



Predicting Fracture Network Development in Crystalline Rocks

JESSICA McBECK,¹  J. M. AIKEN,² B. CORDONNIER,¹ Y. BEN-ZION,³ and F. RENARD^{1,4}

Abstract—The geometric properties of fractures influence whether they propagate, arrest, or coalesce with other fractures. Thus, quantifying the relationship between fracture network characteristics may help predict fracture network development, and perhaps precursors to catastrophic failure. To constrain the relationship and predictability of fracture characteristics, we deform eight one centimeter tall rock cores under triaxial compression while acquiring in situ X-ray tomograms. The tomograms reveal precise measurements of the fracture network characteristics above the spatial resolution of 6.5 μm . We develop machine learning models to predict the value of each characteristic using the other characteristics, and excluding the macroscopic stress or strain imposed on the rock. The models predict fracture development more accurately in the experiments performed on granite and monzonite, than the experiments on marble. Fracture network development may be more predictable in these igneous rocks because their microstructure is more mechanically homogeneous than the marble, producing more systematic fracture development that is not strongly impeded by grain contacts and cleavage planes. The varying performance of the models suggest that fracture volume, length, and aperture are the most predictable of the characteristics, while fracture orientation is the least predictable. Orientation does not correlate with length, as suggested by the idea that the orientation evolves with increasing differential stress and thus fracture length. This difference between the observed and expected relationship between orientation and length highlights the influence of mechanical heterogeneities and local stress perturbations on fracture growth as fractures propagate, link, and coalesce.

Keywords: Fracture, machine learning, X-ray tomography, triaxial compression, granite, marble.

1. Introduction

Mechanical weaknesses, such as fractures, control the macroscopic strength of brittle solids (Griffith, 1924). Increasing differential stress may cause some of these preexisting weaknesses to propagate, producing an evolving fracture network that may further weaken the material. Depending on how the fractures propagate, arrest, coalesce, and localize, the evolving fracture network may lead to catastrophic, system-size failure. Constraining the criteria that determine how a fracture network evolves toward macroscopic failure is a fundamental goal in geoscience and engineering. Following linear elastic fracture mechanics, whether a fracture propagates depends on the applied loading and the geometric properties of the fracture that determine the stress intensity factor, such as the fracture length and aperture (Isida, 1971). Whether a fracture arrests its growth depends on the loading as well as the microstructure of the rock, including the mechanical heterogeneity produced by, for example, interlocking minerals with varying elastic moduli or cemented granular aggregates (e.g., Fredrich & Wong, 1986; Howarth, 1987; Moore & Lockner, 1995; Tapponnier & Brace, 1976). Whether a fault system coalesces and develops a distributed or localized spatial distribution depends on the boundary conditions of the system (constant stress vs. displacement), material healing rate, loading rate, and conditions that determine the rheology, such as temperature, pressure, presence of fluids, and rock type (e.g., Ben-Zion, 2008; Lyakhovskiy et al., 2001). Machine learning analyses of data from laboratory experiments on crystalline rocks under triaxial compression indicate that whether a fracture propagates or closes depends on the factors that control the stress intensity factor (fracture volume, length, aperture, orientation), as well as the distance to its nearest

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00024-021-02908-7>.

¹ Department of Geosciences, Njord Centre, University of Oslo, Oslo, Norway. E-mail: j.a.mcbeck@geo.uio.no

² Department of Physics, Center for Computing in Science Education, University of Oslo, Oslo, Norway.

³ Department of Earth Sciences and Southern California Earthquake Center, University of Southern California, Los Angeles, CA, USA.

⁴ University Grenoble Alpes, University Savoie Mont Blanc, CNRS, IRD, IFTTAR, ISTerre, 38000 Grenoble, France.

neighbor (McBeck et al., 2019). These factors, along with the shape anisotropy of the fracture (one minus aperture/length), also control the timing of catastrophic failure (McBeck et al., 2020).

The ability of these fracture network characteristics to predict the likelihood of individual fracture growth (McBeck et al., 2019) and the timing of catastrophic failure (McBeck et al., 2020) during triaxial compression suggest that examining the relationship between fracture network characteristics could be a valuable effort in our attempts to predict the timing of earthquakes. For example, we may not be able to measure the shape anisotropy of a fracture or fault in the field, but if we can derive relationships between the length and anisotropy using experimental data, we could better constrain this geometric property, which is important for predicting catastrophic failure. Moreover, predicting and quantifying the relationships between fracture characteristics may help to evaluate the applicability of our idealized conceptualizations of fracture growth that are grounded in linear elastic fracture mechanics.

Here, we focus on using the relationships between several fracture characteristics to identify the pre-existing conceptualizations of fracture growth that do not agree with the experimental observations. In particular, we use machine learning techniques to identify the fracture characteristics that help predict the quantities of other fracture characteristics in triaxial compression deformation experiments on marble, granite, and monzonite under the confining stress conditions of the upper crust. This analysis reveals the varying predictability of fracture characteristics in the marble and igneous (monzonite and granite) rocks. The fracture networks in the igneous rocks are more predictable, and thus systematic, than the networks in the marble rocks. This result suggests that fault development that occurs in crust dominated by igneous rocks may be more predictable than fault development in crust dominated by marble. The analysis indicates that the fracture characteristics that are the most predictable, with the highest agreement between the predicted and observed characteristic, include the fracture volume, length and aperture. The characteristic that is the least predictable of the characteristics is the fracture orientation relative to the maximum compression direction. Although

several criteria use the orientation of fractures and faults to assess the likelihood of failure and slip (e.g., Coulomb, 1776), the results prompt reexamination of the idea that the orientation of fractures evolves systematically with increasing differential stress from parallel to the maximum compression direction, to oblique to it (e.g., Cartwright-Taylor et al., 2020; McBeck et al., 2019; Renard et al., 2018, 2019).

2. Background

Laboratory triaxial compression experiments on crustal rocks have provided unparalleled access to the evolving characteristics of fracture networks deformed under upper crustal stress conditions (e.g., Paterson & Wong, 2005). For example, in experiments on granite with in situ acquisition of X-ray tomograms during deformation (e.g., Cartwright-Taylor et al., 2020; McBeck et al., 2020), the fracture networks develop from small, isolated fractures to system-spanning, connected networks (Fig. 1). In these experiments, increasing differential stress, σ_D , localizes fracture development first toward the top portion of the core (Fig. 1). New fractures nucleate and preexisting fractures propagate downward and towards the center of the core with increasing differential stress. Preceding macroscopic failure (rightmost core in Fig. 1), the connected fracture network appears to trend at 30° from the maximum compression direction, σ_1 . The fractures appear to grow longer and thicker with increasing σ_D . We

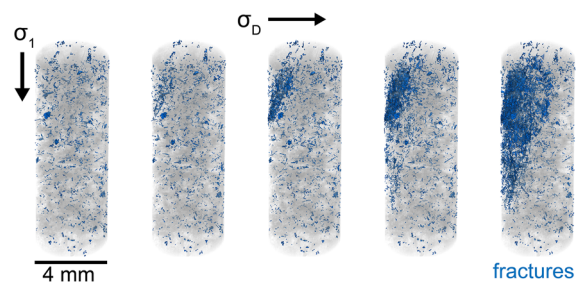


Figure 1
Example of fracture network development in an X-ray synchrotron triaxial compression experiment on granite (experiment granite #4). Fractures are shown in blue and the host rock is shown in gray. The applied differential stress, σ_D , increases from left to right. The rightmost image shows the core immediately preceding macroscopic failure

observed similar behavior in monzonite, another crystalline rock (Renard et al., 2018). Identifying the individual fractures within these networks enables quantifying these observations, and testing if the fracture orientations, lengths and apertures indeed develop with these ostensible evolutions. This approach thereby allows assessing the applicability of existing conceptualizations of fracture network development (Fig. 2).

The central aim of the present study is to use data acquired from in situ X-ray experiments to assess the correctness of preexisting ideas about fracture network development. Here, we describe the concepts that we will test with the evolving fracture networks revealed in X-ray tomography. Abundant experimental evidence indicates that under upper crustal stress and temperature conditions, new fractures nucleate and preexisting fractures grow with increasing differential stress (Tapponier & Brace, 1976; Ayling et al., 1995; Moore & Lockner, 1995; Popp et al., 2001; Stanchits et al., 2006; Bonamy & Bouchaud, 2011; Renard et al., 2018, 2019; Kandula

et al., 2019; Cartwright-Taylor et al., 2020). Fracture nucleation and propagation is expected to increase the total volume of fractures within some subvolumes and decrease the distance between fractures (Fig. 2a). While experiments have documented increasing fracture volume with differential stress in rocks (e.g., Tapponier & Brace, 1976; Paterson & Wong, 2005; Renard et al., 2018) and the importance of the initial spacing between fractures (Vasseur et al., 2017), to our knowledge none have yet reported the expected decreasing distance between fractures.

Experimental observations support the idea from linear elastic fracture mechanics that as fractures grow under approximately constant stress loading, their stress intensity factors increase, producing a positive feedback loop that propels increasingly faster and unstable fracture growth (Jaeger et al., 1979). Tracking the length of fractures throughout X-ray tomography experiments indicate that the fracture length increases with differential stress in crystalline rocks (e.g., Cartwright-Taylor et al., 2020; Renard et al., 2018). Similarly, observing fracture development on the surface of rocks and rock-analogs reveals the variety of fracture network geometries that occur when wing-cracks grow from the tips of preexisting fractures, and multiple fractures coalesce (e.g., Ashby & Hallam, 1986; Ashby & Sammis, 1990; Moore & Lockner, 1995; Morgan et al., 2013; Park & Bobet, 2009; Wong et al., 2001). The propagation and the coalescence of fractures increase the volume, length, and perhaps anisotropy of fractures, and may decrease the average aperture (Fig. 2b). As larger fractures are likely associated with greater localization, the fracture volume, length, and aperture are closely linked to deformation localization. Localization may be a key characteristic of the preparation process leading to system-sized failure in the laboratory and large earthquakes (Ben-Zion & Zaliapin, 2020; Kato & Ben-Zion, 2021; Lockner et al., 1991; Renard et al., 2019; Stanchits et al., 2006).

Experimental observations support the idea that heterogeneities can produce local dilatant zones in rocks undergoing compressive macroscopic loading (e.g., Moore & Lockner, 1995). Because such local dilatant zones can develop, and rocks have lower tensile strength than shear strength, we expect that the first fractures that develop under triaxial compression

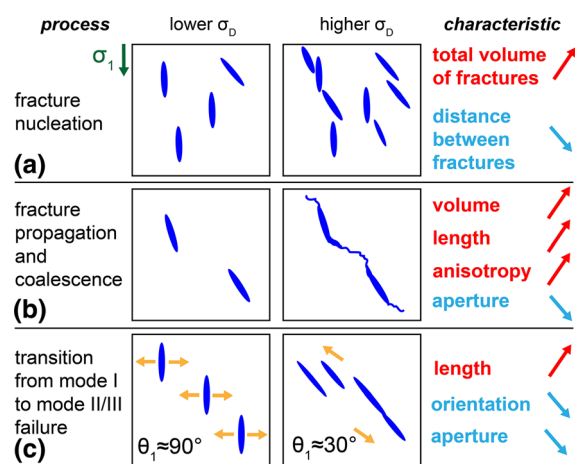


Figure 2

Expected development of fracture network characteristics in triaxial compression experiments. Based on the processes listed to the left of the figure, we expect the fracture network characteristics to increase (red), or decrease (blue) with increasing differential stress, σ_D . Fracture nucleation (a) may increase the total fracture volume and decrease the distance between fractures. Propagation and coalescence (b) may increase the volume, length and anisotropy, and decrease the aperture. A transition in failure mode (c) may decrease the aperture and orientation of the minor axis relative to the maximum compression direction, and increase the length. The three kinds of fracture development (a–c) may occur simultaneously, but here we show them separately for clarity

should be mode I, tensile fractures that open perpendicular to the maximum compression direction and trend parallel to σ_1 (e.g., Tapponnier & Brace, 1976). Then with increasing differential stress, fractures may propagate from the tips of the preexisting fractures, producing fractures that trend oblique to σ_1 , such as wing cracks. Observations from X-ray tomography experiments support the idea that the preferred orientation of fractures evolves with increasing differential stress towards failure (Cartwright-Taylor et al., 2020; McBeck et al., 2019; Renard et al., 2018, 2019). Tracking the orientation of the fractures identified in a tomogram at a given differential stress indicates that the mean orientation of the fractures evolves from parallel to more oblique to σ_1 with increasing differential stress (Renard et al., 2018).

However, some experimental observations indicate that the apparent orientation oblique to σ_1 arises from aligned arrays of mode I fractures that trend parallel to σ_1 (e.g., Horii & Nemat-Nasser, 1985; Nemat-Nasser & Horii, 1982; Peng & Johnson, 1972). Early experiments and Griffith theory indicate that inclined cracks will propagate out of the plane of the crack and then parallel to σ_1 because the maximum tensile stress does not occur exactly at the crack tip, but on the fracture boundary near the tip (Brace & Bombolakis, 1963; Hoek & Bieniawski, 1965). Thus, an array of tensile fractures oriented parallel to σ_1 could grow inclined wing-cracks that will then propagate out-of-plane and perhaps ultimately parallel to σ_1 . If the array of fractures become linked, the resulting overall orientation may become oblique to σ_1 . Therefore, the overall transition from mode I to mode II or III fracturing is expected to shift the orientation from parallel to σ_1 to oblique to it, or decrease it following our reference frame of measuring the orientation (Fig. 2c). This transition may decrease the aperture as fractures that predominantly accommodate shear may have thinner apertures than fractures that predominantly host tension.

Based on these expected relationships between differential stress and the fracture network characteristics, we hypothesize that we may accurately predict some characteristics using others, without knowledge of the macroscopic differential stress nor macroscopic axial strain imposed on the rock core.

We attempt to do this prediction without using the differential stress or axial strain because these properties are often inaccessible in the field or include wide error ranges that may be larger than typical precursory signals (Amoruso & Crescentini, 2010). By comparing the ability of the models to predict each characteristic, we may constrain the extent to which each of the idealized processes of fracture development occurs during triaxial compression of crustal rocks under the stress conditions of the upper crust.

3. Methods

3.1. Rock Samples

We examine fracture development in eight experiments using three rock types. We perform two experiments on Carrara marble, three on monzonite, and three on granite. The three rock types have crystalline microstructures with interlocking mineral grains and low porosity (< 1%). Two of them are igneous rocks (granite, monzonite), and one is a metamorphosed sedimentary rock (Carrara marble). Monzonite and granite have similar mineralogy and ranges of material properties. These igneous rocks only differ in their mean grain size: granite has a mean grain size of 100–200 μm , while monzonite has a mean grain size of 300–400 μm (e.g., Aben et al., 2016). Carrara marble has low initial porosity (0.2%) and a grain size of 100–200 μm (e.g., Malaga-Starzec et al., 2002; Rutter, 1972). These rocks were cored from a block into 10 mm tall and 5 mm wide cylinders.

3.2. X-ray Synchrotron Imaging During Triaxial Deformation

We perform eight experiments at the X-ray synchrotron tomography beamline ID19 at the European Synchrotron and Radiation Facility (ESRF) using the Hades deformation apparatus (Renard et al., 2016). The time series of X-ray tomograms from each experiment are publicly available (Renard, 2018, 2019, 2021). Figure 3 shows the experimental conditions. Previous studies describe the

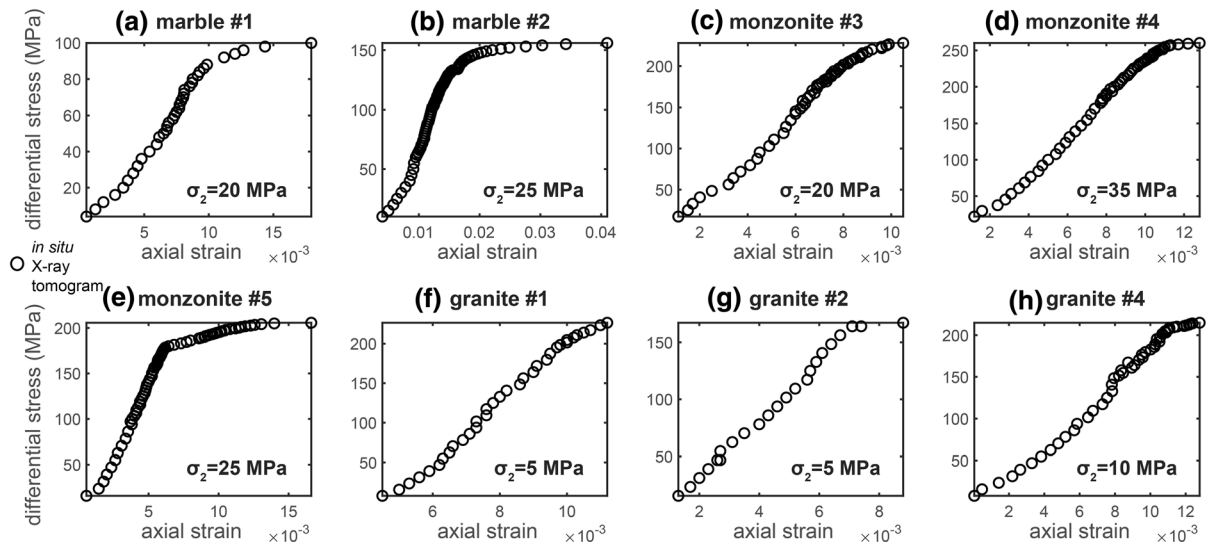


Figure 3

Differential stress and axial strain relationships for eight experiments: **a** marble #1, **b** marble #2, **c** monzonite #3, **d** monzonite #4, **e** monzonite #5, **f** granite #1, **g** granite #2, and **h** granite #4. Each black circle shows the stress and strain conditions when an X-ray tomogram was acquired. The applied confining stress, σ_2 , of each experiment is shown in the corresponding plot. All of the experiments were performed without pore fluid, except monzonite #4, which included 5 MPa of pore fluid pressure

experimental protocol (Kandula et al., 2019; Renard et al., 2018), and we summarize it here. In each experiment, we apply a constant confining stress using oil applied to the jacket surrounding the sample. Then, we increase the axial stress in steps until the rock fails macroscopically. We reduce the size of the stress step after yielding as the rock approaches macroscopic failure in order to attain higher temporal resolution closer to failure. We apply the axial stress increase at rate between 5 MPa per five minutes far from failure, to 1 MPa per five minutes closer to failure. After each axial stress step, we acquire a 3D X-ray tomogram when the rock core is under load inside the deformation apparatus. Each tomogram is acquired in less than 2 min with a spatial sampling, and thus voxel side length of 6.5 μm , and a spatial resolution around 20 μm . The imposed confining stresses range from 5 to 35 MPa for each experiment (Fig. 3). We perform the experiments at ambient temperature (23–25 $^{\circ}\text{C}$). Each experiment was performed without fluid pressure, except for monzonite #4, which included 5 MPa of deionized water as pore fluid pressure with drained conditions at the sample boundaries. However, due to the generally low permeability of the sample, the

fluid may have been drained or undrained inside the sample during the experiment. Macroscopic failure occurs in a sudden stress drop, typically within < 0.5 MPa of increased axial stress after the acquisition of the final scan.

3.3. Fracture Network Feature Extraction

We chose the features using concepts from linear elastic fracture mechanics (Table S1). The X-ray tomograms provide 3D distributions of local X-ray adsorption, which is proportional to the local density. We identify fractures from these adsorption fields by segmenting the voxels dominated by air (i.e., pores and fractures) from those dominated by solid using the procedure described in Renard et al. (2018). We first apply a non-local-mean filter (Buades et al., 2005) to reduce noise in the tomography data. Then, we separate the fractures from the solid by calculating a global threshold using the gray scale values of the tomogram (e.g., Renard et al., 2018). The histograms of the gray scale values contain two peaks that correspond to the solid and air in the tomogram, so we may identify the threshold from the local minimum between the peaks. To identify this

local minimum, we fit two overlapping Gaussian distributions to the two peaks that correspond to the solid- and air-dominated voxels. We then find the summation of these curves and calculate the second derivative of this curve. The minimum absolute value of the second derivative of this curve indicates the inflection point between the two populations of solid- and air-dominated voxels. This method of identifying fractures may misclassify some voxels at the tail ends of the distributions of solid- and air-dominated voxels (e.g., Fousseis et al., 2014). Although machine learning methods and local adaptive thresholding methods can provide higher accuracy than global thresholding methods (e.g., Fousseis et al., 2014), at low levels of noise, global thresholding methods can perform with similar accuracy as the other methods (Andrew, 2018; Wang et al., 2011). Without laboratory measurements of the porosity of the rock cores at each differential stress, it is difficult to assess the accuracy of the calculated segmentation of the fractures and pores. Ongoing work is assessing the accuracy of this global thresholding method and other techniques, such as machine learning.

After we segment the tomograms into a binary field of solid and air, we identify unique pores and fractures by finding connected groups of voxels labelled as air. We find groups of voxels with the most conservative connectivity in 3D (26-fold), ensuring the highest level of connectivity. Next, we apply a noise threshold to remove identified connected voxel clusters below a certain threshold (3000 voxels) in order to remove noise from the data. After we calculate the features of each fracture, we split the volumes into subvolumes with a side length of 200 voxels, and thus a volume of 200^3 voxels, from which we calculate statistics of the fracture characteristic populations (i.e., features). Previous machine learning analyses indicate that using this noise threshold and subvolume size produces the most accurate model predictions of the tested combinations of parameters in these X-ray tomography experiments (McBeck et al., 2020). We report the minimum, 25th percentile, 50th percentile, 75th percentile and maximum value of each population of feature. We chose to predict the statistics of populations of features, rather than the characteristic for individual fractures, because we aim to test whether accurate predictions

require classifying the complete fracture network population, or only require measuring the extreme values, such as the longest or most volumetric fractures.

Table S1 lists all of the 41 features used in the models. In addition to the features derived from the fracture network, we also include a random number as a feature in order to compare the importance of the other features to this meaningless value. We selected each feature derived from the fracture network following formulations from linear elastic fracture mechanics, and preexisting knowledge about fracture development (i.e., Fig. 2). We calculate the features using the number of voxels of each identified fracture (fracture volume), center of mass or centroid of the fracture (distance between fractures), or the eigenvalues and eigenvectors of the covariance matrix of the connected voxels of each fracture (orientation relative to the σ_1 direction, fracture length, fracture aperture, shape anisotropy) (Fig. S1). In particular, the distance between fractures is the distance between the centroid of a fracture and all the other fracture centroids. To calculate the orientation of the fractures relative to the σ_1 direction, we find the orientation of the smallest, θ_1 , and largest, θ_3 , eigenvectors of the fracture relative to the σ_1 direction. We consider both eigenvectors to account for fractures that are penny-shaped (in which the intermediate axis has a similar length to the maximum axis) or cigar-shaped (in which the intermediate axis is closer in length to the minimum axis). When the fractures are penny-shaped, the smallest eigenvector provides a more reliable indicator of orientation than the largest eigenvector. In contrast, when the fractures are cigar-shaped, the orientation of the largest eigenvector provides a more reliable indicator. The fracture length and aperture are approximated by the maximum and minimum eigenvalues, respectively. The shape anisotropy of a fracture is one minus the ratio aperture/length.

3.4. Machine Learning Model Development

We develop XGBoost regression models to predict the fracture characteristics (Chen & Guestrin, 2016). These gradient boosted models depend on a network of decision trees to predict an outcome.

Unlike classification machine learning models which predict discrete values, regression models can predict a continuum of values.

Boosting is an iterative technique of solving statistical models that first fits a model to data, computes the residual errors of that model, fits a new model to the residuals of the first model, then creates a new model that combines the first and second model (Friedman, 2001). This process may be iterated until the total error is minimized. The following equations formalize this process:

$$\hat{y} = F_0 + \sum_{m=1}^M \rho_m X F_m(X) \quad (1)$$

$$F_m(X) = F_{m-1}(X) + \rho_m h_m(X) \quad (2)$$

where \hat{y} is the outcome variable, F_0 is the initialized model (e.g., minimizing a loss function such as mean squared error), $F_m(X)$ is the function being fit to data, X , (e.g., a decision tree) and thus the linear combination of the model from the previous iteration and the model trained on the residuals of this iteration, ρ_m is a weighting term that determines the effect of this iteration on the overall model, $h_m(X)$ is the function being fit to the residuals of iteration $m - 1$, m is the boosting iteration, and M is the total number of boosting iterations.

This process allows the boosted model to learn different areas of the overall space of available data by correcting itself in the places where it was wrong in previous iterations (Friedman, 2001). This process enables boosted models to better fit data than models solved using maximum likelihood methods (Bühlmann & Yu, 2003). Here, we use an implementation of gradient boosting known as Extreme Gradient Boosting or XGBoost (Chen & Guestrin, 2016). XGBoost implements a number of additional optimizations that reduces overfitting, and so has won many machine learning competitions. XGBoost and other gradient boosted tree methods have been used in a number of recent machine learning analyses in geoscience (e.g., Hulbert et al., 2019; Shreedharan et al., 2021).

We develop models using data from the eight experiments, and models from the two dominant rock types: the marble and igneous rocks. We estimate the influence of the features on the model predictions

using a widely used metric in the machine learning community: Shapely Additive Explanations (SHAP) (Lundberg & Lee, 2017; Pedregosa et al., 2011). SHAP values may be calculated for the feature of each sample (i.e., unique data point provided to the model). Here, we report and compare the mean absolute value of the SHAP (mean |SHAP|) of each feature. We thus focus on the net influence of that feature on the model prediction, and not on the influence of that feature on particular samples. We perform a grid search over the hyperparameter-space to find the appropriate set of hyperparameters for each model (Lundberg & Lee, 2017).

We separate the training (80%) and testing (remaining 20%) datasets with no overlap between these two sets. Each sample provided to the models are distinct in time (stress step) and space. We split the samples into training and testing datasets based on time, such that a particular time (differential stress) occurs in either the training or testing dataset. This method of splitting by time reduces the potential for autocorrelation in the results. We randomize the times that occur in the training and testing datasets such that times earlier and later in the experiment can occur in either dataset.

For each model that predicts a given fracture characteristic, we develop ten unique models that differ only in the separation of the training and testing datasets. We develop these models in order to reduce the influence of random variations in the training and testing dataset on the calculated model performance.

To assess the influence of autocorrelation with this method of splitting the data, we also develop models using testing datasets that are continuous in time, and occur at the end of the experiment, close to macroscopic failure (Fig. S2). This method of splitting is expected to lower the model performance because characteristics of fracture networks evolve toward failure. Thus, the fracture characteristics near failure may be outside the range of those characteristics earlier in the experiment. Despite this natural evolution, when we split the training and testing datasets such that the testing dataset encompasses a continuous interval of time with 20% of the data at the end of the experiment, the R^2 scores of the models do not generally decrease in the models that predict the characteristics using all of the rock types, and in

the models that predict the characteristics in the igneous rocks. In these cases, only the models that predict the orientation perform with R^2 scores that are lower than 0.2 of the models that are developed with the time random approach of splitting the training and testing data. The time continuous method of splitting the data produces greater decreases of performance in the models developed with the marble data than the other rocks. The mean decrease in the R^2 scores across the eight predictions of fracture characteristics is 0.07, 0.06 and 0.17 for the three type of models (Fig. S2). Regardless of the shift in some of the R^2 scores, the general conclusions of the analysis remain unchanged when we change the method of splitting the data. We discuss these points in greater detail in subsequent sections. Moreover, previous work demonstrates that the distribution of SHAP values does not change significantly when changing the method of splitting the data (McBeck et al., 2020), indicating that these results remain unchanged as well.

4. Results

4.1. Success of Predicting Fracture Characteristics

We develop the models to predict one fracture network characteristic at a given moment in time, such as the fracture volume or length, using the other characteristics as features. In particular, we predict the median value of one characteristic using the full suite of all the other characteristics. We assess the model performance using the R^2 score between the observed and predicted values. Higher R^2 scores indicate more accurate models: R^2 scores > 0.8 indicate strong positive correlations between the predicted and observed values.

The performance of the models depends on the rock type used to develop the model, and the fracture characteristic that the model predicts (Fig. 4). Figure 4 shows the R^2 scores for the models that predict each fracture characteristic, including the total fracture volume, individual fracture volume, fracture aperture, fracture length, shape anisotropy, distance between fractures in a subvolume, and fracture orientation. The models perform better when they

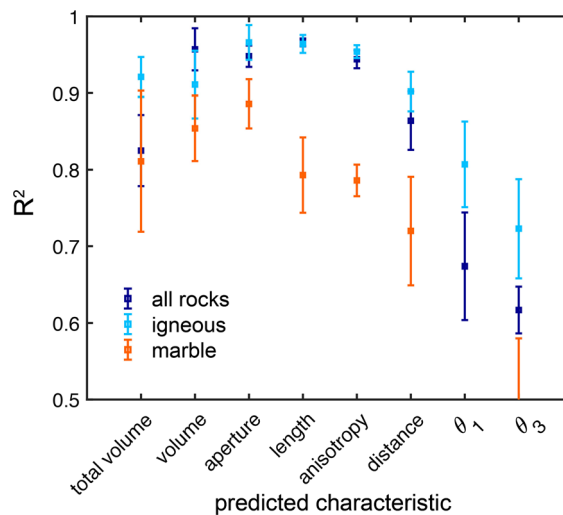


Figure 4

Success of predicting each fracture network characteristic in all of the examined rock types (dark blue), only the igneous rocks (light blue) and only the marble rocks (orange), shown with the R^2 score. We train and test models ten times using different divisions of training and testing data in order to account for random variations. We report the mean \pm one standard deviation of the R^2 scores of the ten models with different training and testing data sets. Thus, the orange symbol above the total volume label on the horizontal axis shows the mean \pm one standard deviation of the scores for the models that predict the total volume of fractures in a subvolume

use data from the igneous rocks than the marble (Fig. 4). The higher number of samples used in the igneous rock models (1551) compared to the marble models (730) may contribute to the higher R^2 scores of these models (Fig. S3). Alternatively, the fracture characteristics may be more predictable and systematic in the igneous rocks than the marble. To test this idea, we develop models with the same number of samples for both rock type models (Fig. S4). When the igneous and marble models have the same number of samples, the igneous models continue to perform better than the marble models (Fig. S4). Thus, the fracture characteristics appear more predictable in the igneous rocks than the marble. When we split the training and testing data using a time continuous method, rather than the time random method, this observation remains unchanged; the models perform better when they are developed with the igneous rock data than the marble data (Fig. S2).

The success of the models also depends on the fracture characteristic that the model predicts. For all of the characteristics that we predict except the total

volume, we predict the median value of that characteristic within a subvolume of the rock core. In particular, we measure the lengths of all the fractures within a subvolume. Then we provide as features to the model the minimum, 25th percentile, 50th percentile, 75th percentile and maximum value of all of these lengths within that particular subvolume. When we predict the length, we predict only the 50th percentile (median) length, and exclude all of the features that include information about the length (i.e., minimum and maximum length, etc.). Thus, the models predict one representative value of each characteristic. When we predict either orientation measure, θ_1 or θ_3 , we exclude all of the features with information about the orientation as these features are closely related to each other. When we predict the individual fracture volume or total number of fractures in a subvolume, we exclude all of the features with information about the fracture volume. The number of features provided to the models can influence the model success: we expect lower R^2 with lower numbers of features. However, the models shown in Fig. 4 all have approximately the same number of features as we calculate the same set of statistics for each fracture characteristic.

The characteristics that are the most predictable, with $R^2 > 0.8$ for the igneous rocks and all rock types models, include the total volume of fractures in a subvolume, volume of individual fractures, fracture aperture, fracture length, fracture anisotropy, and distance between fractures. The characteristics that are the least predictable, with the lowest R^2 , are the orientations of the minimum eigenvector, θ_1 , and maximum eigenvector, θ_3 . The scores for the marble models that predict θ_1 are below 0.5, and thus not visible on Fig. 4. This observed lower range of the scores of models that predict the orientation remains unchanged when we alter the method of splitting the training and testing datasets (Fig. S2). The lower range of the R^2 scores for models that predict the orientations suggest that the other fracture characteristics behave more systematically and predictably than the orientations. Although the R^2 scores are lower for the models that predict the orientations, the models that predict the other characteristics could use the orientations to make their predictions. In the next

section, we examine this question by comparing the importance of features in each model.

4.2. Key Characteristics of Features Required to Predict Fracture Characteristics

To compare the importance of each feature, and thereby examine the degree to which the models rely on the fracture orientations to make successful predictions, we compare the mean |SHAPI value, s , of each feature (Fig. 5). We report the mean |SHAPI value across all the samples rather than for each sample (Fig. 5a–c). Because we develop ten different models that only differ in how their training and testing data are split, we simplify the comparison by calculating the normalized importance as $\bar{s}/\max(\bar{s})$, where \bar{s} is the mean s from the ten models, and thus the mean of the ten mean |SHAPI values (Fig. 5d–f). The similarity of the distribution of s across the ten models indicates that finding the features with the highest normalized importance will yield the same results as examining the s values. To further aid the identification of the features that most strongly influence the model predictions across all three rock types, we calculate a cumulative feature importance across all three rock type models as $\sum R^2(\bar{s}/\max(\bar{s}))$, where R^2 is the mean R^2 score of the ten different models of each rock type (Fig. 6). We weight the normalized \bar{s} by R^2 so that more accurate models, with higher R^2 , have a greater impact on the results than less accurate models. We then identify the top three features with the highest cumulative normalized importance as the features that have the strongest influence on the predictions of a particular fracture network characteristic (Fig. 6).

The fracture characteristics identified as the most important in predicting each fracture characteristic are consistent with some of our previous knowledge and intuition about the dependence of each characteristic on others, following linear elastic fracture mechanics (i.e., Fig. 2). The total volume of fractures in a subvolume and the volume of individual fractures both depend on the aperture and length, but the total volume of fractures also depends on the distance between fractures (Fig. 6a, b). Thus, the spatial localization of the fracture network is also critical to predict the overall fracture density.

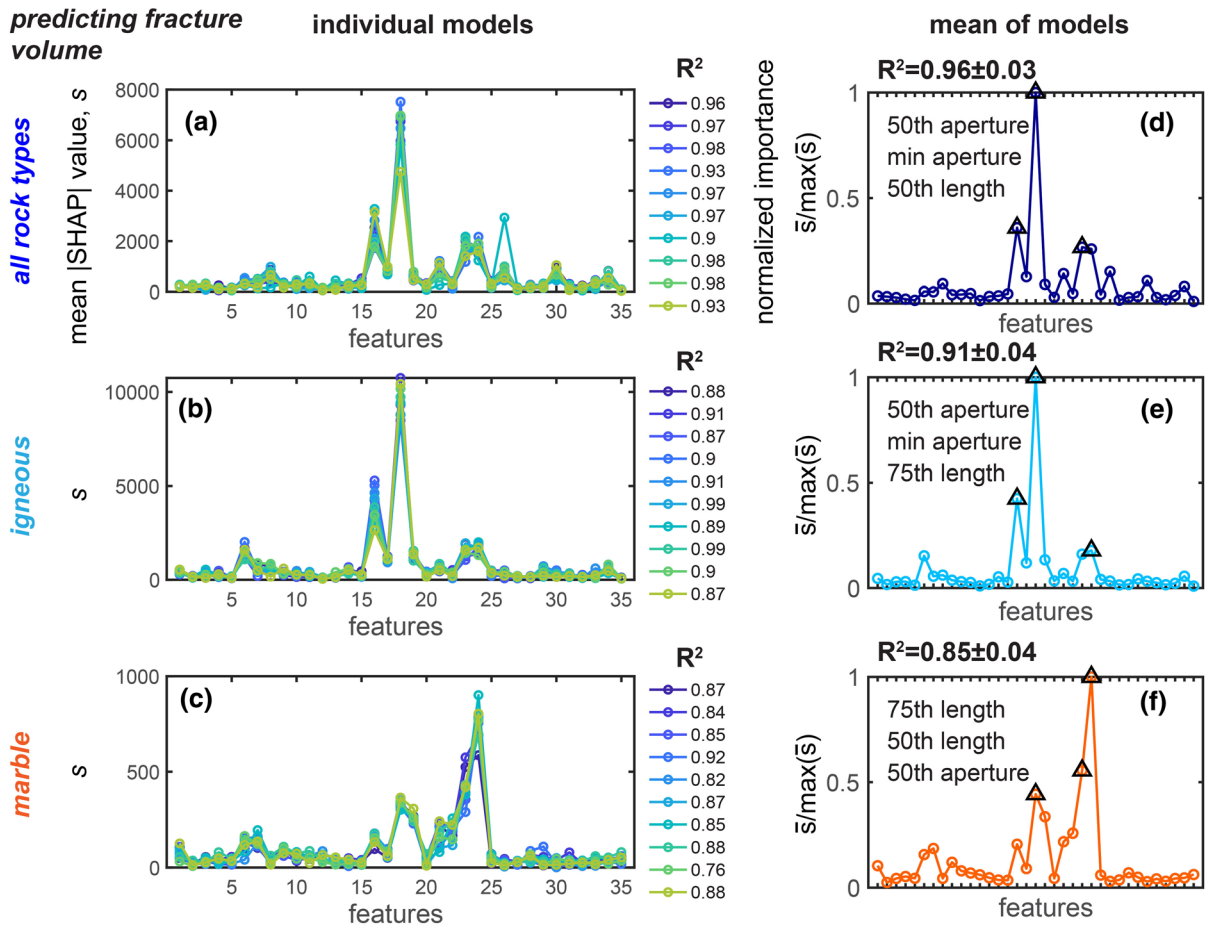


Figure 5

Example distribution of mean |SHAP| values, s , for models designed to predict the fracture volume for all rock types (a, d), igneous rocks (b, e) and marble (c, f). The first column (a–c) shows the results for each individual model and the corresponding R^2 score. The second column (d–f) shows the results as the normalized mean of each |SHAP| value returned for each feature, $\bar{s}/\max(\bar{s})$, and the top three features with the highest importance (triangles and text). The similarity of the distributions of s justifies reporting the mean s of each feature and normalizing this distribution. The features listed in the top left corner of (d–f) are the top three identified from the $\bar{s}/\max(\bar{s})$ distribution. Table S1 lists all of the features

Similarly, predicting the distance between fractures in a subvolume depends primarily on the total volume of fractures in a subvolume (Fig. 6f), as well as the individual fracture volume and aperture. Predicting the aperture and length both depend on the anisotropy, but the aperture depends on the fracture volume as well (Fig. 6c, d). Similarly, the anisotropy depends on the length and aperture, as expected from the fixed relationship between them as $a = 1 - A/L$, where a is the anisotropy, A is the aperture, and L is the length. However, note that because the models predict the median value of each fracture

characteristic, the relationship between the features is not the equation listed above. Predicting the orientations depend on the length and anisotropy, but predicting the orientation of minimum eigenvector, θ_1 , also depends on the aperture (Fig. 6g, h).

None of the top three identified most important features for each of the eight predictions of the fracture network characteristics include the fracture orientation (Table 1, Fig. 6). Similarly, removing the fracture orientations from the features provided to the models does not reduce the R^2 scores of the models (Fig. S5). These results suggest that the fracture

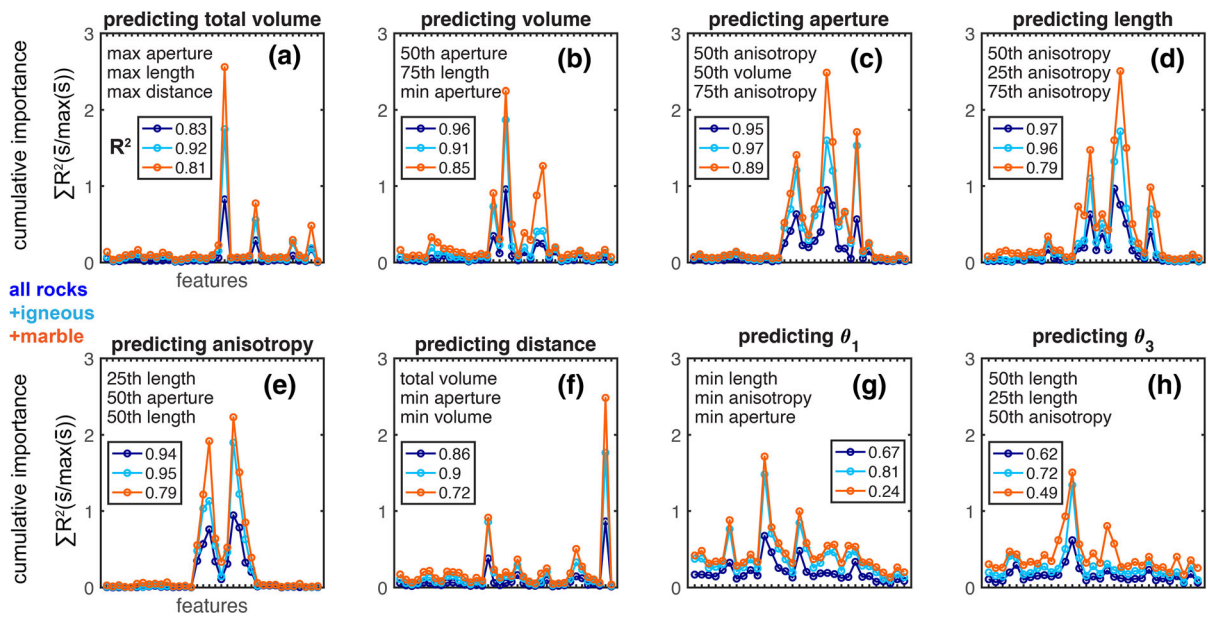


Figure 6

Cumulative importance of features for each predicted fracture characteristic: **a** total volume of fractures in a subvolume, **b** individual fracture volume, **c** fracture aperture, **d** fracture length, **e** fracture shape anisotropy, **f** distance between fractures in a subvolume, and orientations of the minimum eigenvector, θ_1 (**g**), and maximum eigenvector, θ_3 (**h**). Cumulative feature importance is shown as $\sum R^2(\bar{s}/\max(\bar{s}))$, where R^2 is the mean R^2 score of the ten different models of each rock type. We weight the normalized \bar{s} such that more accurate models, with higher R^2 , have a greater impact on the results than less accurate models. The top three most important features are listed in the top left corner of each plot

Table 1

Fracture characteristics of the most important features identified in models. None of the most important features rely on the orientation of the fractures

Prediction	Important characteristics
total Volume	Aperture, length, distance
Volume	Aperture, length
Aperture	Volume, anisotropy
Length	Anisotropy
Anisotropy	Length, aperture
Distance	Volume, aperture
θ_1	Length, aperture, anisotropy
θ_3	Length, anisotropy

orientation is the least systematic of the characteristics, and thus the least reliable in predicting each of the characteristics.

Examination of the statistics used to calculate each feature suggests that analyzing the complete fracture network may produce the most accurate predictions, rather than the extreme members of the

population (Fig. 6). The most important characteristics used to make each prediction generally do not depend on only the extreme values of the population, such as the maximum or minimum, but require the intermediate percentile statistics. This trend applies to five of the predicted characteristics, including the individual fracture volume, aperture, length, anisotropy, and θ_3 . However, the three remaining predicted characteristics (the total fracture volume, distance between fractures, and θ_1) primarily depend on the extreme values. The high R^2 scores of the models that predict the total fracture volume and distance between fractures indicate that we may only need to classify the extreme values of the fracture network characteristics in order to make successful predictions of these characteristics.

4.3. Predicting Fracture Characteristics with Limited Data

To further test the idea that the fracture orientation is one of the least predictable and systematic

characteristics of fractures, we develop models that use fewer and fewer features (Fig. 7). In particular, we remove the fracture characteristics from the features provided to the models that we identify as the top three most important fracture characteristics (Table 1, Fig. 6). When the models have less information about the fracture network, they may begin to use the orientation for their predictions.

As expected, the R^2 scores decline after removing the most important features for several of the fracture characteristics. These declines are most rapid when predicting the total volume, individual volume, aperture, anisotropy and θ_1 (Fig. 7). This result suggests that the other characteristics, the distance between fractures and θ_3 , can successfully use the remaining fracture characteristics to make predictions with similar success as the original models that have

access to all of the features. Excluding the most important features reduces the R^2 scores of the marble models by larger magnitudes than the models developed with the other rock types for all but one prediction (θ_1) (Fig. 8). The R^2 scores generally decrease by 0.15–0.50 for the marble models, 0–0.35 for the igneous models, and 0–0.30 for the models that include all of the rock types (Figs. 7, 8).

Removing the most important features for each predicted characteristic causes the fracture orientation to become one of the most important features for all of the predictions (Fig. 9, Table 2). However, only removing one of the three identified most important fracture characteristics does not promote the orientation to the class of highly important features for the majority of the predicted characteristics (Fig. S6). Only after removing all top three characteristics do

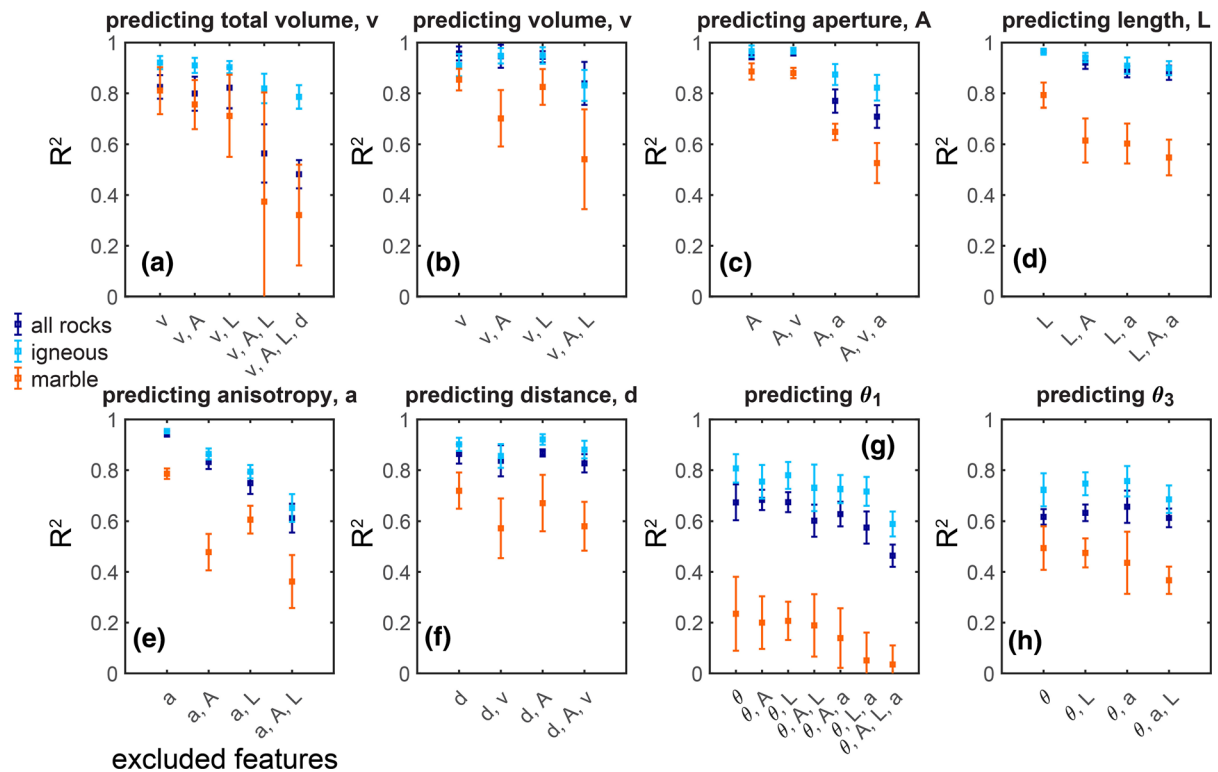


Figure 7

Success of predicting fracture network characteristics upon removal of the most important features (i.e., Table 1, Fig. 5). The horizontal axis lists the features that were removed before developing each model that predicts the total volume, v , (a), individual fracture volume, v , (b), aperture, A , (c), length, L , (d), anisotropy, a , (e), distance between fractures, d , (f), θ_1 (g), and θ_3 (h). For example, the scores listed in the position marked v, A, L, d in (a) show the scores for models developed to predict the total volume without using features that include information about the volume, v , aperture, A , length, L , and distance, d

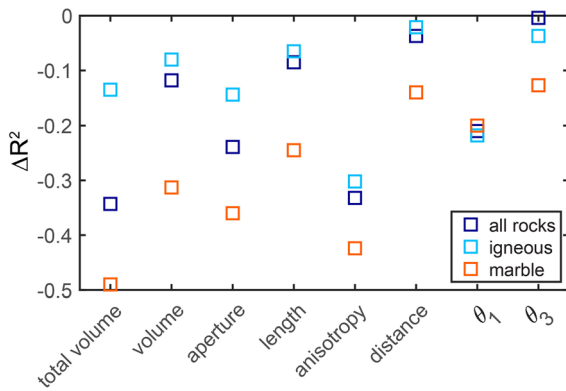


Figure 8

Difference in the mean R^2 score for models that include all of the features, R_a , and models that exclude all of the most important features, R_e , as $\Delta R^2 = R_e - R_a$ for models developed with the three groups of rock types. The decrease in R^2 is the largest for the marble models for all but one fracture characteristic, θ_1

all of the models use the orientation to predict the examined fracture characteristic. This result suggests that the fracture orientation is systematic enough to provide reasonable predictions about the other fracture characteristics. However, the R^2 scores of the models that rely on the orientation are generally at least 0.1 lower than the models that use all of the characteristics (Fig. 8). For example, models that predict the aperture for all the rock types and use all of the characteristics as features have a mean R^2 score of 0.95, and depend primarily on the anisotropy and volume (Fig. 9c). This type of model, but developed without the anisotropy and volume, produce a mean R^2 score of 0.71, and depends primarily on the length and θ_1 . Furthermore, the models in which the top three features only include information about θ_1 and θ_3 (predicting the anisotropy, Fig. 9e) have the lowest R^2 scores (0.61, 0.65, 0.36 for all of the rock types, igneous rocks, and marble, respectively) of the models that exclude the identified most important characteristics.

5. Discussion

This analysis assesses the accuracy of preexisting conceptualizations of fracture network development, derived from linear elastic fracture mechanics and laboratory observations (Fig. 2). The varying success

with which the machine learning models predict the selected fracture network characteristics, and the features they use to make these predictions, help constrain the processes of fracture network development that occur during triaxial compression deformation under the conditions of the upper crust. Consistent with the idea that increasing differential stress increases the total volume of fractures and decreases the distance between fractures (Fig. 2a), the models that predict the total fracture volume depend on the distance between fractures, and the models that predict the distance depend on the total fracture volume (Fig. 6). Consistent with the idea that fracture propagation and coalescence should increase the fracture length and anisotropy (Fig. 2b), the models that predict the fracture length depend on the anisotropy, and the models that predict the anisotropy depend on the length (Fig. 6). Consistent with the idea that increasing differential stress should increase the length of fractures and change the fracture orientation (Fig. 2c), the models that predict the fracture orientation depend on the length (Fig. 6).

Thus, the evolution of these fracture characteristics matches our expectations. However, we note that previous laboratory experiments that analyzed these evolutions did not typically have access to the four dimensional observations of fracture network characteristics. The few other experiments that used X-ray tomography to make such observations (e.g., Cartwright-Taylor et al., 2020) did not apply machine learning to quantify the applicability of previous conceptualizations of fracture network development. Thus, this work is the first, to our knowledge, to quantify the correctness of these fundamental ideas about fracture network development.

In this discussion section, we examine why the fracture network characteristics are more unpredictable in the marble rocks than the igneous rocks. Then, we discuss the observed relationship between fracture length and orientation, and the associated lower predictability of the orientation compare to the other characteristics. Finally, we link these results to observed precursors to some earthquakes. We note that we did not explicitly test the ability of the fracture characteristics to predict the timing of catastrophic failure in the present analysis. However, we suggest that the predictability of the

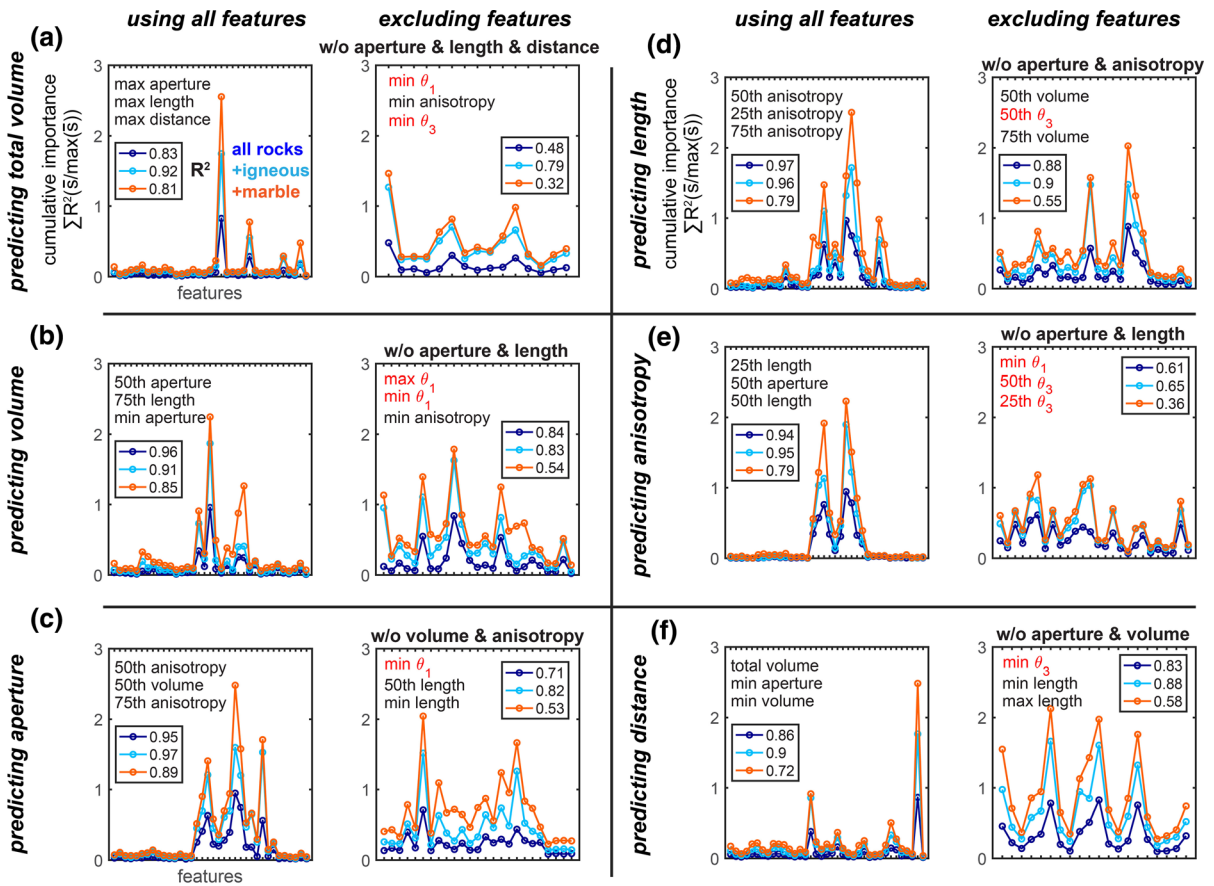


Figure 9

Cumulative importance of features for each predicted fracture characteristic: **a** total volume of fractures in a subvolume, **b** individual fracture volume, **c** fracture aperture, **d** fracture length, **e** fracture shape anisotropy, and **f** distance between fractures in a subvolume. Cumulative feature importance is shown as $\sum R^2(\bar{s}/\max(\bar{s}))$, where R^2 is the mean R^2 score of the 10 different models of each rock type and s is the mean ISHAPI value of a given feature. We weight the normalized feature importance \bar{s} such that more accurate models, with higher R^2 , have a greater impact on the results than less accurate models. Left columns show the results for models developed using all of the features (except the one being predicted). Right columns show the results for models developed using a subset of the features that excludes the identified most important features (listed in the left column)

Table 2

Most important characteristics identified in models with the listed features excluded

Prediction	Excluded features	Important characteristics
Total volume	Aperture, length, distance	θ_1, θ_3 , anisotropy
Volume	Aperture, length	θ_1 , anisotropy
Aperture	Volume, anisotropy	θ_1 , length
Length	Anisotropy, aperture	θ_3 , volume11
Anisotropy	Length, aperture	θ_1, θ_3
Distance	Volume, aperture	θ_3 , length

characteristics may help crustal monitoring efforts that aim to forecast the timing of large earthquakes. For example, machine learning analyses suggest that the fracture shape anisotropy is a key characteristic for predicting the timing of macroscopic failure in triaxial compression experiments. However, it is difficult to measure this property in the field. If we can derive equations or relationships between the measurable fault characteristics (such as length) and the characteristics that are useful for predicting the timing of failure (such as anisotropy), then we may improve efforts to forecast the timing of large earthquakes. This approach may also prove fruitful

for processes that have the most documented, systematic precursors, such as volcanic eruptions and avalanches.

5.1. *Erratic Nature of Fracture Development in Marble*

Comparing the R^2 scores of the models developed using data from the experiments performed on igneous rocks and those performed on marble suggest that fracture development in marble is less predictable (with lower R^2 scores) than in igneous rocks (Fig. 4). Identifying the fracture characteristics that produce the largest discrepancy in the R^2 scores between the rock types highlights the aspects of fracture network development that are the least predictable, and thus least systematic. The mean R^2 scores of the marble and igneous rock models are similar when the models predict the total volume of fractures in a subvolume, the individual fracture volume and fracture aperture. In contrast, the mean R^2 scores differ by more than 0.1 when the models predict the fracture length, anisotropy, distance between fractures, and orientation. Thus, these characteristics are less systematic and thus less predictable in marble than in the igneous rocks.

By examining the correlation between sets of fracture characteristics in the marble and igneous rock experiments, we may gain further insight into the difference between fracture network development in these rock types, and their varying degrees of predictability. As mentioned above, the R^2 scores are similar in the marble and igneous rock models when they predict the total volume of fractures (Fig. 4). The top three most important features that determine this model prediction are the maximum aperture, maximum length and maximum distance. Because the two rock type models perform similarly when they predict the total volume, we expect that the three most important features are correlated to the total volume with similar strengths in both rock types. To quantify potential correlations between the fracture characteristics, we calculate the linear correlation coefficient between these features and predictions. Because the XGBoost machine learning model can develop non-linear relationships between features and predictions, the linear correlation coefficient may not

fully represent the strength of the function between the feature and prediction in the model. However, the linear correlation coefficients provide rough approximations of the relationships between these fracture characteristics, and are thus useful to examine.

As expected, calculating the linear correlation coefficient between each set of features for both rock types (Fig. 10a, b) indicates that both the marble and igneous rocks have strong correlations between the total volume and maximum aperture (0.73, 0.82), and moderate correlations between the total volume and maximum length (0.60, 0.61). The two rock types differ in their strength of correlation of the total volume and maximum distance; the marble experiments have a correlation coefficient of 0.63, while the igneous rock experiments have a correlation coefficient of 0.20. Although the third most important feature (maximum distance) has a wider range of correlations between the rock types, the similarity of the R^2 scores of the different rock type models indicates that the strong correlations of the top two features outweigh the influence of the lower correlation strength of the third most important feature.

As a corollary to the idea that the correlation coefficients help describe the differences in the model performance of the different rock types, we expect that models with lower R^2 scores should have lower correlations between the predicted fracture characteristic and the most important features. When the models predict the fracture length, the marble models have R^2 scores at least 0.10 lower than the igneous rock models (Fig. 4). Thus, we expect stronger correlations between the fracture length and the top three important features in the igneous rocks than the marble. Indeed, the correlation coefficients between the length and the three percentile statistics of the anisotropy range from 0.58 to 0.63 in the igneous experiments and 0.49–0.57 in the marble experiments (Fig. 10c, d). Although the ranges of these coefficients differ by only 0.01, the R^2 scores of the models vary significantly, highlighting that even apparently small differences in the systematic nature of the fracture characteristics produce large changes in the predictability of these characteristics.

The observed difference in the predictability of fracture network development in the marble and igneous rocks likely arises from the varying

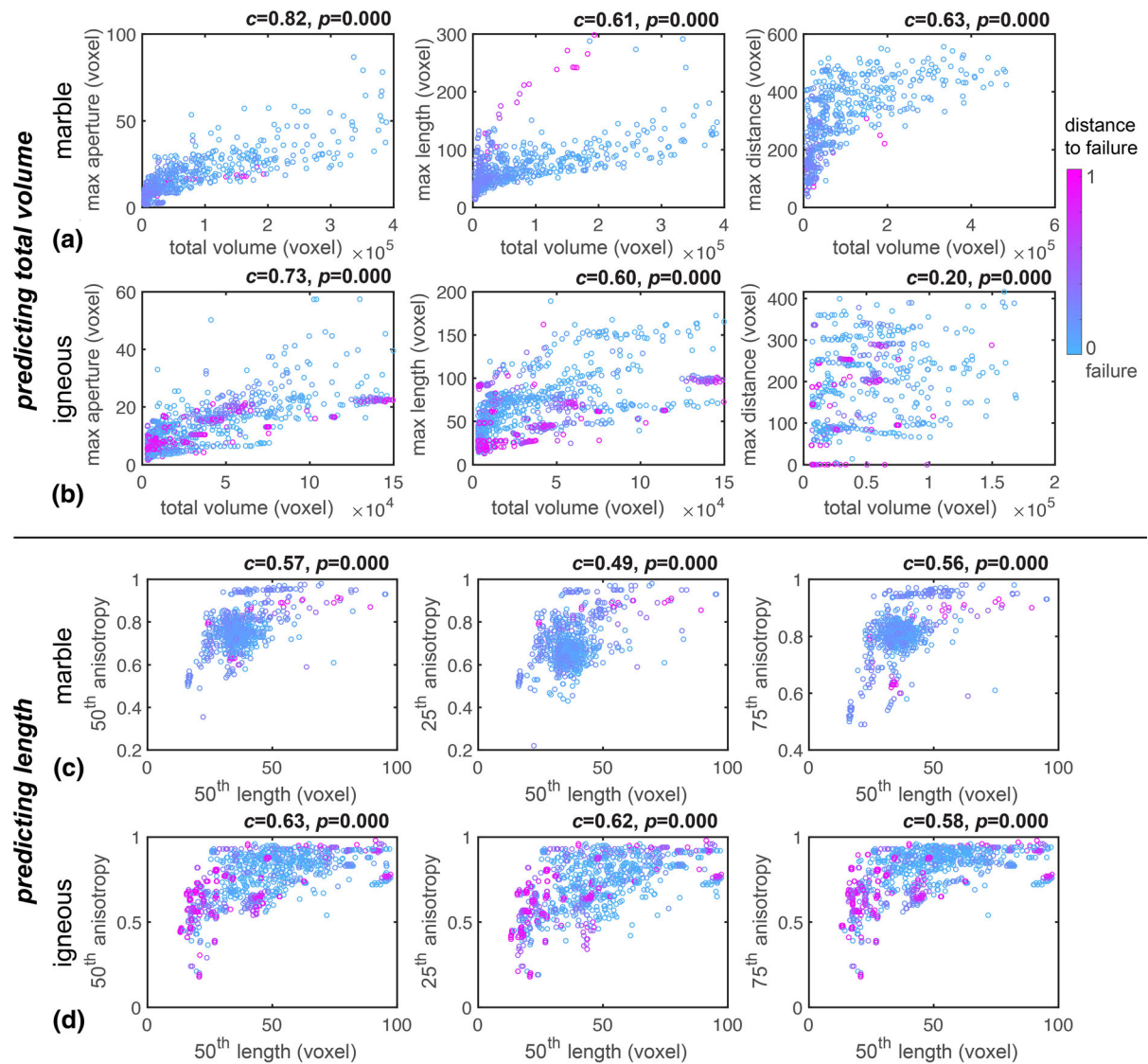


Figure 10

Relationship between fracture characteristics that the models predict (horizontal axes) and the top three most important features identified for that model (vertical axes). The values at the top of each plot are the linear correlation coefficient, c , and the p-value, p , of the two fracture characteristics shown in the plot. Coefficients with $p < 0.05$ are considered significant. We present the data from the marble (a, c) and igneous rocks (b, d) separately to better understand the varying predictability of the total volume (a, b) and fracture length (c, d) in these rock types. The color of the symbols indicates the distance to failure, $(\sigma_F - \sigma_D)/\sigma_F$, where σ_F is the differential stress at macroscopic failure, and σ_D is the differential stress when the tomogram was acquired. Zero is at macroscopic failure. The marble and igneous rock models have similar R^2 scores when they predict the total volume, but different R^2 scores when they predict the fracture length. The differences in c between the marble data and igneous rock data for each pair of characteristics agree with these differences. In particular, the correlations between the total volume and the top three highly important features are similar in the igneous rocks and marble data, but differ between the fracture length and top three highly important features

microstructures that control fracture nucleation, propagation, and arrest in both rocks. Marble is composed of cemented and metamorphosed calcite

grains that contain weak cleavage planes, whereas monzonite and granite are composed of stronger and interconnected feldspar and granite minerals.

Uniaxial compression experiments support the idea that cleavage planes and grain boundaries have a strong influence on fracture development in marble (e.g., Fredrich et al., 1989; Kandula et al., 2019; Tal et al., 2016). Fredrich et al. (1989) observed that fractures often nucleate at sites of local stress perturbations, such as twin boundaries, twin terminations and the intersection of twin lamellae. Tal et al. (2016) observed that fractures first develop near grain boundaries in marble. With increasing differential stress, these fractures grow along grain boundaries and then transect the boundaries near peak stress. Calculation of the local strain during fracture development suggests that the fractures that form within grains could exploit the cleavage plans as mechanical twins (Tal et al., 2016). Although mineral boundaries can influence fracture propagation in crystalline rock (e.g., Moore & Lockner, 1995; Tapponnier & Brace, 1976), the mechanical heterogeneities in sedimentary and metamorphosed sedimentary rock, including weak cleavage planes within calcite (e.g., Fredrich & Wong, 1986; Howarth, 1987; Tal et al., 2016), appear to exert a stronger influence on fracture propagation and arrest than the heterogeneities in crystalline rock. Thus, fracture propagation in the igneous rocks may more closely align with the idealized expectations of fracture growth during brittle deformation outlined in Fig. 2. In contrast, in marble the rock microstructure may prevent the idealized growth that would lead to a systematic relationship between fracture characteristics, such as fracture length and anisotropy, for example.

The degree of fracture localization may also contribute to the greater predictability of the fracture network characteristics within the igneous rocks than the marble rocks. At the confining stress (20–25 MPa) and ambient (room) temperature conditions of these experiments, ductile processes may accommodate a non-negligible portion of the deformation in the marble rocks (e.g., Griggs, 1960; Turner et al., 1954; Walton et al., 2015). In contrast, under these conditions, brittle processes dominate deformation in the igneous rocks (e.g., Tullis & Yund, 1977). Thus, ductile processes may prevent the localization of the fracture network to a greater extent in the marble rocks than the igneous rocks. Previous

work indicates that the localization of deformation is a key characteristic of approaching system-scale failure (e.g., Ben-Zion & Zaliapin, 2020; Kato & Ben-Zion, 2021; Lockner et al., 1991; Renard et al., 2019; Stanchits et al., 2006) that we may use to accurately predict the timing of failure (McBeck et al., 2020). Thus, fracture networks that localize toward failure may be more predictable than fracture networks that remain more distributed. The greater predictability of the fracture networks in the igneous rocks relative to the marble rocks may arise from the dominance of brittle deformation that tends to localize the fracture network in the igneous rocks. These results thus suggest that fault network development in crust dominated by igneous rocks may be more predictable than fault network development in crust dominated by marble.

5.2. Fracture Orientation as an Unreliable Predictor

Consistent with the idea that increasing differential stress should increase the length of fractures and change the fracture orientation (Fig. 2c), the models that predict the fracture orientation depend on the length (Fig. 6). However, the models that predict the fracture length do not depend on the orientation (Fig. 6). Moreover, the models that predict the fracture orientation are the least accurate of the models, with the lowest R^2 (Figs. 4, S1), and removing the orientation from the available features does not reduce the R^2 scores (Fig. S5). In addition, the top three highly important fracture characteristics used to predict each fracture characteristic do not rely on information about the orientation (Fig. 6, Table 1). We must remove these top three characteristics from the data used to develop the models in order for the models to depend on the orientation (Fig. 9). Importantly, these models perform worse than the models that do not depend on orientation, with at least 0.1 lower R^2 scores (Figs. 7, 8).

The observed unpredictable nature of fracture orientation suggests that the idealized view about the transition from mode I failure to mode II/III failure requires reexamination (Fig. 2). In this conceptualization, under lower differential stress fractures first open perpendicular to the maximum compression direction, σ_1 , in mode I failure. Then with increasing

differential stress, as the fractures coalesce and grow longer, wing cracks and other linking fractures may propagate between the older fractures, producing an overall fracture network that trends oblique to σ_1 .

Qualitatively, we observe fracture network development in X-ray synchrotron experiments that agree with this conceptualization (e.g., Cartwright-Taylor et al., 2020; Renard et al., 2018, 2019). For example, in one of the granite experiments analyzed here, visual inspection suggests that the fracture network trends obliquely to σ_1 in the scans preceding macroscopic failure (Fig. 1). Moreover, tracking the orientation of all the fractures identified in a tomogram at a given differential stress indicates that the mean orientation of these fractures evolves from parallel to more oblique to σ_1 with increasing differential stress (Renard et al., 2018, 2019). However, the mean fracture orientation is not a complete representation of the distribution of fracture orientations present at a particular stress step. Rather than the simplistic idea that the majority of the fractures gain an oblique orientation with increasing differential stress, a more precise description of the fracture network may be that while many fractures gain an oblique orientation, some also continue to trend parallel to σ_1 , and thus maintain their original orientation.

We may test the idea that fractures evolve to different orientations with increasing differential stress with our experimental data. If fractures develop with the idealized conceptualization described above, we expect that the length of fractures should correlate with the orientation. Smaller fractures should tend to trend parallel to σ_1 while longer fractures have more oblique orientations. By calculating the linear correlation coefficient between the fracture length and orientation, we may quantify the degree to which the observations match these expectations. (Fig. 11). These coefficients indicate a weak or non-existent relationship between the orientation and the fracture length. This result holds for the orientation of the smallest dimension of the fracture relative to the maximum compression direction, θ_1 (Fig. 11), and the longest dimension, θ_3 (Fig. S7). Thus, longer fractures do not tend to have systematically different orientations from shorter ones.

The difference between the observed and the expected relationship of length and orientation may arise from the influence of local stress perturbations on fracture growth, and the resulting misapplication of the Coulomb criterion to heterogeneous rocks at the local scale. The Coulomb criterion, and corresponding relationship between fault orientation, principal stress directions, and friction coefficient, was originally developed to describe the orientation of faults within granular aggregates that behave plastically (Coulomb, 1776). Unsurprisingly, the predictions of this criterion for faults within granular aggregates, and material with similar rheology, closely match observations (e.g., Dahlen, 1984, 1990; Dahlen et al., 1984; Davis et al., 1983; Huiqi et al., 1992; Lallemand et al., 1994; McBeck et al., 2017; Mulugeta, 1988; Vermeer, 1990). However, abundant experimental observations such as decreasing seismic wave velocities and macroscopic radial dilation indicate that increasing differential stress can open fractures. Fracture opening changes the stress state around the fracture (Inglis, 1913) and reduces or eliminates the relevance of the coefficient of friction used in the Coulomb criterion (Peng & Johnson, 1972). The influence of opening may also explain the observed curvature of the shear stress and normal stress relationship at less compressive/more tensional stresses, and linearity of this relationship at more compressive stresses. Thus, the physical justification of applying the Coulomb criterion to brittle materials with many similarly-sized fractures, rather than one system-size fault, may be unfounded (e.g., Peng & Johnson, 1972). This apparent lack of physical justification agrees with our experimental observations: we only observe a weak or non-existent correlation between fracture length and orientation (Figs. 11, S7), and the fracture orientation is the least predictable of the fracture characteristics (Fig. 4).

Examining the geometry of the fracture network in the final tomogram captured immediately preceding macroscopic failure provides further insight into the lack of clear relationship between fracture length and orientation (Fig. 12). In experiments on granite and marble, we observe that the complete fracture network appears to trend at the expected oblique orientation to σ_1 (middle in Fig. 12). However, when

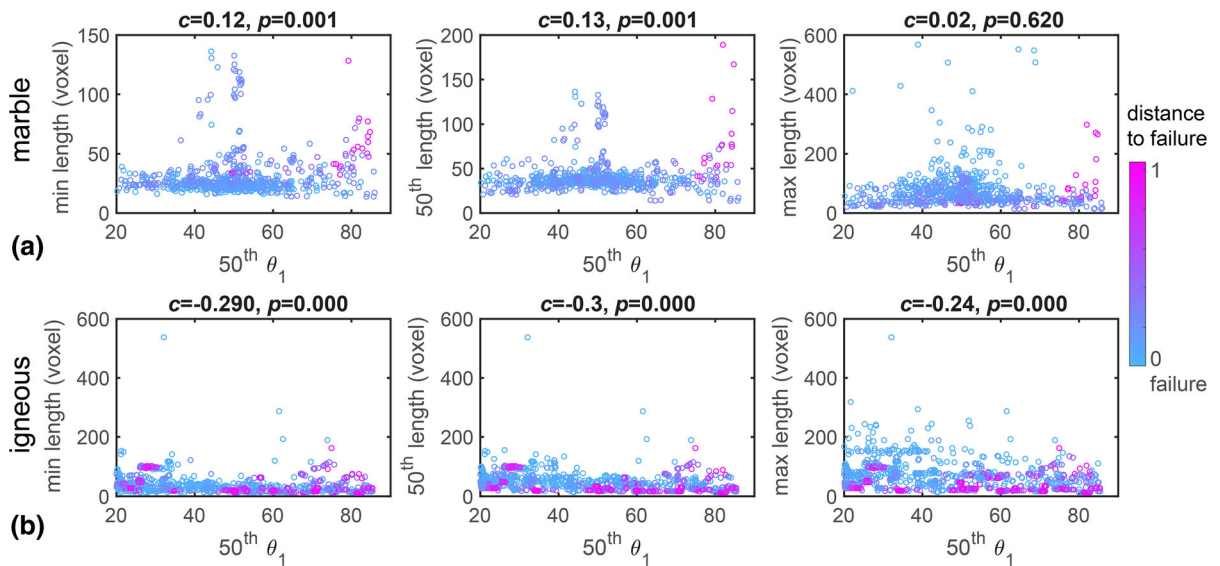


Figure 11

Relationship between the fracture orientation, θ_1 , and representative statistics of the fracture length for the **a** marble and **b** igneous rocks. The values at the top of each plot are the linear correlation coefficient, c , and the p-value, p , of the two fracture characteristics shown in the plot. Coefficients with $p < 0.05$ are considered significant. The color of the symbols indicates the distance to failure, $(\sigma_F - \sigma_D)/\sigma_F$, where σ_F is the differential stress at macroscopic failure, and σ_D is the differential stress when the tomogram was acquired. Zero is at macroscopic failure. There is no clear trend between the applied differential stress (i.e., distance to failure), and the fracture orientation. The low c between the orientation and length indicate that longer fractures are not more likely to attain the expected orientation oblique to σ_1 . Rather, long and short fractures have a wide range of orientations. Figure S7 shows the relationship between these fracture length statistics and the orientation relative to the maximum dimension of the fracture, θ_3

we examine the individual and largest fractures with volumes > 5000 voxels (right in Fig. 12), we observe that few of the largest fractures trend at the expected orientation. For example, although the dark blue fracture that dominates the marble core hosts this orientation (Fig. 12b), it is surrounded by many other large fractures that do not trend at 30° from σ_1 . Moreover, examination of the 2D slice of the tomogram (left in Fig. 12) indicates that the fracture network that appears to have an overall orientation near 30° from σ_1 is dominated by many smaller fractures that are oriented near parallel to σ_1 . This observation agrees with previous experimental results (e.g., Peng & Johnson, 1972).

These experimental observations suggest that the behavior shown in Fig. 13 is closer to reality than the idealized view of fracture development outlined in Fig. 2. Under lower differential stress, mode I fractures may first develop parallel to σ_1 , as expected from the lower tensile strength than shear strength of brittle rocks (e.g., Tapponnier & Brace, 1976). With

increasing differential stress, these preexisting fractures propagate, some coalesce, and new fractures may nucleate. New fracture nucleation and preexisting propagation increase the total fracture volume and decrease the distance between fractures. Propagation and coalescence increase the fracture length and anisotropy while perhaps decreasing the aperture. New fractures may nucleate at an oblique orientation to σ_1 when the system is under higher differential stress and the local stress perturbations favor this formation. However, local stress perturbations may cause the local orientation of σ_1 to deviate from the macroscopic orientation. When the fractures are small enough, these local perturbations may exert a significant impact on the direction of their growth. Thus, depending on the initial configuration of the fractures, the preexisting fractures that formed in tension parallel to σ_1 may maintain a similar orientation (fractures on the left in Fig. 13) or may form a fracture network that has a more oblique orientation to σ_1 (fractures on the right). In other

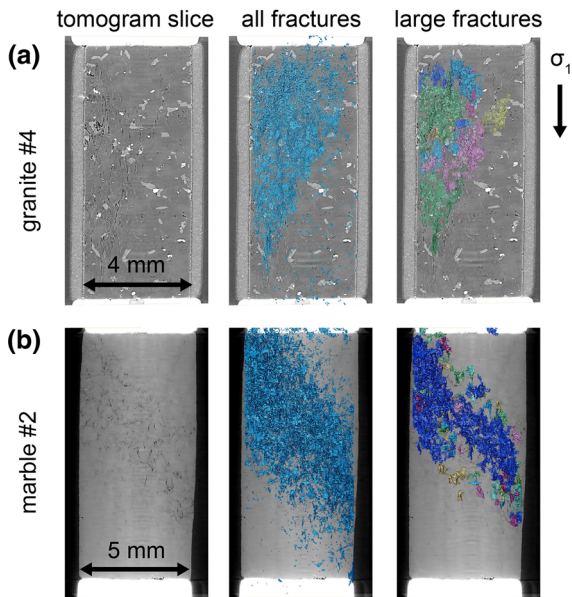


Figure 12

Fracture network development in the final tomograms acquired immediately preceding macroscopic failure in the granite #4 (a) and marble #2 (b) experiment. The columns show the 2D slice of the tomogram (left), the complete fracture network (middle), and the fractures with volumes > 5000 voxels (right). Each of the large fractures are colored differently from each other in order to highlight their varying orientations

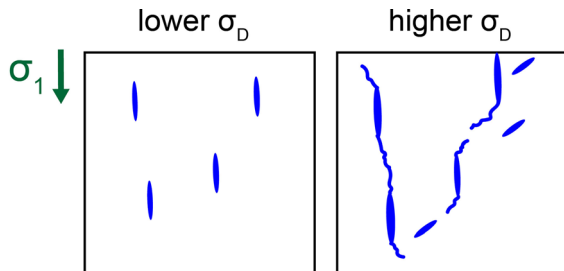


Figure 13

Conceptualization of fracture network development revealed by machine learning models. The fracture orientations under higher differential stress may not match the expected Coulomb criterion due to the initial configuration of the fractures, and resulting local stress perturbations

words, fractures may coalesce with a geometry that maintains their original orientation if they happen to form with only a small offset from each other. In contrast, fractures may coalesce with a geometry that produces a connected fracture network that trends oblique to σ_1 if they have a larger offset from each other. The influence of local stress perturbations likely contributes to the nonsystematic relationship

between fracture orientation and length observed in our data. This conceptualization agrees with experimental observations (e.g., Fig. 12) that indicate that aligned arrays of mode I fractures that trend parallel to σ_1 comprise the larger fracture networks that may appear to trend oblique to σ_1 (e.g., Horii & Nemat-Nasser, 1985; Nemat-Nasser & Horii, 1982; Peng & Johnson, 1972). This idea also agrees with observations of acoustic emissions in laboratory experiments. These analyses indicate a shift from tensile to shear-dominated deformation with increasing differential stress; however, both forms of deformation operate throughout deformation (e.g., Graham et al., 2010). Thus, the prevalence of both shear and tensile deformation indicate that fractures may be aligned with a range of orientations indicative of tensile to shear dominated deformation, such as parallel to the maximum compression direction, to oblique to it.

5.3. Fracture Volume as a Reliable Predictor

While the fracture orientation is the least predictable fracture network characteristic, the fracture volume is one of the most predictable in our experiments (Fig. 4). This result agrees with previous machine learning analyses on the suite of experiments analyzed here (McBeck et al., 2019, 2020). In these analyses, the fracture volume is one of the most important characteristics that determines whether a fracture propagates (McBeck et al., 2019), and that successfully predicts the timing of catastrophic failure (McBeck et al., 2020). The importance of the fracture volume in these three different predictions agrees with the dilatancy diffusion hypothesis. This idea builds from observations of changes in hydrologic activity preceding some earthquakes (Nur, 1974). It proposes that as fractures propagate, open and coalesce, the evolving fracture networks change hydrologic activity, as well as the P- and S-wave seismic velocities and effective moduli. Dilation of the rock, through fracture network development, thus produces the observed geophysical activity. Such accelerations in geophysical activity have been observed prior to some earthquakes, but certainly not all earthquakes (e.g., Amoroso & Crescentini, 2010). The growing fractures, and increasing volume of fractures increases the background damage that

weakens the rock volume and makes it more susceptible for a large failure event (Kurzon et al., 2019; Lyakhovsky et al., 2001). Such a process of increasing background damage before large failure events has also been observed before several $M > 7$ earthquakes in southern and Baja California (Ben-Zion & Zaliapin, 2019, 2020). Such an increase in fracture volume may also enhance the hydrologic and related geophysical activity, as reported before some large earthquakes (e.g., Nur, 1974). In particular, the opening and lengthening of fractures, and fracture coalescence, increases the porosity, reduces the effective moduli, and can alter hydrologic activity in the volume of crust surrounding the main fault preceding catastrophic failure. Thus, fracture network development that produces dilation can trigger geophysical precursors. The results of the present study indicate that the fracture network characteristics that produce the dilation of the rock are among the best predicted characteristics. This result thus provides additional evidence that geophysical precursors linked to dilation may be useful for predicting the timing of catastrophic failure.

However, the success of using the dilation of the crust to predict the timing of failure depends on the repeatability of such signals. In the best case scenario, such dilation would occur in the same location along a fault, and with a magnitude required to exceed some noise threshold. Given the heterogeneous nature of earthquake nucleation, and limited resources for instrumenting active seismic fault zones, using evolving fracture characteristics to predict the timing of failure may not yet improve our ability to predict earthquakes significantly beyond existing methods. Future work should focus on identifying the locations surrounding an active fault that are most likely to predict the strongest signals of precursory deformation. Identifying these locations may help the ability of this approach to improve crustal monitoring efforts.

6. Conclusions

We use X-ray tomography data and machine learning to test several conceptualizations of fracture network development in crystalline rock under

triaxial compression at the laboratory scale. We develop the models to predict fracture network characteristics using other characteristics, without knowledge of the macroscopic stress or strain imposed on the rock. We find that the models perform worse when they predict the fracture characteristics in the marble experiments, rather than the experiments on granite and monzonite. The varying mechanisms of fracture nucleation, growth and arrest in marble and igneous rocks may produce this difference in the predictability of fracture network development. Although the grain boundaries in granite and monzonite can influence fracture development, the grain boundaries and weak cleavage planes within calcite in the marble appear to exert a stronger influence on fracture development than the grain boundaries in the igneous rocks. This stronger mechanical heterogeneity may cause fracture development to deviate from the idealized conceptualizations of fracture growth (i.e., Fig. 2) more in the marble rocks than the igneous rocks. The varying dominance of brittle and ductile deformation mechanisms in the marble and igneous rocks and the corresponding influence on the degree of fracture localization may also contribute to the greater predictability of the fracture networks in igneous rocks than the marble rocks. These results indicate that fault network development in rocks dominated by igneous rocks may be more predictable than that development in rocks dominated by marble.

We find that the fracture volume, length and aperture are the most predictable of the examined characteristics. This result supports the idea that the geophysical signals indicative of dilation are critical for accurate predictions of the timing of catastrophic failure, such as earthquakes, avalanches, and volcanic eruptions. The least predictable examined fracture characteristic is the fracture orientation relative to the maximum compression direction. The fracture characteristics that are identified as the most important in predicting each characteristic do not include the orientation. Moreover, when we remove the orientation from the available features, the performance of the models does not decline. Some of the models depend on the orientation only after we exclude the top three most important characteristics used to make each prediction from the available features. However,

these models perform significantly worse than the models that do not depend on the orientation. Thus, in contrast to the idea that fractures transition from mode I to mode II/III, the orientation of a fracture is not a predictor of its length. Instead, there appears to be no strong or moderate correlation between fracture length and orientation in these experiments. The influence of local stress perturbations and resulting deviation from the fracture orientation expected from the Coulomb criterion may explain this difference in the observed and expected relationship between length and orientation.

Acknowledgements

We thank Elodie Boller, Paul Tafforeau, and Alexander Rack for providing advice on the design of the tomography setup, and Sanchez Technology for building the deformation apparatus. The Norwegian Research Council (grant 300435 to JAM) and U.S. Department of Energy (award DE-SC0016520 to YBZ) funded this work. Beamtime was allocated at the European Synchrotron Radiation Facility (Long Term Proposal ES-295). The experimental data (time series of X-ray tomograms) are available on Norstore: including the three monzonite samples (Renard, 2017, 2018), the two Carrara marble samples (Renard, 2019) and the three Westerly granite samples (Renard, 2021). The code used to build the machine learning models are available on Github (https://github.com/jmbeck/Machine_Learning_Rock_Deformation). We thank Editor Christian David, Ian Main, and an anonymous reviewer for suggestions that improved the manuscript.

Funding

Open access funding provided by University of Oslo (incl Oslo University Hospital). The Norwegian Research Council (grant 300435 to JAM) and U.S. Department of Energy (award DE-SC0016520 to YBZ) funded this work.

Availability of data

The experimental data (time series of X-ray tomograms) are available on Norstore: including the three monzonite samples (Renard, 2017, 2018), the two Carrara marble samples (Renard, 2019) and the three Westerly granite samples (Renard, 2021).

Code availability

The code used to build the machine learning models are available on Github (https://github.com/jmbeck/Machine_Learning_Rock_Deformation).

Declarations

Conflicts of interest There are no conflicts of interest.

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

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Aben, F. M., Doan, M. L., Mitchell, T. M., Toussaint, R., Reuschlé, T., Fondriest, M., Gratier, J.-P., & Renard, F. (2016). Dynamic fracturing by successive coseismic loadings leads to pulverization in active fault zones. *Journal of Geophysical Research: Solid Earth*, *121*(4), 2338–2360.
- Amoruso, A., & Crescentini, L. (2010). Limits on earthquake nucleation and other pre-seismic phenomena from continuous strain in the near field of the 2009 L'Aquila earthquake. *Geophysical Research Letters*, *37*, L10307.
- Andrew, M. (2018). A quantified study of segmentation techniques on synthetic geological XRM and FIB-SEM images. *Computational Geosciences*, *22*(6), 1503–1512.

- Ashby, M. F. A., & Hallam, S. D. (1986). The failure of brittle solids containing small cracks under compressive stress states. *Acta Metallurgica*, 34(3), 497–510.
- Ashby, M. F., & Sammis, C. G. (1990). The damage mechanics of brittle solids in compression. *Pure and Applied Geophysics*, 133(3), 489–521.
- Ayling, M. R., Meredith, P. G., & Murrell, S. A. (1995). Microcracking during triaxial deformation of porous rocks monitored by changes in rock physical properties, I. Elastic-wave propagation measurements on dry rocks. *Tectonophysics*, 245(3–4), 205–221.
- Ben-Zion, Y. (2008). Collective behavior of earthquakes and faults: Continuum-discrete transitions, progressive evolutionary changes, and different dynamic regimes. *Reviews of Geophysics*, 46(4), 1–7.
- Ben-Zion, Y., & Zaliapin, I. (2019). Spatial variations of rock damage production by earthquakes in southern California. *Earth and Planetary Science Letters*, 512, 184–193.
- Ben-Zion, Y., & Zaliapin, I. (2020). Localization and coalescence of seismicity before large earthquakes. *Geophysical Journal International*, 223(1), 561–583.
- Bonamy, D., & Bouchaud, E. (2011). Failure of heterogeneous materials: A dynamic phase transition? *Physics Reports*, 498(1), 1–44.
- Brace, W. F., & Bombolakis, E. G. (1963). A note on brittle crack growth in compression. *Journal of Geophysical Research*, 68(12), 3709–3713.
- Buades, A., Coll, B., & Morel, J. M. (2005). A non-local algorithm for image denoising. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 2, pp. 60–65). *IEEE*.
- Bühlmann, P., & Yu, B. (2003). Boosting with the L 2 loss: Regression and classification. *Journal of the American Statistical Association*, 98(462), 324–339.
- Cartwright-Taylor, A., Main, I. G., Butler, I. B., Fousseis, F., Flynn, M., & King, A. (2020). Catastrophic failure: How and when? Insights from 4-D in situ X-ray microtomography. *Journal of Geophysical Research: Solid Earth*, 125(8), e2020JB019642.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In: *Proceedings of the 22nd acm sigkdd International Conference on knowledge discovery and data mining* (785–794).
- Coulomb, C. A. (1776). Essai Sur Une Application Des Règles Des Maximis et Minimis à Quelques Problèmes De Statique Relatifs à l'Architecture. *Mémoires de l'Académie Royale des Sciences*, 7, 343–382.
- Dahlen, F. (1984). Noncohesive critical Coulomb wedges: An exact solution. *Journal of Geophysical Research*, 89, 10125–10133. <https://doi.org/10.1029/JB089iB12p10125>
- Dahlen, F. A. (1990). Critical taper model of fold-and-thrust belts and accretionary wedges. *Annual Review of Earth and Planetary Sciences*, 18, 55–99.
- Dahlen, F. A., Suppe, J., & Davis, D. (1984). Mechanics of fold-and-thrust belts and accretionary wedges: Cohesive Coulomb theory. *Journal of Geophysical Research*, 89, 10087–10101. <https://doi.org/10.1029/JB089iB12p10087>
- Davis, D., Suppe, J., & Dahlen, F. (1983). Mechanics of fold-and-thrust belts and accretionary wedges. *Journal of Geophysical Research*, 88, 1153–1172. <https://doi.org/10.1029/JB088iB02p01153>
- Fredrich, J. T., Evans, B., & Wong, T. F. (1989). Micromechanics of the brittle to plastic transition in Carrara marble. *Journal of Geophysical Research: Solid Earth*, 94(B4), 4129–4145.
- Fredrich, J. T., & Wong, T. F. (1986). Micromechanics of thermally induced cracking in three crustal rocks. *Journal of Geophysical Research: Solid Earth*, 91(B12), 12743–12764.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 29, 1189–1232.
- Fousseis, F., Xiao, X., Schrank, C., & De Carlo, F. (2014). A brief guide to synchrotron radiation-based microtomography in (structural) geology and rock mechanics. *Journal of Structural Geology*, 65, 1–16.
- Graham, C. C., Stanchits, S., Main, I. G., & Dresen, G. (2010). Comparison of polarity and moment tensor inversion methods for source analysis of acoustic emission data. *International Journal of Rock Mechanics and Mining Sciences*, 47(1), 161.
- Griffith, A. A. (1924). Theory of rupture. In: *Proceedings of the First International Congress of Applied Mechanics*, Delft, pp. 55–63.
- Griggs, D. (1960). Deformation of rocks at 500 °C to 800 °C. *Geological Society of America Memoirs*, 79, 39–104.
- Hoek, E., & Bieniawski, Z. T. (1965). Brittle fracture propagation in rock under compression. *International Journal of Fracture Mechanics*, 1(3), 137–155.
- Horii, H., & Nemat-Nasser, S. (1985). Compression-induced microcrack growth in brittle solids: Axial splitting and shear failure. *Journal of Geophysical Research: Solid Earth*, 90(B4), 3105–3125.
- Howarth, D. F. (1987). The effect of pre-existing microcavities on mechanical rock performance in sedimentary and crystalline rocks. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 24, 223–233.
- Huiqi, L., McClay, K. R., & Powell, D. (1992). Physical models of thrust wedges. *Thrust Tectonics*, 71, 8. <https://doi.org/10.1007/978-94-011-3066-0>
- 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.
- Inglis, C. E. (1913). Stresses in a plate due to the presence of cracks and sharp corners. *Trans Inst Naval Archit*, 55, 219–241.
- Isida, M. (1971). Effect of width and length on stress intensity factors of internally cracked plates under various boundary conditions. *International Journal of Fracture Mechanics*, 7(3), 301–316.
- Jaeger, J. C., Cook, N. G. W., & Zimmerman, R. (1979). Rock mechanics. *Fundamentals of Rock Mechanics*.
- Kandula, N., Cordonnier, B., Boller, E., Weiss, J., Dysthe, D. K., & Renard, F. (2019). Dynamics of microscale precursors during brittle compressive failure in Carrara marble. *Journal of Geophysical Research: Solid Earth*, 124(6), 6121–6139.
- Kato, A., & Ben-Zion, Y. (2021). The generation of large earthquakes. *Nature Reviews Earth & Environment*, 2, 26–39. <https://doi.org/10.1038/s43017-020-00108-w>
- Kurzon, I., Lyakhovskiy, V., & Ben-Zion, Y. (2019). Dynamic rupture and seismic radiation in a damage-breakage rheology model. *Pure and Applied Geophysics*, 176, 1003–1020.
- Lallemand, S., Schnrle, P., & Malavieille, J. (1994). Coulomb theory applied to accretionary and nonaccretionary wedges: Possible causes for tectonic erosion and/or frontal accretion. *Journal of Geophysical Research*, 99, 12033–12055. <https://doi.org/10.1029/94JB00124>

- Lockner, D., Byerlee, J. D., Kuksenko, V., Ponomarev, A., & Sidorin, A. (1991). Quasi-static fault growth and shear fracture energy in granite. *Nature*, *350*(6313), 39–42.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In: *Advances in neural information processing systems* (pp. 4765–4774).
- Lyakhovskiy, V., Ben-Zion, Y., & Agnon, A. (2001). Earthquake cycle, fault zones, and seismicity patterns in a rheologically layered lithosphere. *Journal of Geophysical Research: Solid Earth*, *106*(B3), 4103–4120.
- Malaga-Starzec, K., Lindqvist, J. E., & Schouenborg, B. (2002). Experimental study on the variation in porosity of marble as a function of temperature. *Geological Society, London, Special Publications*, *205*(1), 81–88. <https://doi.org/10.1144/GSL.SP.2002.205.01.07>
- McBeck, J. A., Aiken, J. M., Mathiesen, J., Ben-Zion, Y., & Renard, F. (2020). Deformation precursors to catastrophic failure in rocks. *Geophysical Research Letters*, *47*, e2020GL090255.
- McBeck, J., Cooke, M., & Madden, E. (2017). Work optimization predicts the evolution of extensional step overs within anisotropic host rock: Implications for the San Pablo Bay, CA. *Tectonics*, *36*(11), 2630–2646.
- McBeck, J., Kandula, N., Aiken, J. M., Cordonnier, B., & Renard, F. (2019). Isolating the factors that govern fracture development in rocks throughout dynamic in situ X-ray tomography experiments. *Geophysical Research Letters*, *46*(20), 11127–11135.
- Moore, D. E., & Lockner, D. A. (1995). The role of microcracking in shear-fracture propagation in granite. *Journal of Structural Geology*, *17*(1), 95–114.
- Morgan, S. P., Johnson, C. A., & Einstein, H. H. (2013). Cracking processes in Barre granite: Fracture process zones and crack coalescence. *International Journal of Fracture*, *180*(2), 177–204.
- Mulugeta, G. (1988). Modelling the geometry of Coulomb thrust wedges. *Journal of Structural Geology*, *10*, 847–859.
- Nemat-Nasser, S., & Horii, H. (1982). Compression-induced non-planar crack extension with application to splitting, exfoliation, and rockburst. *Journal of Geophysical Research: Solid Earth*, *87*(B8), 6805–6821.
- Nur, A. (1974). Matsushiro, Japan, earthquake swarm: Confirmation of the dilatancy-fluid diffusion model. *Geology*, *2*(5), 217–221.
- Park, C. H., & Bobet, A. (2009). Crack coalescence in specimens with open and closed flaws: A comparison. *International Journal of Rock Mechanics and Mining Sciences*, *46*(5), 819–829.
- Paterson, M. S., & Wong, T. F. (2005). *Experimental rock deformation—the brittle field*. Springer Science & Business Media.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, *12*, 2825–2830.
- Peng, S., & Johnson, A. M. (1972). Crack growth and faulting in cylindrical specimens of Chelmsford granite. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *9*(1), 37–86.
- Popp, T., Kern, H., & Schulze, O. (2001). Evolution of dilatancy and permeability in rock salt during hydrostatic compaction and triaxial deformation. *Journal of Geophysical Research: Solid Earth*, *106*(B3), 4061–4078.
- Renard, F. (2017). Critical evolution of damage towards system size failure in a crystalline rock. *Norstore*. <https://doi.org/10.11582/2017.00025>
- Renard, F. (2018). Volumetric and shear processes in crystalline rock during the approach to faulting. *Norstore*. <https://doi.org/10.11582/2018.00023>
- Renard, F. (2019). Dynamics of microscale precursors establish brittle compressive failure in Carrara marble as a critical phenomenon. *Norstore*. <https://doi.org/10.11582/2019.00001>
- Renard, F. (2021). X-ray tomography data of Westerley granite. *Norstore*. <https://doi.org/10.11582/2021.00002>
- Renard, F., Cordonnier, B., Dysthe, D. K., Boller, E., Tafforeau, P., & Rack, A. (2016). A deformation rig for synchrotron microtomography studies of geomaterials under conditions down to 10 km depth in the Earth. *Journal of Synchrotron Radiation*, *23*(4), 1030–1034.
- Renard, F., McBeck, J., Kandula, N., Cordonnier, B., Meakin, P., & Ben-Zion, Y. (2019). Volumetric and shear processes in crystalline rock approaching faulting. *Proceedings of the National Academy of Sciences*, *116*(33), 16234–16239.
- Renard, F., Weiss, J., Mathiesen, J., Ben-Zion, Y., Kandula, N., & Cordonnier, B. (2018). Critical evolution of damage toward system-size failure in crystalline rock. *Journal of Geophysical Research: Solid Earth*, *123*(2), 1969–1986.
- Rutter, E. H. (1972). The influence of interstitial water on the rheological behaviour of calcite rocks. *Tectonophysics*, *14*(1), 13–33. [https://doi.org/10.1016/0040-1951\(72\)90003-0](https://doi.org/10.1016/0040-1951(72)90003-0)
- Shreedharan, S., Bolton, D. C., Rivière, J., & Marone, C. (2021). Machine Learning Predicts the Timing and Shear Stress Evolution of Lab Earthquakes Using Active Seismic Monitoring of Fault Zone Processes. *Journal of Geophysical Research: Solid Earth*, *126*(7), e2020JB21588.
- Stanchits, S., Vinciguerra, S., & Dresen, G. (2006). Ultrasonic velocities, acoustic emission characteristics and crack damage of basalt and granite. *Pure and Applied Geophysics*, *163*(5–6), 975–994.
- Tal, Y., Evans, B., & Mok, U. (2016). Direct observations of damage during unconfined brittle failure of Carrara marble. *Journal of Geophysical Research: Solid Earth*, *121*(3), 1584–1609.
- Tapponnier, P., & Brace, W. F. (1976). Development of stress-induced microcracks in Westerly granite. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, *13*(4), 103–112.
- Tullis, J., & Yund, R. A. (1977). Experimental deformation of dry Westerly granite. *Journal of Geophysical Research*, *82*(36), 5705–5718.
- Turner, F. J., Griggs, D. T., & Heard, H. (1954). Experimental deformation of calcite crystals. *Geological Society of America Bulletin*, *65*(9), 883–934.
- Vasseur, J., Wadsworth, F. B., Heap, M. J., Main, I. G., Lavallée, Y., & Dingwell, D. B. (2017). Does an inter-flaw length control the accuracy of rupture forecasting in geological materials? *Earth and Planetary Science Letters*, *475*, 181–189.
- Vermeer, P. A. (1990). The orientation of shear bands in biaxial tests. *Geotechnique*, *40*(2), 223–236.
- Walton, G., Arzúa, J., Alejano, L. R., & Diederichs, M. S. (2015). A laboratory-testing-based study on the strength, deformability, and dilatancy of carbonate rocks at low confinement. *Rock Mechanics and Rock Engineering*, *48*(3), 941–958.

Wang, W., Kravchenko, A. N., Smucker, A. J. M., & Rivers, M. L. (2011). Comparison of image segmentation methods in simulated 2D and 3D microtomographic images of soil aggregates. *Geoderma*, 162(3–4), 231–241.

Wong, R. H. C., Chau, K. T., Tang, C. A., & Lin, P. (2001). Analysis of crack coalescence in rock-like materials containing three flaws—Part I: Experimental approach. *International Journal of Rock Mechanics and Mining Sciences*, 38(7), 909–924.

(Received July 7, 2021, revised September 30, 2021, accepted November 5, 2021, Published online November 16, 2021)