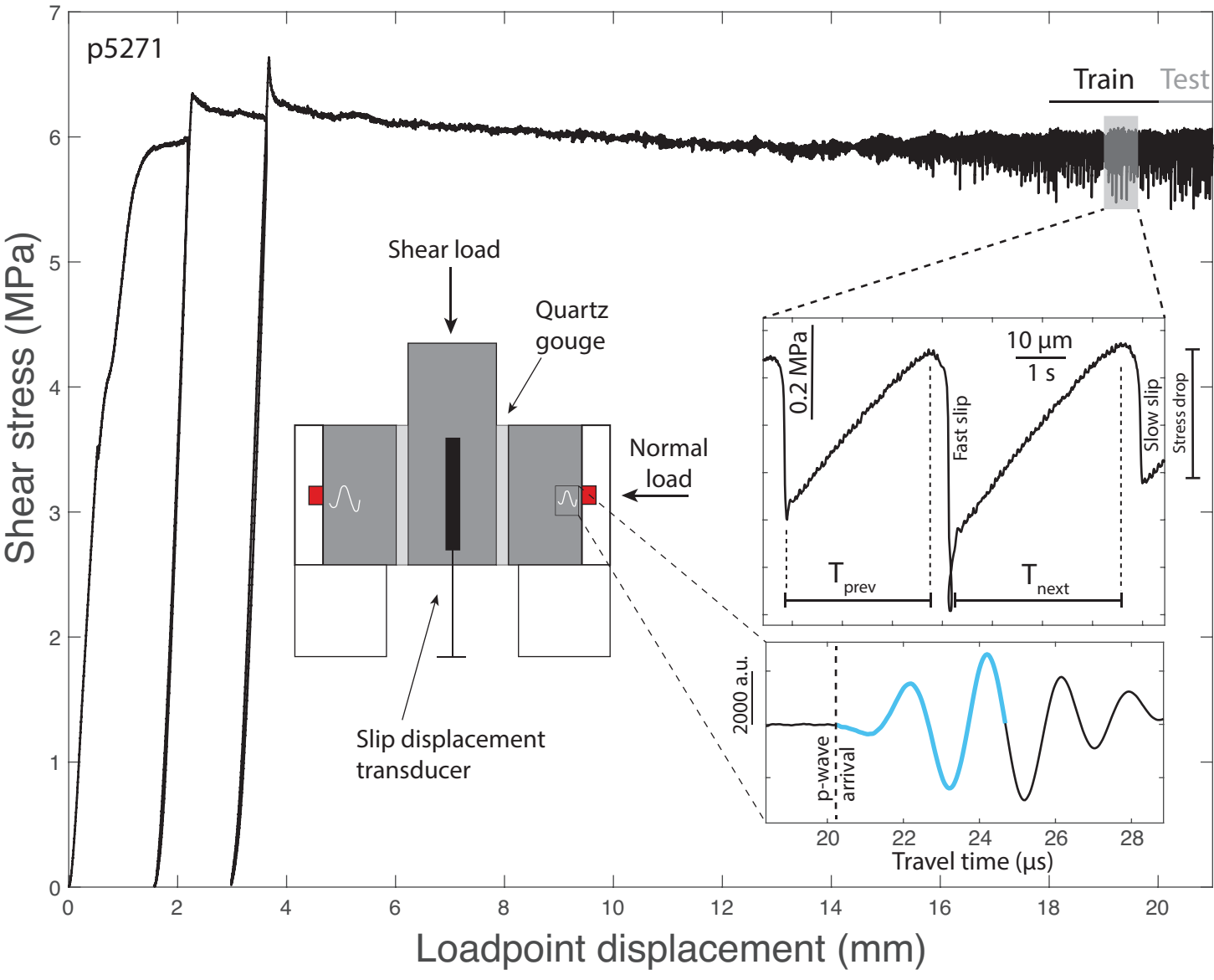

p5271
Train
Test
Shear stress (MPa)
Loadpoint displacement (mm)
Shear load
Quartz gouge
Normal load
Slip displacement transducer
0.2 MPa
10 µm
1 s
Fast slip
Slow slip
Stress drop
T prev
T next
2000 a.u.
p-wave arrival
Travel time (µs)

Figure 2.

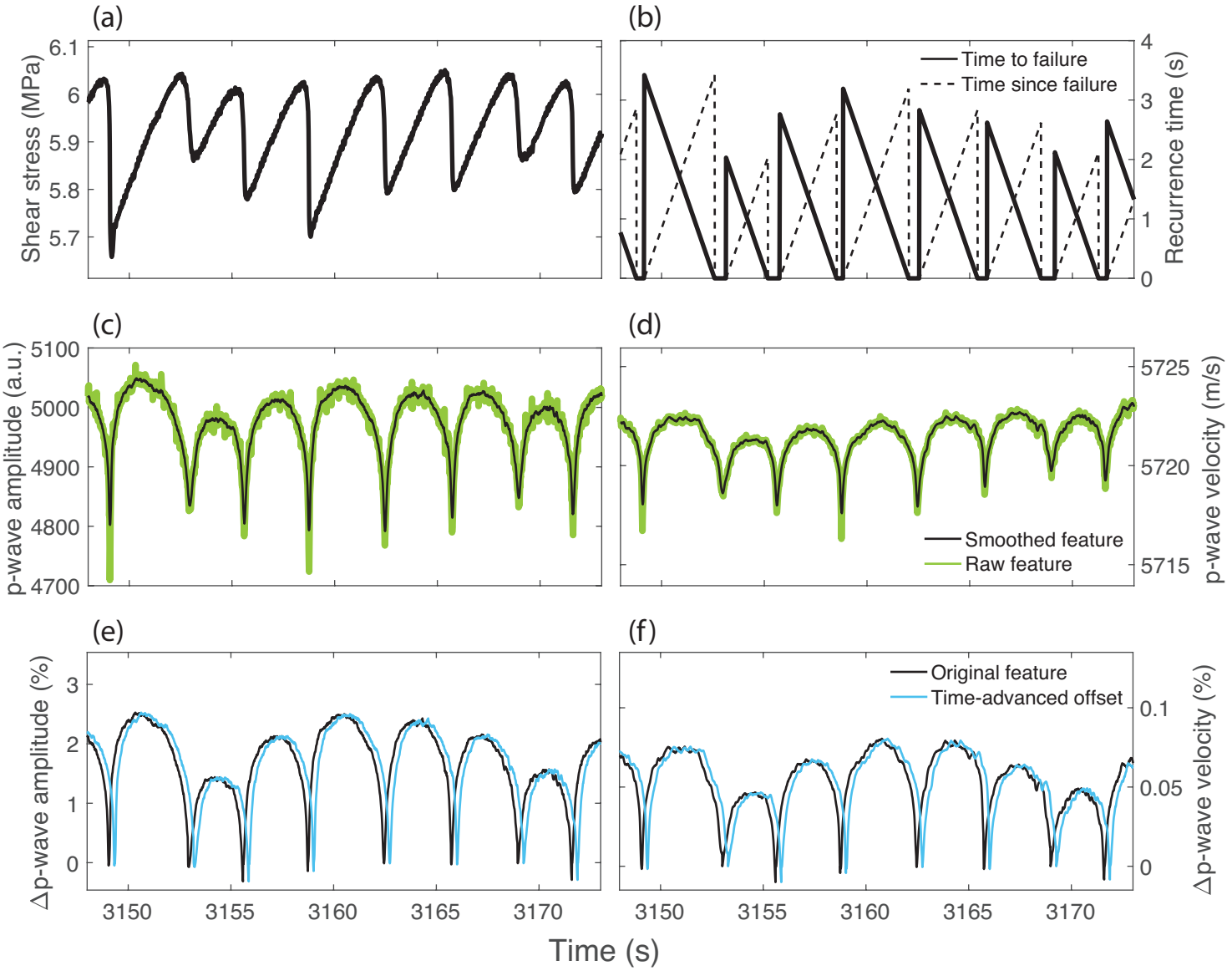

(a)
(b)
Shear stress (MPa)
6.1
6
5.9
5.8
5.7
Time to failure
Time since failure
Recurrence time (s)
4
3
2
1
0
(c)
(d)
p-wave amplitude (a.u.)
5100
5000
4900
4800
4700
Smoothed feature
Raw feature
p-wave velocity (m/s)
5725
5720
5715
(e)
(f)
Δp-wave amplitude (%)
3
2
1
0
Original feature
Time-advanced offset
Δp-wave velocity (%)
0.1
0.05
0
3150
3155
3160
3165
3170
Time (s)

Figure 3.

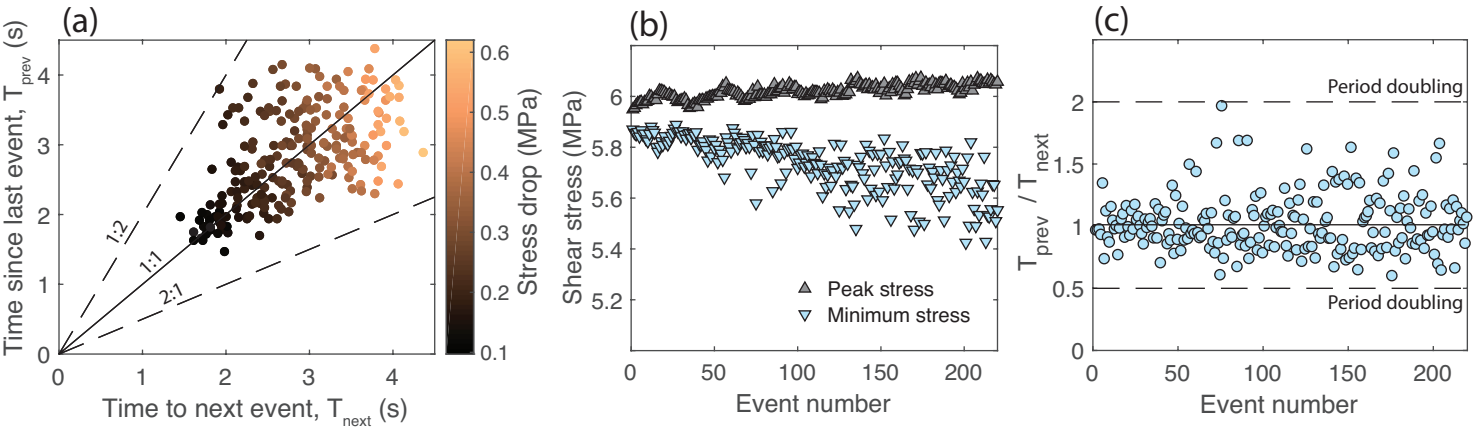

(a)
Time since last event, T_{prev} (s)
Time to next event, T_{next} (s)
1:2
1:1
2:1
Stress drop (MPa)
0.1
0.2
0.3
0.4
0.5
0.6
(b)
Shear stress (MPa)
Event number
Peak stress
Minimum stress
(c)
T_{prev} / T_{next}
Event number
Period doubling
Period doubling

Figure 4.

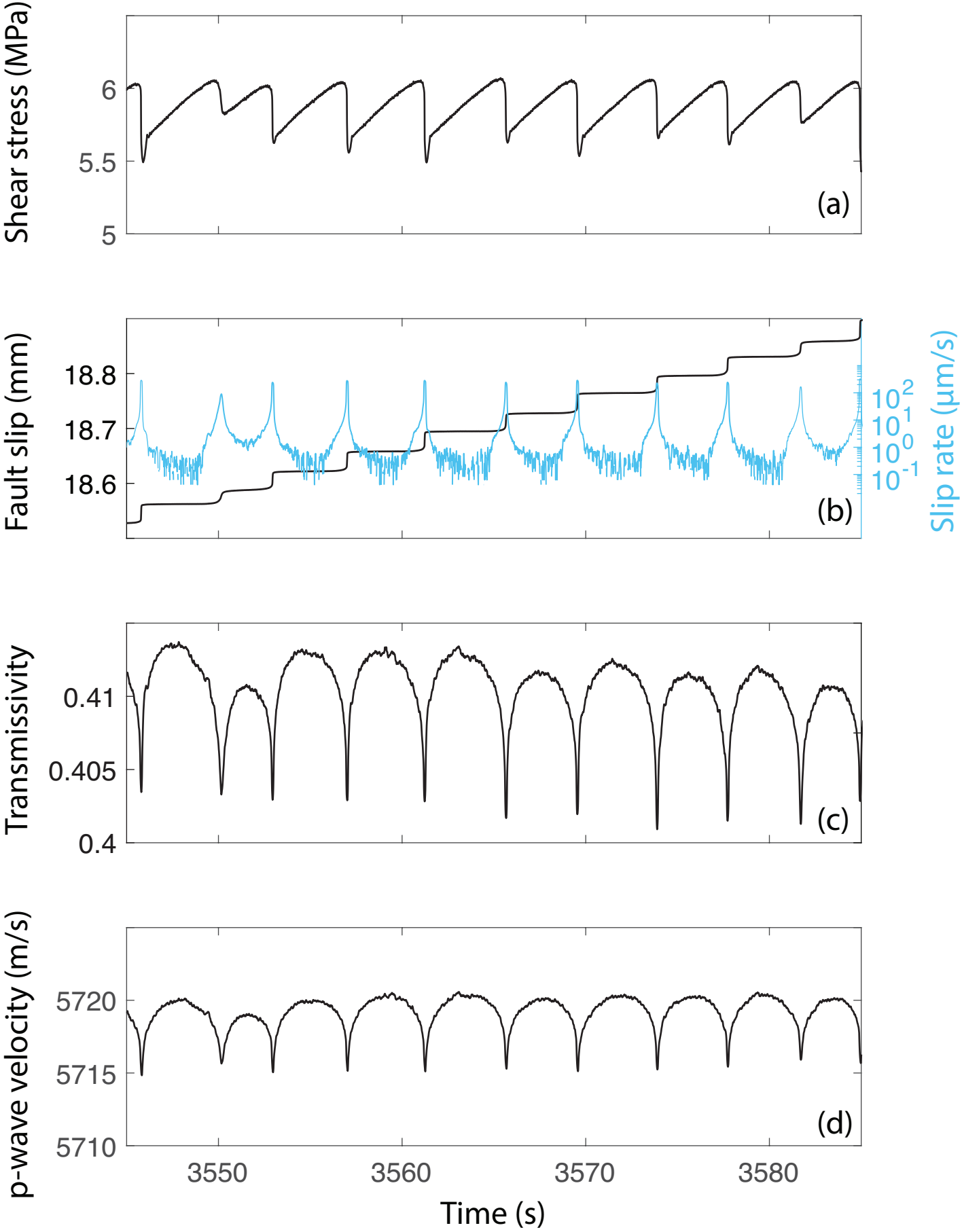

Shear stress (MPa)
6
5.5
5
(a)
Fault slip (mm)
18.8
18.7
18.6
(b)
Slip rate (μm/s)
10^2
10^1
10^0
10^{-1}
Transmissivity
0.41
0.405
0.4
(c)
p-wave velocity (m/s)
5720
5715
5710
(d)
3550
3560
3570
3580
Time (s)

Figure 5.

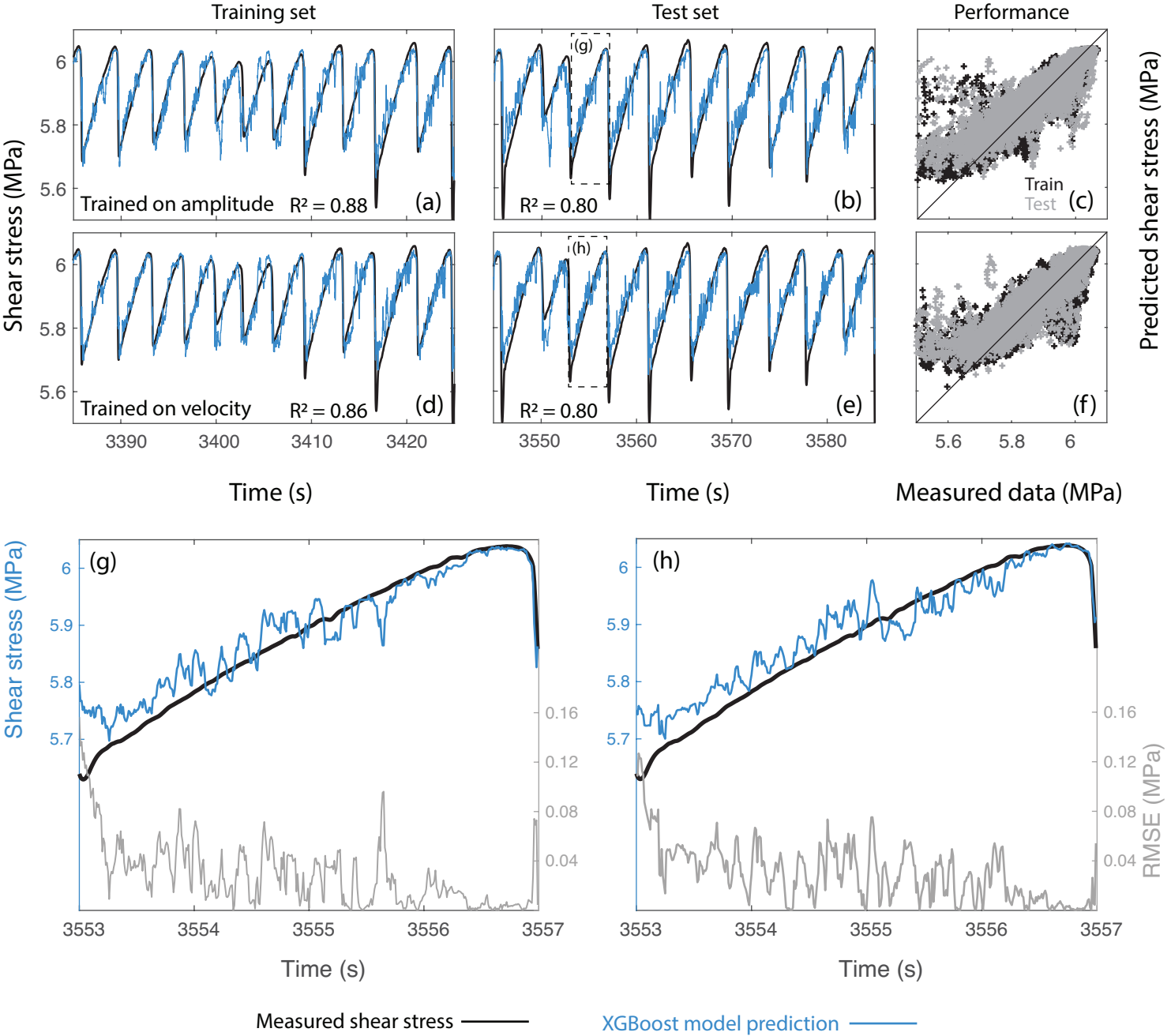

Training set
Test set
Performance
Shear stress (MPa)
Predicted shear stress (MPa)
Trained on amplitude
R² = 0.88
(a)
R² = 0.80
(b)
Train
Test
(c)
Trained on velocity
R² = 0.86
(d)
R² = 0.80
(e)
(f)
(g)
(h)
Time (s)
Time (s)
Measured data (MPa)
(g)
(h)
Shear stress (MPa)
RMSE (MPa)
Time (s)
Time (s)
Measured shear stress
XGBoost model prediction

Figure 6.

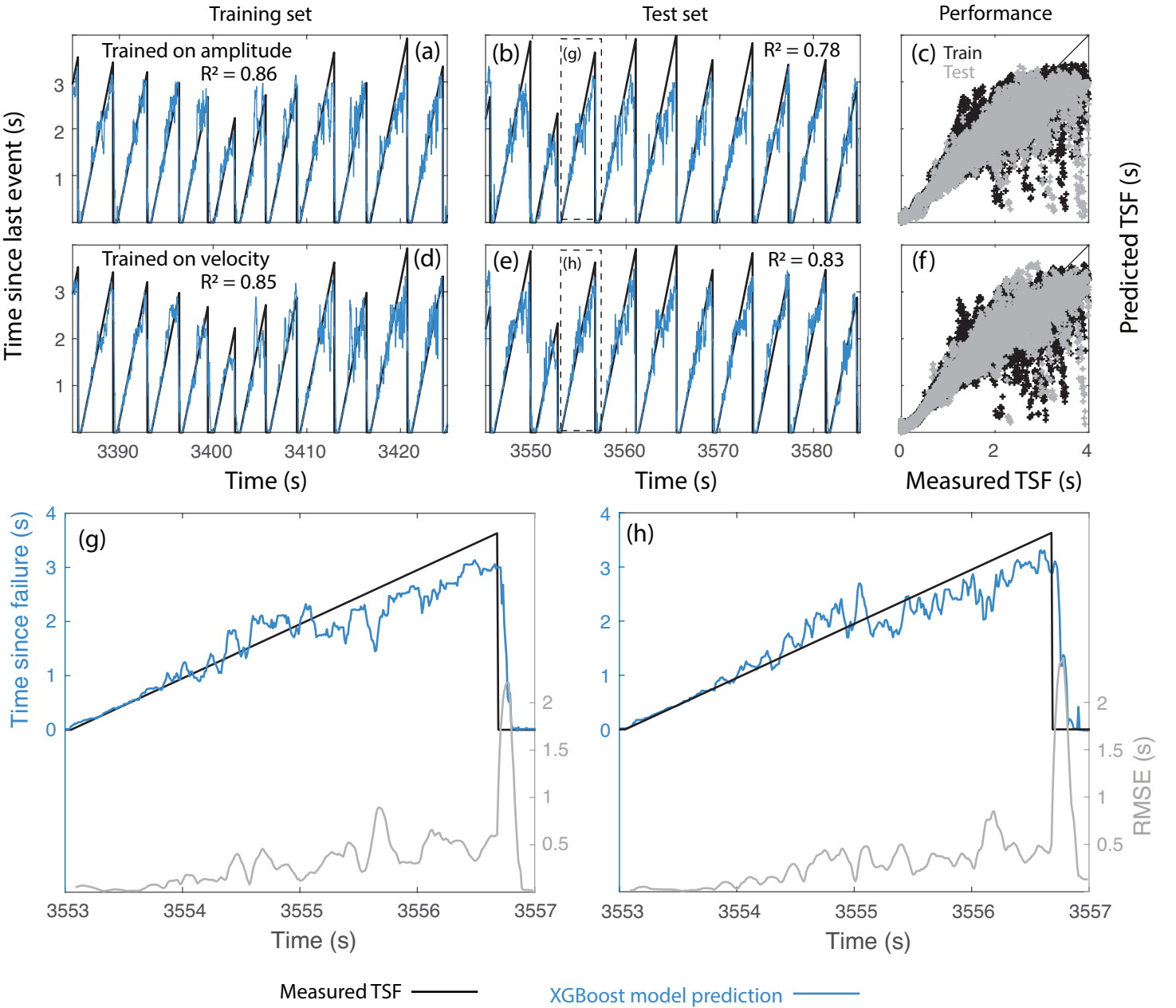

Training set
Test set
Performance
Trained on amplitude
R² = 0.86
(a)
(b)
(g)
R² = 0.78
(c) Train
Test
Trained on velocity
R² = 0.85
(d)
(e)
(h)
R² = 0.83
(f)
Time since last event (s)
Predicted TSF (s)
Time (s)
Time (s)
Measured TSF (s)
(g)
(h)
Time since failure (s)
RMSE (s)
Time (s)
Time (s)
Measured TSF
XGBoost model prediction

Figure 7.

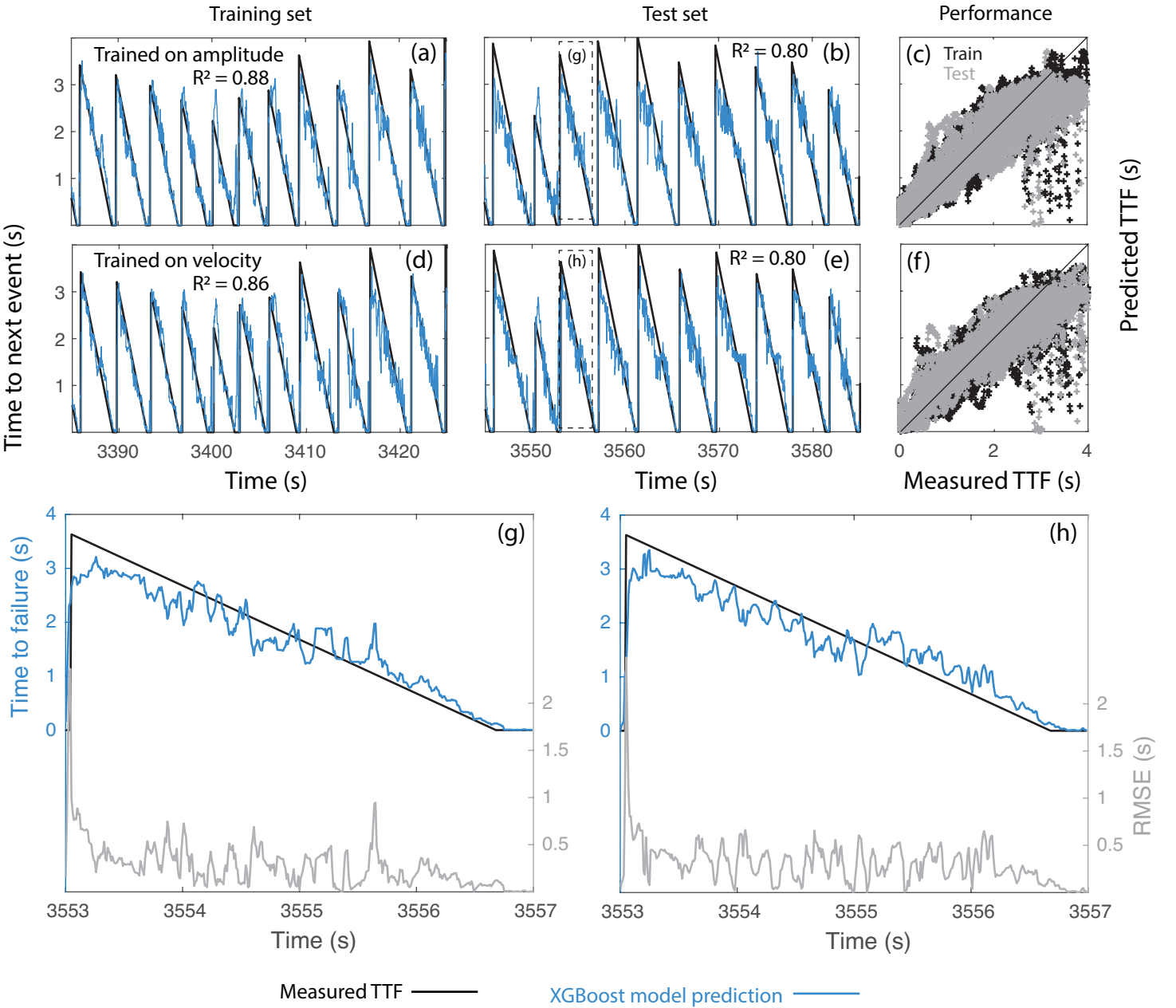

Training set
Test set
Performance
Trained on amplitude
R² = 0.88
(a)
(g)
R² = 0.80
(b)
(c)
Train
Test
Trained on velocity
R² = 0.86
(d)
(h)
R² = 0.80
(e)
(f)
Time to next event (s)
Predicted TTF (s)
Time (s)
Time (s)
Measured TTF (s)
(g)
(h)
Time to failure (s)
RMSE (s)
Time (s)
Time (s)
Measured TTF
XGBoost model prediction

Figure 8.

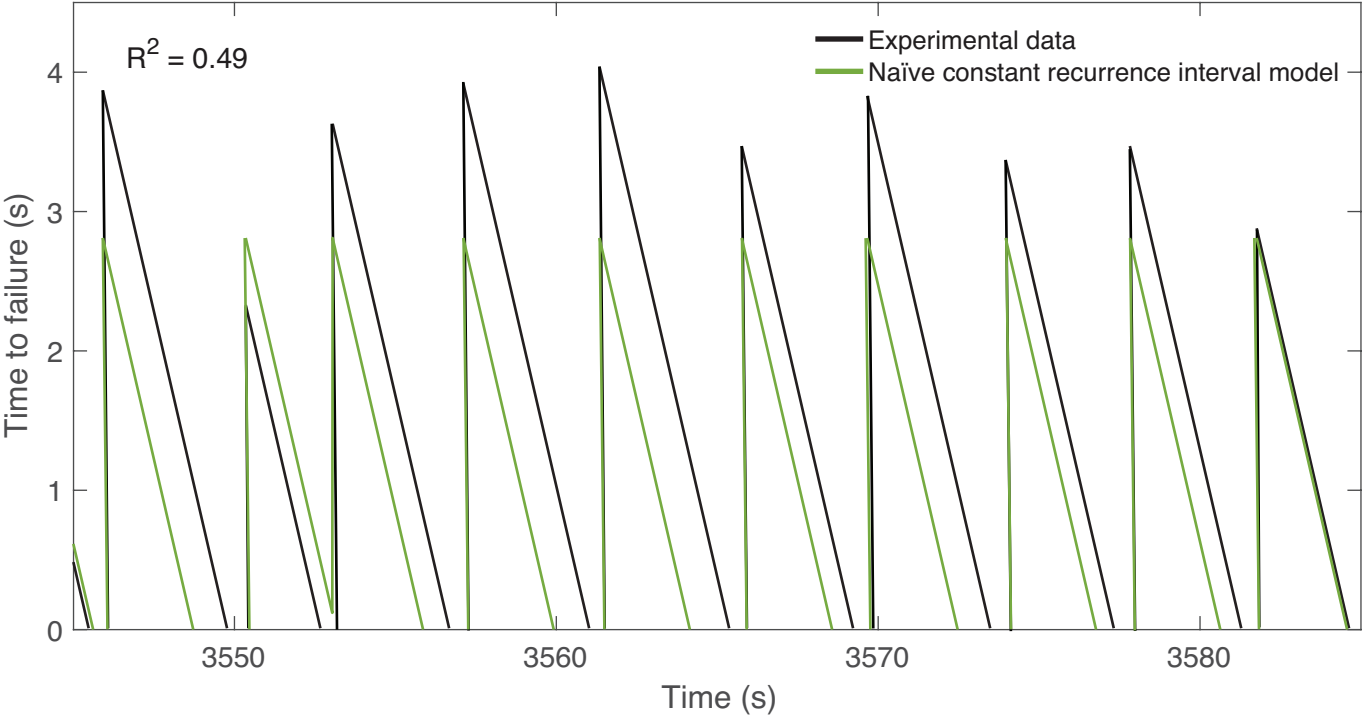
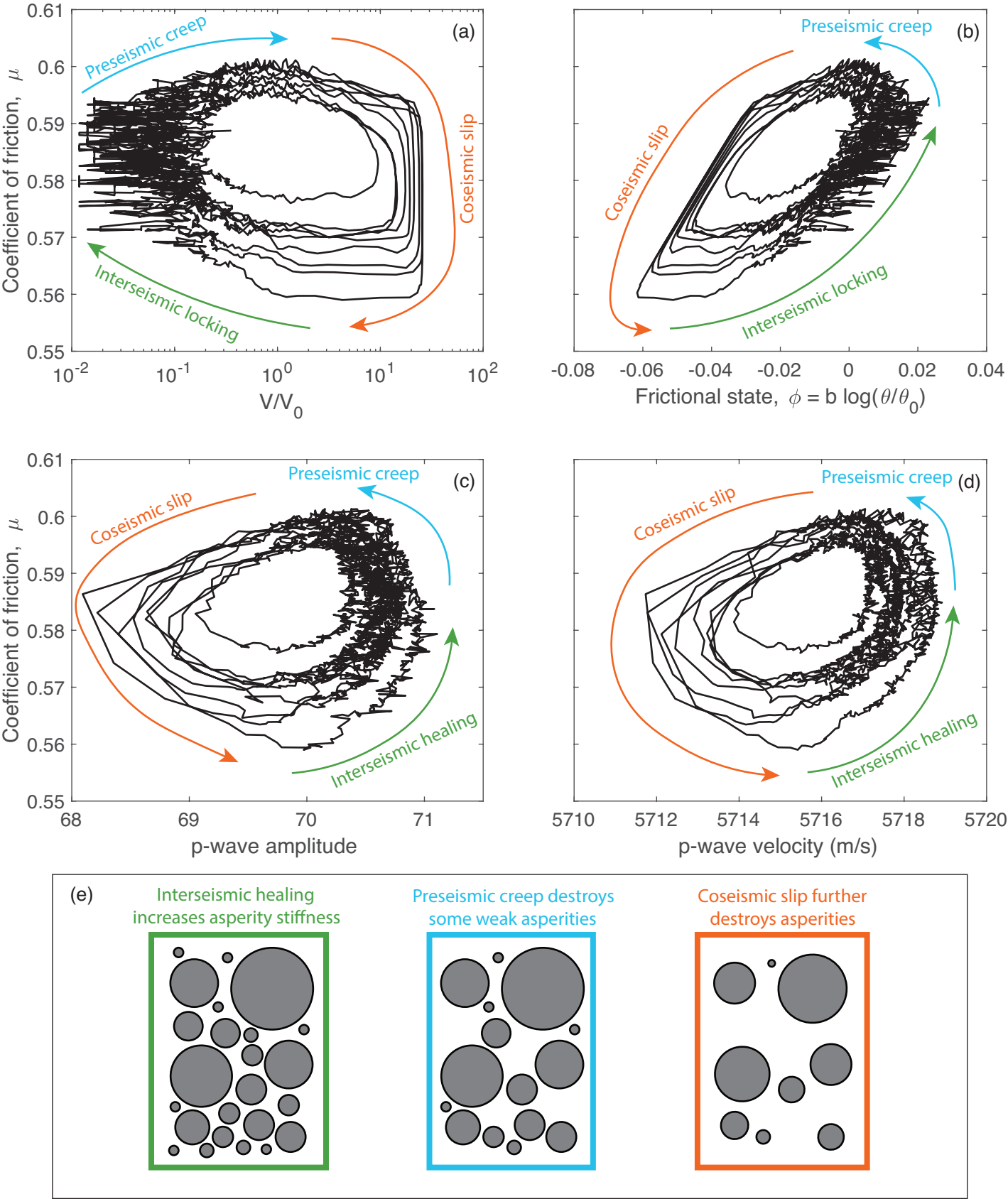

R^2 = 0.49
Experimental data
Naïve constant recurrence interval model
Time to failure (s)
0
1
2
3
4
3550
3560
3570
3580
Time (s)

Figure 9.


(a)
Preseismic creep
Coseismic slip
Interseismic locking
Coefficient of friction, μ
0.61
0.6
0.59
0.58
0.57
0.56
0.55
10^-2
10^-1
10^0
10^1
10^2
V/V0
(b)
Preseismic creep
Coseismic slip
Interseismic locking
-0.08
-0.06
-0.04
-0.02
0
0.02
0.04
Frictional state, ϕ = b log(θ/θ0)
(c)
Preseismic creep
Coseismic slip
Interseismic healing
Coefficient of friction, μ
68
69
70
71
p-wave amplitude
(d)
Preseismic creep
Coseismic slip
Interseismic healing
5710
5712
5714
5716
5718
5720
p-wave velocity (m/s)
(e)
Interseismic healing increases asperity stiffness
Preseismic creep destroys some weak asperities
Coseismic slip further destroys asperities