



Research paper

Enhancing fracture network characterization: A data-driven, outcrop-based analysis

Weiwei Zhu ^a, Xupeng He ^b, Ryan Kurniawan Santoso ^c, Gang Lei ^d, Tadeusz W. Patzek ^b, Moran Wang ^{a,*}

^a Department of Engineering Mechanics, Tsinghua University, Beijing, China

^b Ali I. Al-Naimi Petroleum Engineering Research Center (ANPERC), King Abdullah University of Science and Technology, Thuwal, Kingdom of Saudi Arabia

^c RWTH Aachen University, Aachen, Germany

^d Faculty of Engineering, China University of Geosciences, Wuhan, China



ARTICLE INFO

Dataset link: <https://doi.org/10.4121/14865096.v2>

Keywords:

Fracture
Characterization
Data-driven
Outcrop-based
Stochastic discrete fracture network

ABSTRACT

We utilize a pixel-based fracture detection algorithm to digitize 80 published outcrop maps of different scales at different locations. The key fracture properties, including fracture lengths, orientations, intensities, topological structures, clusters, and flow, are analyzed. Our findings provide significant justifications for statistical distributions used in SDFN modelings. We find that fracture lengths follow multiple (instead of single) power-law distributions with varying exponents. Large fractures tend to have large exponents, possibly because of a small coalescence probability. Most small-scale natural fracture networks have scattered orientations, corresponding to a small κ value ($\kappa < 3$) in a von Mises–Fisher distribution. Large fracture systems analyzed in this research usually have more concentrated orientations with large κ values. Fracture intensities are spatially clustered at all scales. A fractal spatial density distribution, which introduces clustered fracture positions, can better capture the spatial clustering than a uniform distribution. Natural fracture networks usually have a significant proportion of T-type nodes, which is unavailable in conventional SDFN models. Thus, a rule-based algorithm is necessary to mimic the fracture growth and form T-type nodes. Most outcrop maps show good topological connectivity. However, sealing patterns and stress impact must be considered to evaluate the hydraulic conductivity of fracture networks.

1. Introduction

As a general concept of discontinuities in geology, fractures include joints, faults, veins, deformation bands, and pressure solution seams. They are pervasive in the upper crust and usually form complex networks. Fracture networks can affect many important mechanical and hydraulic properties of rocks, including rock strength, stiffness, and permeability (Adler and Thovert, 1999). Therefore, their impacts on many engineering fields are significant, such as underground water movement, nuclear waste disposal, mining, geothermal extraction, and petroleum reservoir development (Berkowitz, 2002; Pogacnik et al., 2016).

However, the three-dimensional, small-scale structures of subsurface fractures are largely inaccessible with current technologies (seismic surveys, well logging, field observations, etc.). Furthermore, complex and irregular fracture shapes (Gertsch, 1995), rough fracture surfaces (Chae et al., 2004), fracture tortuosity, and stress impact on the fracture permeability (Cook et al., 1990) make it extremely difficult to

describe fracture networks in great detail. A stochastic discrete fracture network model (SDFN) is a practical approach to modeling complex fracture systems in the subsurface, where fractures are represented explicitly with simplistic shapes, such as line segments in 2D and polygons or ellipses in 3D (Bour and Davy, 1997; Darcel et al., 2003; Lei et al., 2017; Li et al., 2022).

The key geometric properties of the fracture network, such as fracture orientations, sizes, and center positions, are described with different statistical distributions. Different researchers propose log-normal, gamma, exponential, and power-law distributions to describe fracture lengths (Bonnet et al., 2001). The power-law distribution is prevalent as observed from field outcrops and analog experiments (Sornette et al., 1993). A von Mises–Fisher distribution is widely adopted to describe fracture orientations (Huang et al., 2020). Fracture center positions are described by a Poisson (uniform) or a more general fractal spatial density distribution (Bour and Davy, 1997; Darcel et al., 2003). A fractal dimension characterizes the fractal spatial density

* Corresponding author.

E-mail address: mrwang@tsinghua.edu.cn (M. Wang).

Table 1
Distributions of different fracture geometric properties.

Property	Distribution	Reference
Length	Exponential	Priest and Hudson (1976),
	Gamma	Nur (1982)
	Lognormal	Davy (1993), Sornette and Sornette (1999)
	Power-law	Priest and Hudson (1981), Rouleau and Gale (1985), Segall and Pollard (1983a), Sornette et al. (1993)
Orientation	von Mises–Fisher distribution	Song et al. (2001), Kemeny and Post (2003)
Position	Poisson	Berkowitz (1995), Bour and Davy (1997)
	Fractal	Darcel et al. (2003), Zhu et al. (2018)

distribution. In 2D, the fractal spatial density distribution generates clustered fracture positions and brings clustering effects when the fractal dimension is smaller than 2.0. It reduces to a uniform distribution when the fractal dimension equals 2.0. Correspondingly, the critical dimension for uniformly distributed fracture centers is 3.0 in 3D space. A uniform spatial density distribution is easy to implement but usually unrealistic. Through outcrop observations, natural fracture systems usually show clustering effects (Zhu et al., 2018; Akara et al., 2021). Table 1 summarizes key geometric properties of fracture networks and their commonly adopted statistical distributions.

An SDFN model cannot characterize details of fracture geometry, but it can preserve the essential topological structure of a fracture network, which determines the overall hydraulic diffusivity in fluid flow through low permeability formations (Lei et al., 2017; Zhu et al., 2021c). Important flow results, including flow rate or permeability, depend on the configuration of fracture networks (Zhu et al., 2021c). Therefore, it is necessary to investigate fracture configurations before incorporating impacts of detailed fracture properties, such as roughness, curved shapes, and stress impacts. However, actual subsurface fracture networks are unavailable with current technologies. Therefore, it is impossible to validate the correctness and quality of SDFN models. The confidence of SDFN models depends on whether statistical distributions are representative of real fracture networks. Those statistical distributions are usually summarized from available datasets, including outcrop maps, wellbore images, and seismic maps (Feng et al., 2020; Williams and Johnson, 2004; Prioul and Jocker, 2009). Outcrop maps are essential datasets because they are widely spread and provide abundant information on the geometric properties of fractures, such as fracture orientations, lengths, and intersection relationships. However, outcrop maps require a significant human effort to recognize and detect fractures before summarizing statistics. Geologists usually analyze a few outcrop maps at a given region and hardly extend their findings to more general conditions, considering the tremendous time cost for fracture interpretations. Therefore, a synthetic analysis of fracture geometries with a large number of outcrop maps from different scales and locations is unavailable.

To reduce the human effort in interpreting outcrop maps, we have proposed an automatic fracture interpretation algorithm (Zhu et al., 2020), which automatically interprets typical plan-view maps of fracture networks from a variety of resources, such as seismic reflection horizons, satellite images, aerial photos, etc. The method comprises two stages: (1) converting a raw outcrop image to a binary map that preserves the fracture trace and removes all background features. This stage is denoted as fracture recognition, which is completed by using a deep-learning architecture, U-net (Ronneberger et al., 2015). (2) replacing the binary fracture images with line segments or polylines. This stage is named fracture detection, completed with a pixel-based fracture detection algorithm. The algorithm is further optimized in this

research and discussed in detail in the following section. The deep-learning-based fracture recognition needs many training images, which are unavailable in this research. Therefore, we focus on the fracture detection algorithm and utilize the algorithm to digitize 80 outcrop maps from published literature.

Commonly used lineament detection methods include Hough transform (Wang and Howarth, 1990), Segment Tracing Algorithm (Koike et al., 1995) and more methods are available in a detailed review paper (Ahmadi and Pekkan, 2021). Those lineament detection methods are usually sensitive to curved shapes of real fracture traces, thus unsuitable for complex fracture systems. Our proposed pixel-based detection algorithm instead is robust, accurate, and efficient. All pixels in a fracture trace are recorded, and key information about the fracture trace is available, including fracture orientations, lengths, positions, and the abutment relationship between fractures.

The outcrop maps are collected from different literature at different locations with varying scales from millimeters to tens of kilometers. We synthetically analyze distributions of fracture lengths, orientations, intensities, topological structures, clusters, and flow from those outcrop maps. The findings on fracture geometries provide significant support and justifications for SDFN modeling, and we also point out shortcomings and possible improvements of the commonly adopted SDFN techniques. The analysis of clusters and flow investigates the connected fractures and their permeability, considering the fracture sealing and stress impacts.

The structure of the paper is organized as follows: Section 2 introduces detailed algorithms used in detecting fractures from binary outcrop maps. Section 3 analyzes distributions of fracture lengths, orientations, intensities, their topological structures, clusters and flow. Section 4 discusses insights for SDFN modeling. Important findings and conclusions are summarized in Section 5.

2. Fracture detection

Fracture recognition and fracture detection are the two significant stages for a successful fracture interpretation (Zhu et al., 2020). Deep learning techniques, such as U-net (Ronneberger et al., 2015), and mathematic methods, such as shearlet transform (Reisenhofer, 2014), are suitable for fracture recognition and separate fracture geometries from complex environments. This research collects published binary outcrop maps recognized by different geologists; therefore, the fracture recognition process is considered complete. A fracture detection process is further required to convert binary outcrop maps to line segments or polylines for further in-depth investigations. In the proposed approach, we adopt a pixel-based fracture detection method that is robust, accurate, and efficient. The method is composed of four main steps.

Step 1 Skeletonize a binary image to a one-pixel wide image. ‘bwskel’ function in MATLAB can conveniently convert a binary image to its skeleton. Fig. 1(b) shows an example of the skeletonization. Outcrops have experienced severe weathering and stress-release during the upward movement, which significantly changes fracture apertures. Therefore, the aperture information from an outcrop map is generally unreliable. The skeletonization loses the aperture information but preserves the topological structure of the fracture network, which is essential for the flow behavior in the subsurface.

Step 2 Find initial pixels and intersection pixels of fractures in the skeletonized image. Check each pixel and count the number of its neighboring pixels, N_b . If $N_b = 1$, the pixel is denoted as an initial pixel. If $N_b \geq 3$, the pixel is defined as an intersection pixel. If $N_b = 2$, the corresponding pixel is a transit pixel. Fig. 1(c) shows a sketch map of different types of pixels. The green squares and the blue circles represent initial and intersection pixels, respectively. Several intersection pixels can coexist at one intersection area. Therefore, blue adjacent intersection pixels are merged to their centroid point, referring to the merged intersection pixel (red triangle).

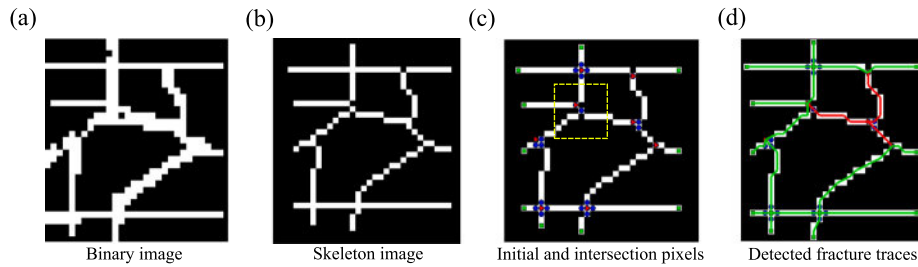


Fig. 1. (a) A binary fracture image; (b) Skeletonization of the binary fracture image in (a); (c) Initial and intersection pixels of the skeletonized image. Green squares represent initial pixels. Blue circles represent intersection pixels. Red triangles represent merged intersection pixels. (d) Fracture traces interpreted by the detection method. Type 1 and 2 fractures found in Steps 3 and 4 are shown in green and red, respectively.

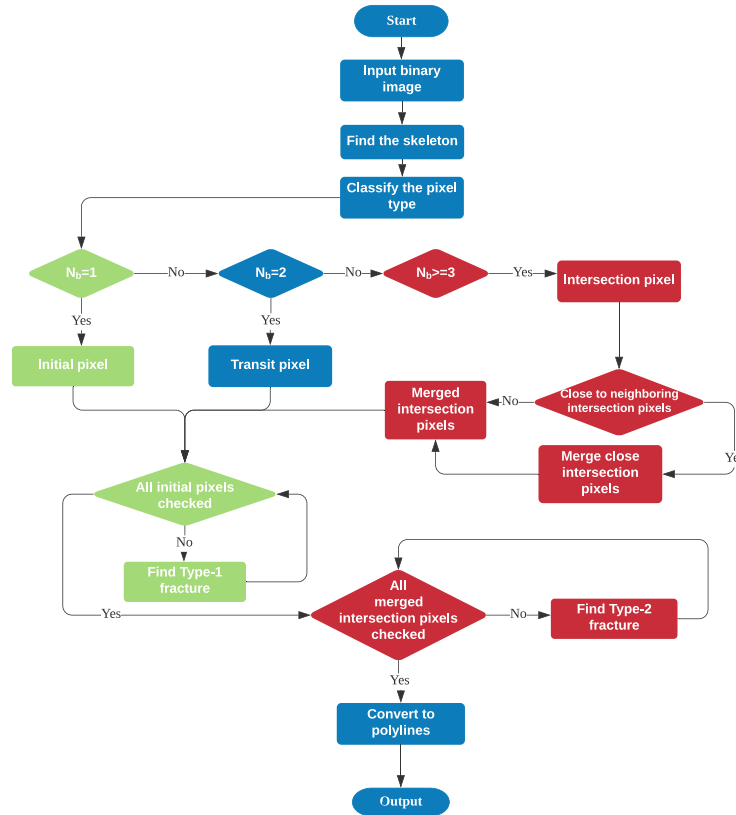


Fig. 2. Flow chart of the fracture detection algorithm. N_b is the number of neighboring pixels. The green and red colors refer to steps involving initial pixels and intersection pixels, respectively; The blue color refers to ordinary steps and the recognition of transit pixels.

Step 3 Define the Type 1 fracture and track the fracture trace. The Type 1 fracture is constrained by at least one initial pixel, either by a pair of initial pixels or an initial pixel and a merged intersection pixel. The tracking process visits all initial pixels. The trace tracking starts at an unvisited initial pixel and records the transit pixels in the fracture trace. When the trace encounters the other initial pixel or an isolated intersection pixel (after which no valid subsequent pixel exists), the tracking stops for the corresponding fracture trace and restarts the tracking with another unvisited initial pixel. The pixels in a specific fracture trace, including the initial, transit, and intersection pixels, are recorded. Fig. 1(d) shows the results of Type 1 fractures, marked in green.

Step 4 Define the Type 2 fracture and track the fracture trace. A pair of merged intersection pixels compose the start and end points of a Type 2 fracture trace. The trace tracking process is implemented on all merged intersection pixels. The tracking starts at a merged intersection pixel and records the transit pixels in the fracture trace. When the trace encounters the

first merged intersection pixel, the tracking process stops for the corresponding fracture trace and restarts the tracking with another unvisited merged intersection pixel. Fig. 1(d) shows the results of Type 2 fractures, marked in red.

Fig. 2 presents a flow chart of the fracture detection algorithm for a clear demonstration. The fracture curvature can be captured since all pixels in a specific fracture trace are recorded. Fig. 1(d) shows curved fracture traces with polylines. The number of segments in a polyline is controllable, and the maximum value is the number of pixels in the fracture trace. Finding the correct pixel when encountering an intersection pixel is the most troublesome part of trace tracking in Step 3. Two steps are proposed to find the correct pixel to solve this issue.

- Find all possible pixels around the closest merged intersection pixel. Possible pixels are the pixels intersecting the yellow box in Fig. 1(c). The size of the yellow box is dynamically adjusted from 1 to 3-pixel lengths.
- Find the pixel which best fits the trace trend. The pixel is recorded as the subsequent pixel if its deviation is insignificant. Otherwise,

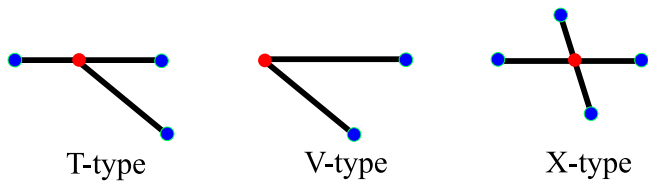


Fig. 3. Demonstration of T-type, V-type and X-type intersections. Blue points refer to the I-type nodes.

the closest merged intersection pixel is recorded, and the tracking process stops for the corresponding trace. In this case, the trace is constrained by an initial pixel and a merged intersection pixel. The current intersection pixel is the isolated intersection pixel mentioned in **Step 3**. The deviation criterion is adjustable for different outcrop maps. $\pi/6$ is a good choice for most cases as a deviation criterion to distinguish the pixel orientation and the averaged trace orientation.

The process of finding the next pixel can also be used to find branches originating from one merged intersection pixel, where the number of branches defines the T-type and X-type intersections, as demonstrated in Fig. 3. If the number of branches is two or three, the corresponding merged intersection pixel is a T-type intersection. When the number of branches is two, the merged intersection pixel is a V-type intersection, where two tips coincide. However, we do not distinguish these two types and regard both of them as T-type intersections because the probability of a V-type intersection occurring in a natural network is negligible (Sanderson and Nixon, 2015). When the number of branches is more than three, the corresponding merged intersection pixel is an X-type intersection. I-type (isolated fractures), T-type, and X-type intersections are used for topological analysis in the next section.

Fig. 4 shows one example of the digitized outcrop maps from the Achnashellach Culmination field area (Figs. 7B and 7D in Watkins et al., 2015). Detected fracture traces are overlapped with the original outcrop map in Fig. 4(c), which shows accurate detection results. The synthetic analysis of geometric properties and topological structures of fractures are obtained with digitized outcrop maps and presented in the following section.

3. Analysis of fracture geometries, topological structures, clusters and flow

We implement the fracture detection algorithm on 80 published outcrop maps (Segall and Pollard, 1983b; Gillespie et al., 1993; Barton, 1995; Odling, 1997; Odling et al., 1999; Holland et al., 2009; Jafari, 2011; Bertrand et al., 2015; Watkins et al., 2015; Bisdom, 2016; Duffy et al., 2017; Healy et al., 2017; Thiele et al., 2017; Becker et al., 2018; Wyller, 2019; Prabhakaran et al., 2021). These outcrop maps are collected from different parts of the world as shown in Fig. 5 and they have a wide range of scales from millimeters to tens of kilometers. Their geometric patterns are summarized and analyzed in detail, including fracture lengths, orientations, intensities, and topological structures. All digitized outcrop maps are available online (<https://doi.org/10.4121/14865096.v2>).

It is worthwhile to mention that published outcrop maps are usually biased because regions with well-developed fractures are preferred for sampling, indicating that the analysis below cannot represent a complete picture of natural fracture networks. However, well-developed fractures are also essential for fluid transportation in the subsurface compared with poorly-developed fractures. Therefore, the analysis based on collected outcrop maps is still helpful in deepening the understanding of natural fracture systems.

3.1. Length distribution

Exponential, log-normal, and power-law distributions are usually adopted to describe fracture length, and the power-law distribution is prevalent. Cladouhos and Marrett (1996) and Olson (2007) simulated the fracture growth and linkage process and observed a power-law distribution of fracture lengths. Researchers also observed two-/three dimensional power-law distribution in outcrop maps (Davy, 1993), where the full-length distributions are separated into two/three regions and described with a power-law distribution with different exponents. Sano et al. (1981) derived the power-law distribution of fracture length based on the assumption of homogeneous materials and uni-axial stress conditions. However, the origin of power-law distribution is still an unsolved issue, and in-depth explanations are insufficient.

Makarov (2007) found the universal fractality of solids through the process of destruction in loaded solids, which leads brittle fracture and plastic deformation to be self-similar at different scales, and in turn, their lengths follow a power-law distribution. However, many experiments and outcrop observations show the opposite process, where large fractures are formed from the linkage and coalescence of numerous small, isolated fractures, and such linkage and coalescence happen in a wide range of scales (mm–km) (Cartwright et al., 1995; Otsuki and Dilov, 2005). Here, we propose a simplistic but elegant model (Appendices A and B) to explain a possible origin of a power-law distribution for fracture lengths based on the self-similar characteristics of natural fractures and their coalescence. The model reproduces the iterative growth process of natural fractures, and we conclude that the fracture lengths should follow a power-law distribution at a later generation:

$$\left(\frac{N_N}{(1-p)N_0}\right) = \left(\frac{l_N}{l_0}\right)^{\frac{\ln(p)}{\ln(n_s)}-1}, \quad (1)$$

where N_N is the number of fractures with a fracture length of l_N at the N th generation; N_0 is the initial number of fractures; l_N is the fracture length at the N th generation; l_0 is the initial length of fractures; p is the coalescence probability at each generation; n_s is the number of small fractures coalesced to form a large fracture. A simpler expression of Eq. (1) is:

$$N_N \sim l_N^{-a}, \quad (2)$$

where $a = 1 - \frac{\ln(p)}{\ln(n_s)}$.

The coalescence probability varies between 0 and 1.0 and the number of coalesced small fractures n_s usually varies between 2 and 5 (Cartwright et al., 1995; Otsuki and Dilov, 2005), so $\frac{\ln(p)}{\ln(n_s)}$ term is always negative. The corresponding exponent should always be larger than 1.0, which is consistent with most outcrop observations (Bonnet et al., 2001).

If the length distribution follows a power-law distribution, the cumulative distribution function (CDF) of fracture lengths should also follow a power-law distribution:

$$C_N \sim l_N^{1-a} \quad (3)$$

If p and n_s are constant for each generation, the cumulative length distribution should be a single power-law distribution with the exponent equal to $1 - a$. However, this scenario is over-idealized, and a typical length distribution from the real outcrop map at the Achnashellach Culmination field area (Fig. 7B in Watkins et al., 2015) is shown in Fig. 6. For clarity, the fracture length is represented by the number of pixels of each fracture trace instead of line segments after converting a fracture trace to polylines.

Fig. 6(a) shows a curved CDF with varying slopes. A single or two-/three-dimensional power-law distribution is insufficient to describe the length distribution. Suppose that we regard each short growth period as a straight line segment. In that case, the varying slopes indicate fracture lengths in different short segments following a power-law distribution with different values of exponents. The variation of exponents comes from variations of p and n_s .

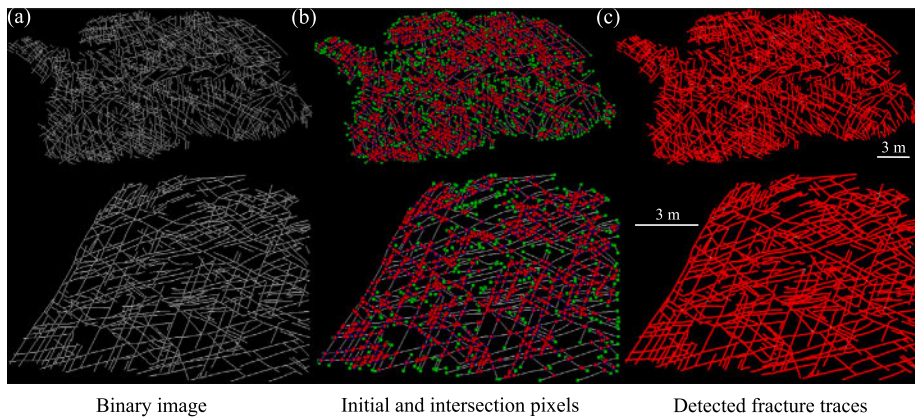


Fig. 4. Digitized fracture outcrop map at Achnashellach Culmination field area (Upper panel: Fig. 7B and Lower panel: Fig. 7D in Watkins et al., 2015).

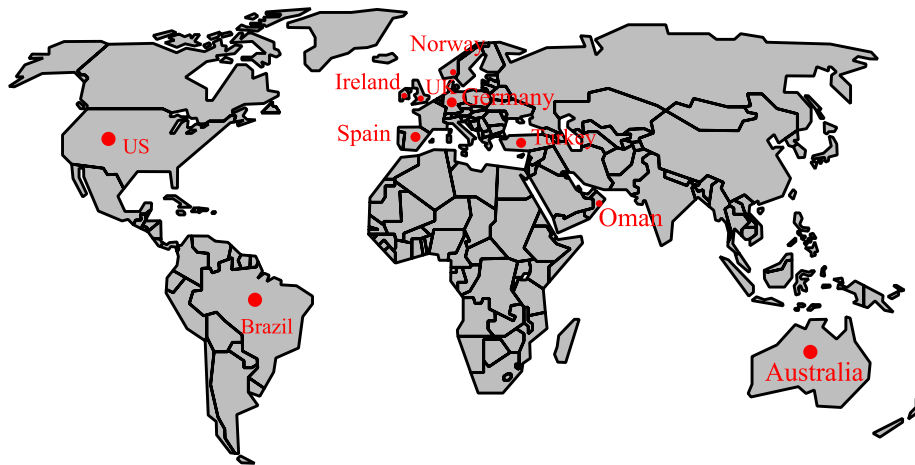


Fig. 5. A world map showing locations of collected outcrop maps.

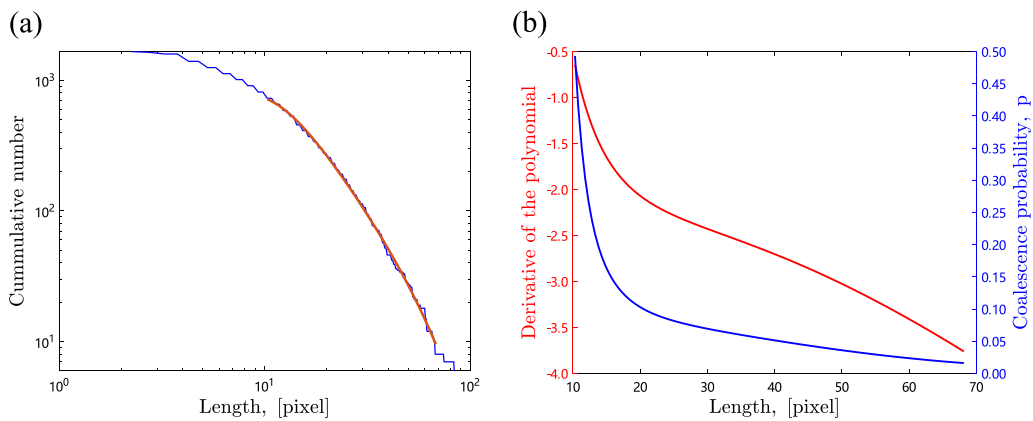


Fig. 6. (a) Length distribution of outcrop maps from Achnashellach Culmination field area (Fig. 7B in Watkins et al., 2015). The red curve is the result of a fourth-order polynomial fitting. (b) The red curve refers to the derivative of the fourth-order polynomial fitting in (a). The blue curve refers to the derived coalescence probability based on Eq. (4).

Due to the finite range effect (Pickering et al., 1995), the exponent obtained from CDF fitting is biased. We adopt the iterative method introduced in Pickering et al. (1995) with 1 ~ 3 iterations to remove the majority of bias. The number of iterations is case-dependent, and usually, three iterations is a good option. An intermediate range of fracture lengths is chosen for the fitting because data points of large and small fracture lengths are inaccurate. To be specific, [0.1 0.8] of the full range is selected. Large fracture trace length is usually affected by censoring effects (Riley, 2005), where the trace length is less than or equal to that of an entire trace. Small fracture trace length is

inaccurate because of two possible reasons. One is incomplete sampling or truncation because of limited resolutions of the sampling method. The second one is the misinterpretation of close and small fractures caused by the limited resolution of published outcrop maps. Therefore, an intermediate range of lengths is more appropriate for the fitting.

We fit the cumulative length distribution with a fourth-degree polynomial and get their derivatives, corresponding to $1 - a$ at each short growth period. The slope of fitted polynomials is shown in Fig. 6(b). With fracture lengths increasing, the slope gets smaller, which corresponds to a smaller value of $\ln(p)/\ln(n_s)$. If we take an intermediate

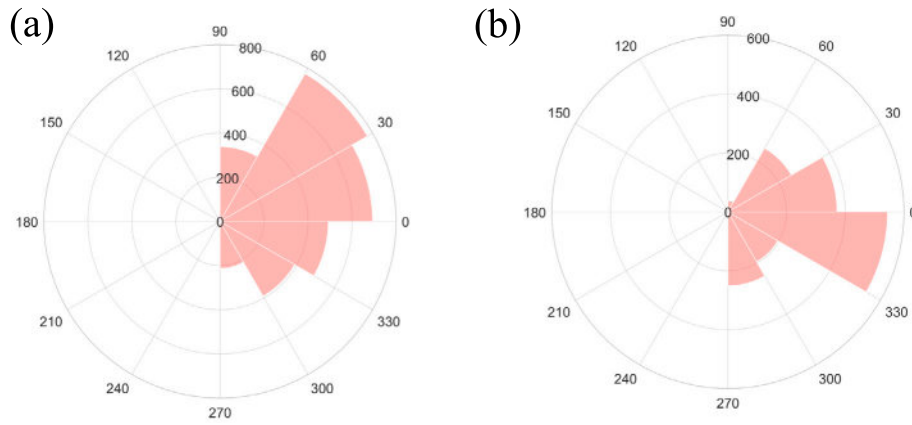


Fig. 7. Orientation distribution of outcrop maps at Achnashellach Culmination field area (Figs. 7B and 7D in Watkins et al., 2015).

value of n_s ($n_s = 3$) and keep constant, Eq. (4) is the formula to calculate the coalescence probability at different length scales, and the corresponding values of p are shown in Fig. 6(b).

$$p = \exp(k \ln n_s), \tag{4}$$

where k is the slope of the cumulative length distribution.

The coalescence probability decreases with increasing fracture lengths. This observation is valid in 51 out of 80 collected outcrop maps, indicating that large fractures are less likely to merge and form a larger fracture. The decreasing coalescence probability may attribute to the fact that relative fracture intensity is small in large fracture systems, and the stress condition required for the coalescence of large fractures is tough.

The simplistic model is inaccurate since the natural fracturing process is extremely complex, considering the heterogeneous materials and complex stress conditions. However, it provides a straightforward and possible scenario to explain the origin of the power-law distribution by capturing essential patterns in natural fractures (self-similarity and coalescence). This model can be further improved by incorporating concrete physical constraints.

3.2. Orientation distribution

A von Mises–Fisher distribution usually describes the orientation of fractures (Song et al., 2001). For a D-dimensional orientation vector \vec{x} , the probability distribution function is:

$$p(\vec{x} | \vec{\mu}, \kappa) = C_D \exp(\kappa \vec{\mu}^T \vec{x}), \tag{5}$$

where $C_D(\kappa)$ is

$$C_D(\kappa) = \frac{\kappa^{D/2-1}}{(2\pi)^{D/2} I_{D/2-1}(\kappa)}, \tag{6}$$

where I_ν is the modified Bessel function of the first kind at the order of ν ; $\vec{\mu}$ is the mean direction, and κ is concentration parameter, which controls the concentration degree of the orientation with respect to $\vec{\mu}$. If $\kappa = 0$, the distribution will reduce to a uniform distribution. If κ is large, the distribution will be close to a Gaussian distribution with $\vec{\mu}$ as the mean and $1/\kappa$ as the variance.

Here, we focus on the two-dimensional outcrop maps. Fig. 7 shows rose diagrams of two outcrop maps at the Achnashellach Culmination field area (Figs. 7B and 7D in Watkins et al., 2015). Fractures with different orientations possibly belong to different fracture sets because of different stress states during the geologic history (Laubach, 1988). In each fracture set, the orientation is highly concentrated and usually has a large κ value (Kemeny and Post, 2003). To distinguish fracture sets on an outcrop map is nontrivial because the fracture orientation is only one of the important factors in distinguishing different fracture

sets, and the abutting and overprinting criteria between fractures are essential as well (Weismüller et al., 2020). More importantly, the stress history, fracture configurations, and stress relief/shadowing effects in the subsurface are usually inaccessible. Therefore, we regard each outcrop map as an integrated fracture system instead of investigating each fracture set. The corresponding κ values for outcrop maps in Fig. 7 are 2.3 and 2.9, respectively. The small value of κ makes the orientation distribution close to a uniform distribution, which is a widely used assumption in many SDFN modeling cases (Bour and Davy, 1997; Zhu et al., 2018).

Summarized κ value for all 80 outcrop maps are presented versus the outcrop scale in double-logarithmic coordinates in Fig. 8(a). The collected outcrop maps have their size similar regarding the number of pixels in the longest distance, therefore, the resolution (m/pixel) is chosen to represent the scale of the outcrop map. Most small-scale outcrop maps have their orientations scattered and yield a small κ value ($\kappa < 3$), which indicates that fracture systems may be composed of many fracture sets with different orientations. Large outcrops tend to have more concentrated orientations, and the largest outcrop has the most concentrated orientation with $\kappa = 100.2$. The correlation coefficient between the outcrop scale and κ in double-logarithmic coordinates is 0.48, implying a positive correlation between these two parameters. However, the positive correlation is mainly caused by four anomalous data points (large-scale faults), which are marked in green. Fig. 8(b) shows κ values after removing four anomalous data points, and the correlation coefficient is close to zero. Usually, it is difficult to collect outcrop maps at large scales, which makes outcrop maps of large faults insufficient. With available datasets in this research, large faults tend to have concentrated orientations. If this trend is valid in reality, it can partially explain the small coalescence probability of large fractures because they are concentrated in orientations and difficult to intersect each other.

3.3. Fracture intensity

Natural fractures are usually regarded as spatially clustered instead of uniformly distributed (Zhu et al., 2018; Akara et al., 2021). Commonly used methods to measure the spatial clustering of fracture networks include one-dimensional sampling, which measures the spacing between fractures in a given fracture set, or two-dimensional sampling, which maps fracture traces exposed on the outcrop (Gillespie et al., 1993). Here we adopt the two-dimensional sampling. We divide outcrop maps into small boxes and calculate the fracture intensity in different boxes to measure the spatial clustering of the fracture network. The box intensity is defined as:

$$P_{21}^i = \frac{L_i}{A_i}, \tag{7}$$

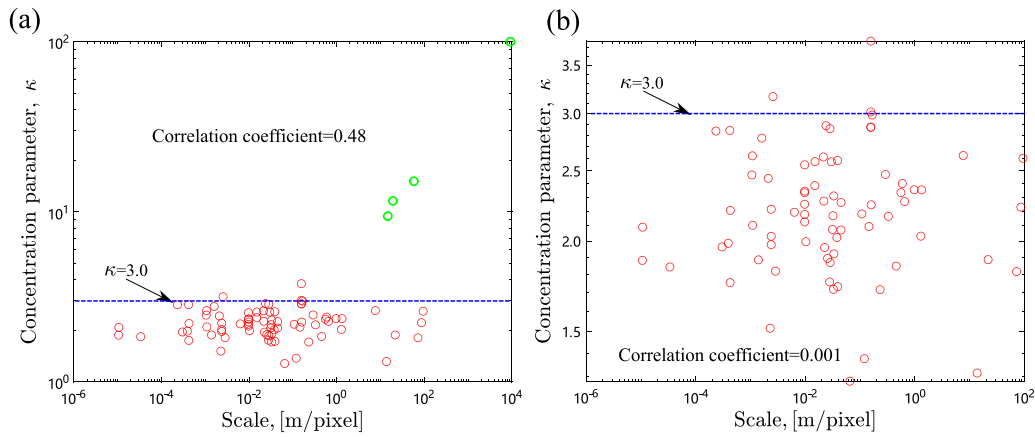


Fig. 8. (a) Concentration parameters, κ , of 80 published outcrop maps; (b) Concentration parameters, κ , of 76 published outcrop maps by removing four anomalous data points (green points in (a)).

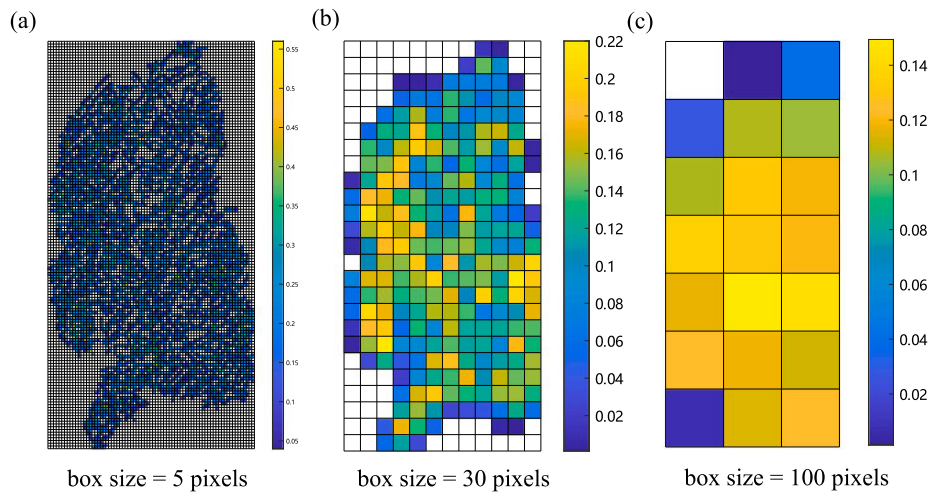


Fig. 9. Box intensities of the fracture outcrop map at Achnashellach Culmination field area (Fig. 7B in Watkins et al., 2015) with different box sizes.

where P_{21}^i follows the notation of fracture intensity proposed by Der-showitz et al. (1992), and it means the total length of fracture traces per unit area in the box i ; L_i is the total length of fracture traces in the box i ; A_i is the area of the box i .

The box size should choose properly according to the size and complexity of the fracture network with trial and error. If the box size is too small (Fig. 9a), the detailed structure of a fracture network is captured, but the void space between fractures is uncovered, so the spatial clustering cannot be measured. If the box size is too large (Fig. 9c), then the domain is over-averaged, and spatial variations of box intensities are insignificant. A proper box size should make most void space inside the fracture network covered, and the spatial variations of box intensities are preserved. Through trial and error, 30 pixels is a proper box size for most outcrop maps in this research, and one example is shown in Fig. 9b.

The spatial variations of the box intensity reveal the spatial clustering of fractures. We use the coefficient of variation (CV) to measure the spatial clustering of a fracture network:

$$CV = \frac{\sigma}{\mu}, \tag{8}$$

where σ is the standard deviation of box intensities; μ is the mean value of box intensities.

The CV value is sensitive to the box size as discussed above. However, 30 pixels as the box size is a good option for most outcrop maps. For a few outcrop maps with sparse fractures, boxes with a size of 30 pixels cannot cover the most void space, and 100 pixels is a better

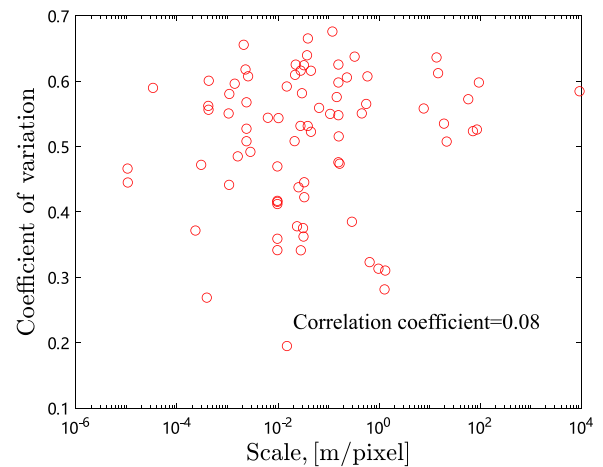


Fig. 10. The coefficient of variations of box intensities of all 80 published outcrop maps.

choice. However, CV values calculated for those outcrop maps with box sizes of 30 and 100 pixels are close. Therefore, we take a box size of 30 pixels for all outcrop maps in this research, and the spatial clustering effects in different fracture maps are comparable. Fig. 10 presents the compilation of all CV values for 80 published outcrop maps.

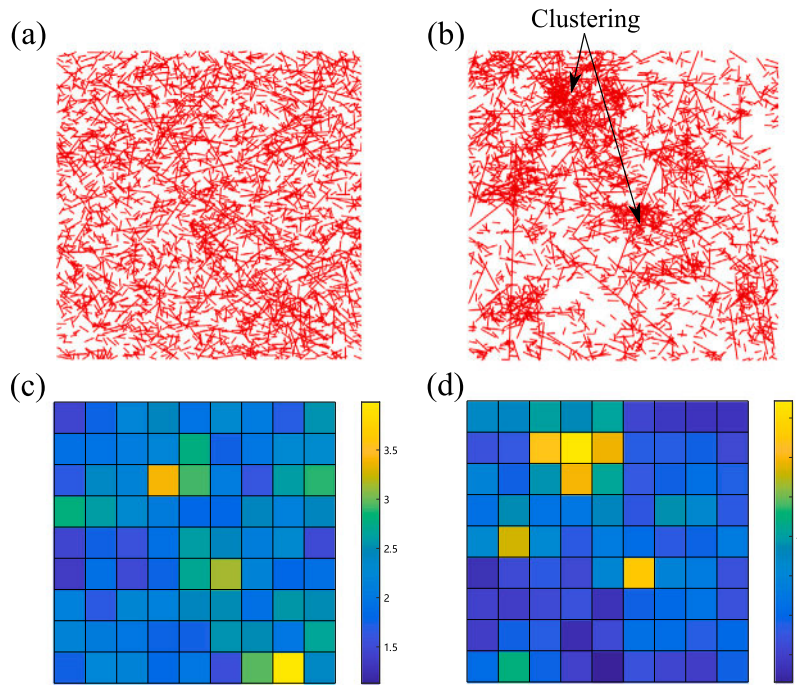


Fig. 11. (a) and (b) are examples of SDNFs generated by HatchFrac (Zhu et al., 2022b). (c) and (d) are box intensities of fracture networks (a) and (b), respectively. Two stochastic discrete fracture networks have their lengths that obey a power-law distribution, and the exponent is 3.0. Fracture orientations follow a von Mises–Fisher distribution with $\kappa = 1.5$. Fracture center positions follow a uniform (left) and fractal (right) spatial density distribution, respectively. The fractal dimension used in the right subfigure is 1.8. The termination criterion is forming a spanning cluster connecting four sides of the domain.

The correlation coefficient between CV and the scale is almost 0, indicating that spatial clustering can exist in all fracture networks regardless of the scale. The maximum, minimum, and mean values of CV are 0.68, 0.20, and 0.51, respectively. The standard deviation of CV is 0.1, implying an insignificant variation of CV in different outcrop maps.

Spatial clustering can be attributed to impacts of fracture geometries, such as fracture positions, lengths, and orientations. It is difficult to separate contributions from each factor. To focus on the impact of fracture positions, we generate two stochastic fracture networks with their fracture centers following a fractal and uniform spatial density distribution, respectively. We keep the other geometric parameters the same and investigate the clustering effect caused by center positions.

Figs. 11(a) and (b) show the two stochastic fracture networks, and their box intensities are shown in Figs. 11(c) and (d). These two fracture networks have lengths that follow a power-law distribution with the exponent equal 3.0, and orientations follow a von Mises–Fisher distribution with $\kappa = 1.5$. The CV values for the two fracture networks are 0.26 and 0.65, respectively. The fracture network with uniformly distributed fracture centers has the CV smaller than most outcrop maps. In contrast, the fracture network with clustered center positions has a much larger CV and is closer to reality. This observation suggests that natural fracture networks may have clustering effects caused by their clustered positions. However, fracture lengths and orientations may have significant impacts on CV as well. Therefore, more detailed and strict variable control should be implemented in future research to evaluate each geometric parameter’s impact comprehensively.

3.4. Topological structures

Fracture connectivity is essential for fluid flow in complex fracture networks, and it depends on fracture lengths, orientations, and fracture intensities (Zhu et al., 2021c). However, topology analysis can bring more insights into the connectivity compared with individual geometrical parameters (Sanderson and Nixon, 2015). Barton and Hsieh (1989) introduced a ternary diagram method to evaluate connectivity, where

a point representing the proportion of the three node types (I, T, X-type) is plotted. In this research, fracture apertures are not considered; therefore, a ternary diagram is sufficient to describe the topological structures of fracture networks. The three node types include isolated nodes (I-type), crossing fracture intersections (X-type), and abutting fracture intersections (T-type). Following Sanderson and Nixon (2015), we adopt C_B , the mean number of connections of each branch, to evaluate connectivity.

$$C_B = \frac{3 \times N_T + 4 \times N_X}{N_B}, \quad (9)$$

where N_T and N_X are the number of T-type and X-type nodes, respectively; N_B is the number of branches, calculated by:

$$N_B = \frac{1}{2}(N_I + 3N_T + 4N_X), \quad (10)$$

where N_I is the number of I-type nodes.

C_B is a dimensionless number varying between 0 and 2.0, and a larger value indicates better connectivity. Fig. 12(a) presents the ternary diagram of all 80 outcrop maps, and the color map refers to the connectivity index C_B . The contour line of C_B is denoted in the figure.

Fracture networks with a higher proportion of T-type and X-type nodes have better connectivity. Most natural outcrops have good connectivity because their C_B value is significant. More importantly, most natural fracture networks have a large proportion of T-type nodes. T-type nodes are essential to improve the connectivity of a natural fracture system because a high proportion of T-type nodes can lead to fewer dead-ends within each connected cluster (Dershowitz and Einstein, 1988; Barton and Hsieh, 1989).

Fig. 12(b) shows the plot between the connectivity index C_B and the scale. The correlation coefficient is -0.42 , indicating a moderate negative correlation between these two parameters. Large-scale fracture networks usually have weaker connectivity than small-scale fracture networks under the data collection in this research. One crucial factor is that large fracture systems are more likely to have concentrated orientations, as discussed in the previous section on fracture orientations. Such

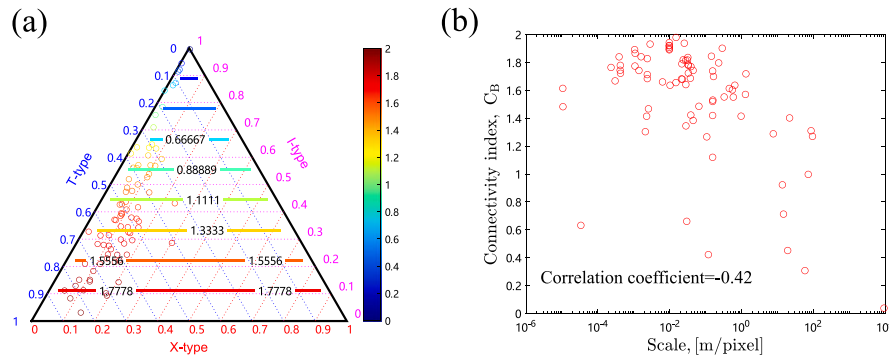


Fig. 12. (a) Ternary diagram of three types of nodes, I-type, T-type, and X-type. The contour lines of C_B are shown in different colors. (b) Connectivity indexes of all 80 published outcrop maps.

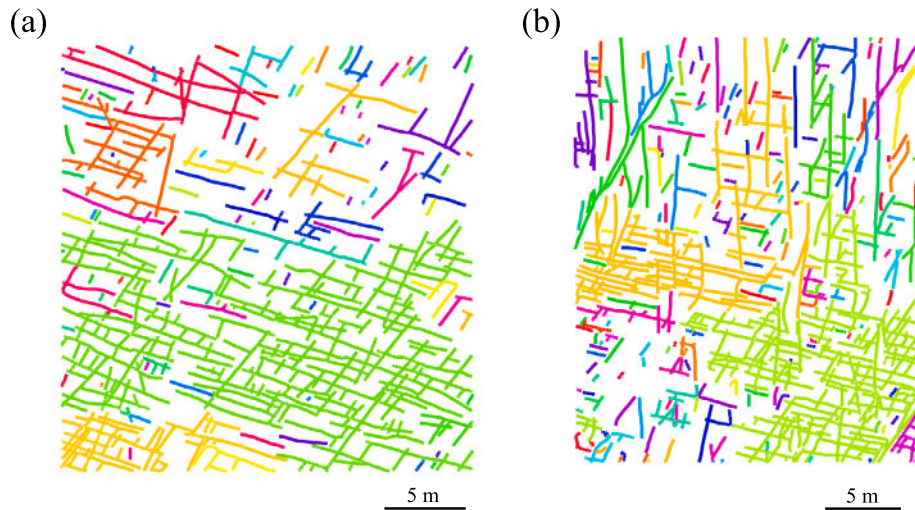


Fig. 13. Fracture outcrop maps at the Spireslack open cast coal pit, south of Glasgow in Scotland (Fig. 7 in Healy et al., 2017). Different colors are applied to distinguish different clusters.

concentrated orientations make large fractures difficult to link each other and form more complex and better-connected fracture networks.

3.5. Clusters and flow analysis

Topological analysis can quantify the connectivity of fracture networks but cannot explicitly demonstrate the flow pathways. Investigating the impact of fracture networks on the subsurface flow requires at least two aspects. One is the cluster analysis, where local and global connected fractures are identified. The other one is the impact of fracture sealing and stress on fracture permeability.

Fluid flow happens in connected instead of isolated fractures for low permeability formations. Therefore, clusters of connected fractures in outcrop maps must be checked. With the DFN modeling software, HatchFrac (Zhu et al., 2022b), we can check and label clusters of those outcrop maps after the digitization. Fig. 13 shows different clusters in outcrop maps from Spireslack open cast coal pit, south of Glasgow in Scotland (Fig. 7 in Healy et al., 2017). We use different neighboring colors to distinguish different clusters. In Fig. 13, no spanning cluster is formed (a spanning cluster connects all boundaries of the outcrop map). Local clusters have different sizes and yield good local instead of global connectivity. Fig. 14 shows clusters in outcrop maps at Achnashellach Culmination field area (Figs. 7B and 7D in Watkins et al., 2015). Spanning clusters are formed in both outcrop maps marked in red. 63 out of 80 natural outcrop maps have formed a spanning cluster, shown as red data points in Fig. 15. For small-scale outcrop maps,

the percentage of forming a spanning cluster is even higher, indicating good global connectivity of fracture networks. Fig. 15 also shows the fracture intensity P_{21} calculated for the entire fracture network in each outcrop map.

A well-connected fracture network is a premise for the subsurface flow in low-permeability formations. However, closure and sealing of fractures caused by compression and cementation over the geologic time can significantly reduce the fracture permeabilities (Ito and Zoback, 2000; Im et al., 2018). Under today's stress field or having severe stress perturbation, e.g., hydraulic fracturing, critically orientated fractures can be critically stressed and slide (Barton et al., 1995). Sliding of critically stressed fractures can induce shear dilation and enlarge the fracture aperture due to fracture roughness (Kim and Inoue, 2003). Non-critically stressed fractures are probably sealed and impermeable after a long geological history.

To better demonstrate the impact of fracture sealing and stress, Fig. 16 shows a global stress analysis of real outcrop maps. In Fig. 16, we assume the principal stress $S_1 = 1.0$ as the reference, with the orthogonal principal stress $S_2 = 0.6S_1$ and pore pressure $P_p = 0.5S_1$. The Coulomb failure criterion (Coulomb, 1773) is adopted for simplicity to distinguish critical and non-critical stressed fractures (Im et al., 2018; Mattila and Follin, 2019):

$$\tau = \mu(S_n - P_p), \quad (11)$$

where μ is the friction coefficient of the fracture plane, P_p refers to the local pore pressure, and τ and S_n are the shear and normal stresses on a fracture plane, respectively.

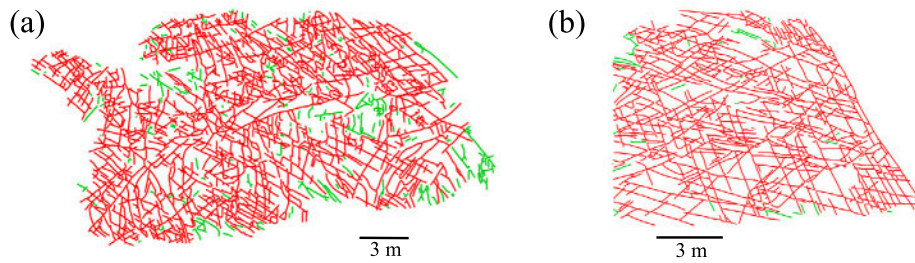


Fig. 14. Fracture outcrop maps at Achnashellach Culmination field area (Figs. 7B and 7D in Watkins et al., 2015). Red and green line segments refer to the largest spanning cluster and local clusters, respectively. After (Zhu et al., 2022a).

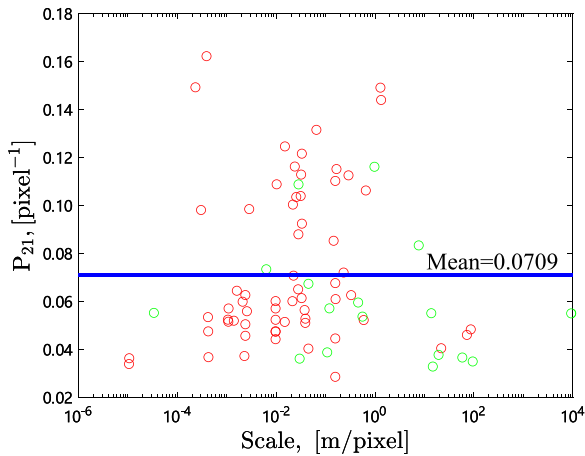


Fig. 15. Fracture intensity (P_{21}) of 80 published outcrop maps. Red and green data points refer to outcrops with and without a spanning cluster formed, respectively. After (Zhu et al., 2022a).

In Fig. 16, red fractures are critically stressed and highly permeable due to sliding; blue fractures are mechanically stable and non-permeable because of sealing. Non-critically stressed fractures could also be partially sealed and yield complex sealing patterns. The local stress impact can be essential for fracture reactivation and propagation. However, those factors are out of the scope of this research, and relevant work is available in our previous (Chen and Wang, 2017; Zhu et al., 2021a) and future research. The stress state of each fracture is shown in a Mohr's diagram in Figs. 16(c, d). Red crosses refer to stress states of critically stressed fractures, and blue dots are stress states of mechanically stable fractures.

To explicitly demonstrate the impact of fractures on the formation permeability, a full-scale, embedded discrete fracture matrix simulation is conducted with UNCONG simulator (Li et al., 2015), and the formation permeabilities in different scenarios are calculated. A square subdomain (green square) from Fig. 16(a) is taken for convenient implementation of boundary conditions, and the size of the subdomain is set as 100 m \times 100 m. A constant pressure boundary condition is applied, where the inflow boundary (left side) has a constant pressure of 2.0 bar, and the remaining three boundaries have zero bar. The pressure difference provides a macroscopic pressure gradient of 2.0 kPa/m, which makes the Reynolds number close to a realistic range, $\mathcal{O}(10^{-3})$. The upper and lower boundaries are impermeable. The matrix has a permeability of 0.1 mD since we are considering a low permeability formation. Critically stressed fractures are assumed to have a permeability of 20,833 mD, six orders of magnitude larger than the matrix permeability. Three scenarios are considered to demonstrate the impact of fractures on the formation permeability: i. no fractures; ii. critically stressed fractures have a permeability of 20,833 mD, and non-critically stressed fractures are impermeable (Fig. 17(a));

iii. both critically stressed and non-critically stressed fractures have a permeability of 20,833 mD (Fig. 17(b)).

The formation permeability of the first scenario is set as the reference, where $K_{ref} = 1.0$. Pressure distributions of three scenarios are presented in Figs. 17(c–e). If all fractures are highly permeable, fractures can significantly increase the formation permeability by 262%. However, if only critically stressed fractures are permeable, the increase is only 8% and insignificant. Although the fracture network in the chosen subdomain has good topological connectivity, its hydraulic conductivity is not guaranteed. Fractures can be essential for fluid flow in low permeability formations. However, their impacts highly rely on many other factors besides topological connectivity, such as fracture sealing degrees and patterns, and local and global stress conditions, which will be further addressed in future work.

4. Discussion

Through digitizing published natural outcrop maps with the pixel-based fracture detection algorithm, we systematically investigate geometric patterns of fractures, including fracture lengths, intensities, orientations, and topological structures. The geometric patterns are essential for discrete fracture network modeling since the critical strategy of SDFN modeling is to describe fracture geometries with statistical distributions. The validation of such discrete fracture networks is impossible because the comprehensive information about the subsurface structures is inaccessible with current technologies. The confidence of SDFN models thus depends on the justification of chosen statistical distributions.

A power-law distribution usually describes fracture lengths. However, this research shows that fracture lengths may not follow a single power-law distribution but multiple power-law distributions with varying exponents. The simplistic model proposed explains a possible origin of the power-law distribution considering fracture growth and linkage. The derived power-law exponents of fracture lengths are larger than 1.0, consistent with most outcrop observations. Two parameters adopted in the simplistic model, the coalescence probability and the number of coalesced minor fractures, can successfully explain the variation of power-law exponents for fractures with different lengths. The exponent is usually large for long fractures, possibly attributed to a small coalescence probability caused by concentrated orientations, sparse fractures, and high requirements on stress conditions.

Over a long geologic history of subsurface rocks, varying stress states can generate fracture sets with different orientations (Barton et al., 1995). Each fracture set has concentrated orientations and thus has a large κ value. However, the stress history is hardly known, and most SDFN models regard all sets of fractures belonging to an integrated fracture system and describe their orientations with a uniform distribution. From observations in this paper, most small-scale fracture systems (< 100 m) have their orientation distribution close to a uniform distribution because the corresponding concentration degree κ is smaller than 3.0 in a von Mises–Fisher distribution. However, large fracture systems usually have more concentrated orientations with a large κ value as observed in the collected datasets in this research. A

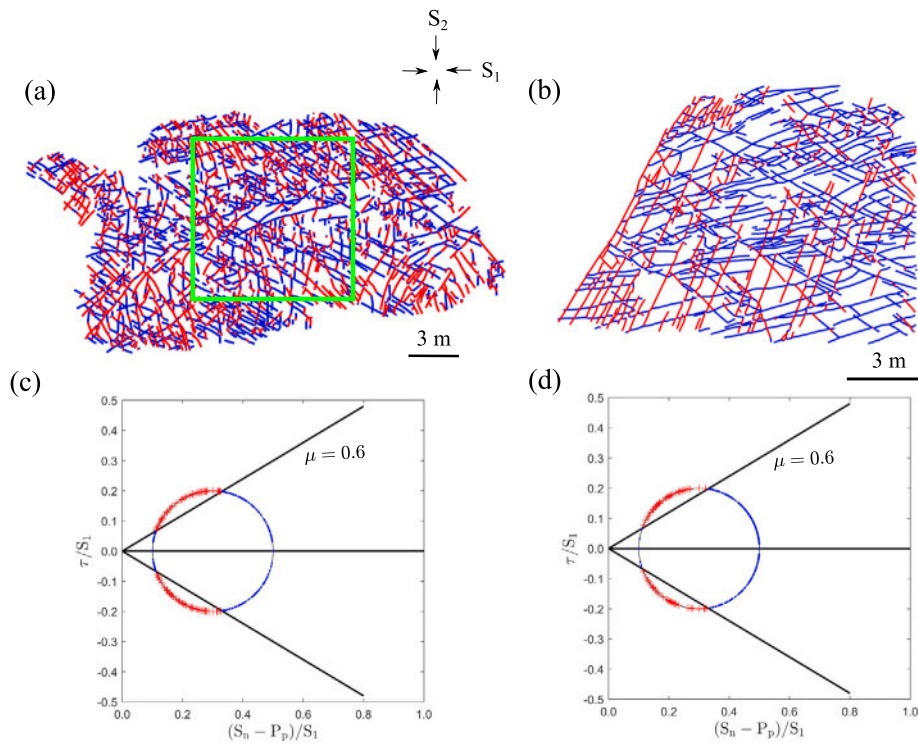


Fig. 16. Fracture outcrop maps at Achnashellach Culmination field area (Figs. 7B and 7D in Watkins et al., 2015). Red line segments refer to critically stressed fractures and blue line segments are mechanically stable fractures. The stress state of each fracture plane is shown in the Mohr's diagram. A subdomain (green square) is selected for the flow simulation.

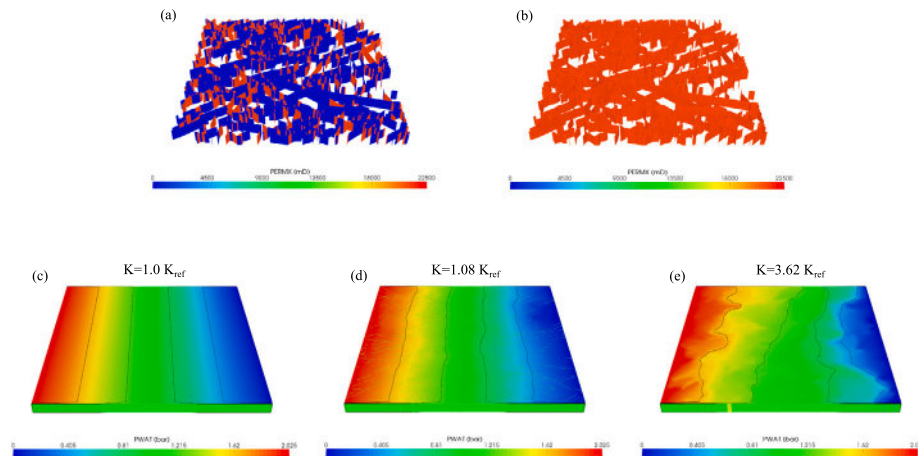


Fig. 17. (a) Fracture permeability considering stress impact; (b) Fracture permeability considering no stress impact (all fractures are open); (c–e) pressure distribution of the formation in different scenarios: (c) no fracture; (d) critically stressed fractures have a permeability of 20,833 mD, and non-critically stressed fractures are impermeable; (e) both critically and non-critically stressed fractures have a permeability of 20,833 mD.

positive correlation exists between the scale and concentration degree (κ). Concentrated orientations of large fracture systems can make them difficult to intersect and merge, thus making large fractures have a small coalescence probability.

Fracture intensities are usually not spatially uniform but clustered. Spatial clustering can improve local connectivity but hardly contribute to global connectivity. Such spatial clustering can exist in all fracture networks regardless of scale, possibly attributed to fracture lengths, orientations, and fracture positions. Evaluating the impact of each geometric parameter on spatial clustering needs more strict variable control and detailed investigations, which is beyond the scope of this paper and will be discussed in future research. A fractal spatial density distribution generates clustered positions and significantly increases the spatial clustering of the system. Therefore, a fractal spatial density

distribution can better capture the spatial clustering of fracture systems compared with a uniform spatial density distribution.

The topological analysis finds that most natural fracture networks have a significant proportion of T-type nodes and good connectivity. However, commonly adopted SDFN models cannot generate T-type nodes but only X-type and I-type nodes. A random truncation of branches at X-type nodes to form T-type nodes is undesirable for developing realistic fracture networks because the abutment relationship reveals the growth history of fractures. Therefore, it is necessary to mimic the fracture growth process and form T-type intersections. Detailed numerical simulation based on fracture mechanics is inapplicable for discrete fracture networks with massive fractures. Rule-based fracture growth process (Davy et al., 2013) is more appropriate for the implementation. Rules, such as nuclei distributions, growth criteria,

growth velocities, and termination criteria, can be summarized from classic theories, experiments, and numerical simulations. Then, growth rules can be incorporated into an SDFN model at each time step to mimic the growth process.

Furthermore, there is a spanning cluster formed in 63 out of 80 natural outcrop maps, indicating good topological connectivity of exposed fracture networks. However, fracture networks are three-dimensional in the subsurface, and outcrops can only be regarded as cross-section maps of the corresponding 3D structures with the ground as the cross-sectional plane. If the surface outcrops and subsurface formations are similar regarding the structural patterns and rock types, they can be relevant to each other. Complex factors, including weathering and stress-release, cause significant differences between outcrops and the subsurface systems (Ukar et al., 2019). Therefore, 2D outcrop maps cannot completely characterize the real subsurface fracture networks. A preliminary investigation of the correlation between the connectivity in 2D and 3D fracture networks is available in our recent work (Zhu et al., 2022a). In 2D outcrop maps, good topological connectivity (formation of the spanning cluster) does not ensure good hydraulic conductivity because fracture permeability can be significantly reduced, attributing to the compression and sealing. Therefore, considering sealing degrees and patterns, local and global stress states are necessary to evaluate the flow contribution from fractures in low permeability formations.

In a nutshell, several improvements listed above are available for current SDFN models to be more representative.

5. Conclusions

We have analyzed fracture geometries in detail with 80 published outcrop maps. The findings and observations from this research help construct a representative discrete fracture network. A meaningful SDFN model is a premise for investigating complex flow behaviors in the subsurface. The key conclusions from our research are:

- A pixel-based fracture detection algorithm can interpret binary outcrop maps as polylines. The algorithm is robust, accurate, and efficient.
- For most outcrop maps, fracture lengths follow multiple power-law distributions instead of a single one. Large fractures usually have large exponents, possibly because of a small coalescence probability.
- Natural fracture systems are usually composed of many fracture sets with different orientations, which results in small κ values in a von Mises–Fisher distribution. Most small-scale fracture systems have their concentration parameter κ smaller than 3.0. Large fracture systems usually have more concentrated orientations with large κ values.
- Fracture intensities are usually spatially clustered instead of uniformly distributed in fracture systems at all scales. A fractal spatial density distribution can better capture spatial clustering by generating clustered fracture positions.
- Natural fracture networks are usually well connected with many T-type intersections. A rule-based algorithm is necessary to mimic fracture growth and form T-type nodes in conventional SDFN modeling.
- Most collected natural outcrop maps form a spanning cluster in this work, indicating good topological connectivity. However, fracture sealing and stress impact on the fracture permeability should be considered to evaluate the conductivity of fracture networks.

In this work, we collect and analyze 2D outcrop maps. However, 2D outcrop maps can be regarded as relevant to the subsurface structures when the rock types and geological settings of outcrop and subsurface formations are similar. Furthermore, outcrops are 2D cross-section maps of actual 3D fracture networks in the subsurface after experiencing severe weathering and stress release. More research is necessary to link the properties, such as fracture intensity, connectivity, and conductivity, in different dimensions.

CRedit authorship contribution statement

Weiwei Zhu: Investigation, Writing – original draft. **Xupeng He:** Data curation. **Ryan Kurniawan Santoso:** Formal analysis. **Gang Lei:** Data curation. **Tadeusz W.Patzek:** Conceptualization, Supervision. **Moran Wang:** Conceptualization, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All the synthetically generated data and digitized outcrop maps are available online: <https://doi.org/10.4121/14865096.v2>.

Acknowledgments

This project was supported by the National Key Research and Development Program of China (No. 2019YFA0708704). The authors would like to thank Dr. X. Li from Ennosoft for providing the UNCONG simulator for the flow simulation. The authors would like to thank all editors and anonymous reviewers for their comments and suggestions.

Appendix A. A simplistic model

We propose a simplistic model to demonstrate a possible origin of the power-law distribution of fracture lengths. The model simplifies the fracture growth process by capturing the two most important fracture patterns: self-similarity and fracture coalescence. Fig. A.18 shows similar fracture geometries in different scales and a sketch map for the coalescence of echelon fracture segments. The basic assumptions are listed below:

- Initial fractures are small fractures with a constant length, l_0 , and the number of initial fractures is N_0 ;
- Large fractures grow from the coalescence of small fractures. n_s is number of small fractures, which coalesce and form one large fracture ($n_s = 3$ in Fig. A.18b);
- The length of coalesced fracture equal $n_s \times l_{i-1}$. The overlapping and underlapping structures are ignored;
- A constant coalescence probability, p , applies on each generation and decides the number of small fractures coalesced to form large fractures in each generation.
- At the i th generation, $\frac{N_{i-1} \times p}{n_s}$ fractures will coalesce and form the i th generation fractures with length l_i . The number of remaining fractures at the $i - 1$ generation is denoted as $N_{i-1(\text{remain})}$ and those fractures are observable with length l_{i-1} , which is equal to $N_{i-1} \times (1 - p)$

At the initial state ($n = 0$), we have

$$\begin{cases} N_0 = N_0 \\ l_0 = l_0 \end{cases} \quad (\text{A.1})$$

At the first generation, $n = 1$

$$\begin{cases} N_1 = \frac{N_0 \times p}{n_s} \\ l_1 = n_s l_0 \\ N_{0(\text{remain})} = N_0(1 - p) \end{cases} \quad (\text{A.2})$$

where $N_{0(\text{remain})}$ is the remaining fractures at the initial state.

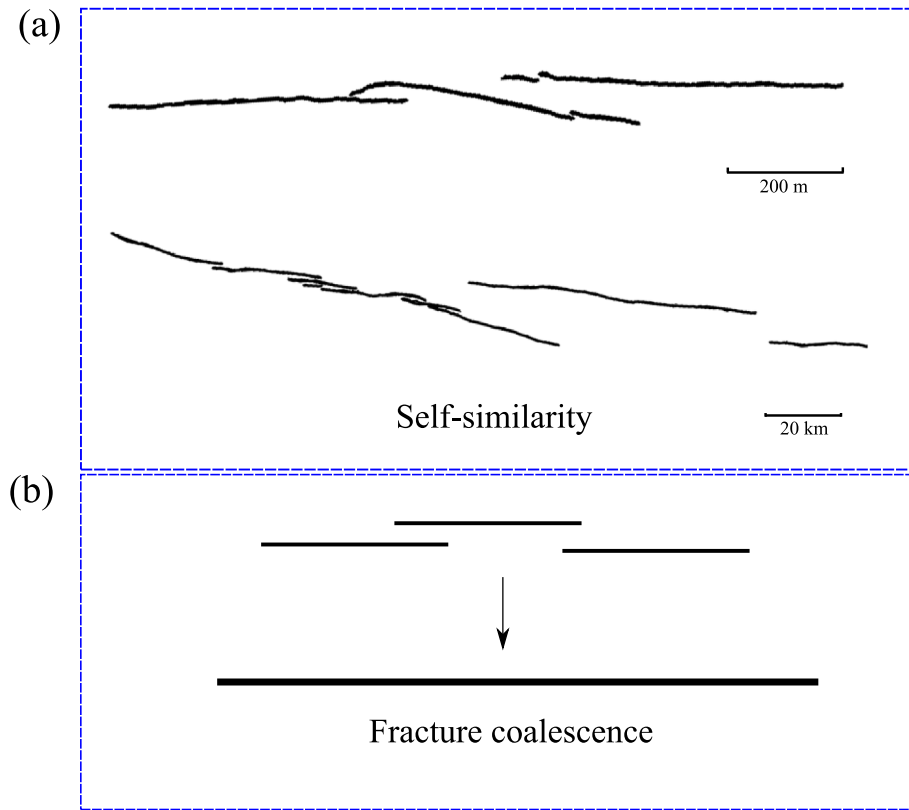


Fig. A.18. (a) Demonstration of self-similar geometries of natural fractures in different scales (after Zhu et al., 2021d); (b) A sketch map showing the coalescence of echelon fracture segments.

At the second generation, $n = 2$

$$\begin{cases} N_2 = \frac{N_1 \times p}{n_s} = \left(\frac{p}{n_s}\right)^2 N_0 \\ l_2 = n_s^2 l_0 \\ N_{1(\text{remain})} = N_1(1-p) = \frac{N_0 \times p}{n_s}(1-p) \end{cases} \quad (\text{A.3})$$

At the N th generation, $n = N$

$$\begin{cases} N_N = \frac{N_{N-1} \times p}{n_s} = \left(\frac{p}{n_s}\right)^N N_0 \\ l_N = n_s^N l_0 \\ N_{N-1(\text{remain})} = N_{N-1}(1-p) = \left(\frac{p}{n_s}\right)^{N-1} (1-p) N_0 \end{cases} \quad (\text{A.4})$$

Take the logarithm of the first two equations in Eqs. (4), we have

$$\begin{cases} \ln\left(\frac{N_N}{N_0}\right) = N \ln\left(\frac{p}{n_s}\right) \\ \ln\left(\frac{l_N}{l_0}\right) = N \ln(n_s) \end{cases} \quad (\text{A.5})$$

Therefore, we have

$$\ln\left(\frac{N_N}{N_0}\right) = \ln\left(\frac{l_N}{l_0}\right) \ln\left(\frac{p}{n_s}\right) / \ln(n_s) \quad (\text{A.6})$$

which is equal to

$$\ln\left(\frac{N_N}{N_0}\right) = \ln\left\{\left(\frac{l_N}{l_0}\right)^{\ln(p)/\ln(n_s)}\right\} \quad (\text{A.7})$$

Therefore,

$$\left(\frac{N_N}{N_0}\right) = \left(\frac{l_N}{l_0}\right)^{\frac{\ln(p)}{\ln(n_s)} - 1} \quad (\text{A.8})$$

Therefore, we see that the length of fractures at different generations (or different scales) follows a power-law distribution, and the exponent should be smaller than -1 (when p equals 1).

However, what we observe in reality is the remaining fractures. With the similar procedures, the number of remaining fractures at the

N th generation follow a power law distribution.

$$\frac{N_{N(\text{remain})}}{(1-p)N_0} = \left(\frac{l_N}{l_0}\right)^{\frac{\ln(p)}{\ln(n_s)} - 1} \quad (\text{A.9})$$

Since both N_N and $N_{N(\text{remain})}$ follow a power-law distribution with the same exponent, but different coefficients, we do not distinguish these two parameters and denote the number of fractures with length l_N as N_N for the following discussion.

- When $p = 0$, which corresponds to the initial state, we have

$$\left(\frac{N_N}{N_0}\right) = \left(\frac{l_N}{l_0}\right)^{-\infty} = 0 \quad (\text{A.10})$$

therefore, the number of all fractures larger than l_0 is zero and we only have fractures at the initial state.

- When $p = 1.0$, which corresponds to the case where all small fractures are coalesced to form large fractures. We have

$$\left(\frac{N_N}{N_0}\right) = \left(\frac{l_N}{l_0}\right)^{-1} = \frac{l_0}{l_N} \quad (\text{A.11})$$

therefore,

$$N_N = \frac{l_0}{l_N} \times N_0 \quad (\text{A.12})$$

This scenario has the exponent of 1.0 and the numbers of remaining fractures at all previous generations are zero.

- Because of complex stress states and interactions between fractures (such as stress shadow), the coalescence probability cannot be 1.0. If we take $p = 0.7$ and $n_s = 3$, we have the exponent

$$a = -\{(\ln(0.7)/\ln(3)) - 1\} = 1.32 \quad (\text{A.13})$$

If $p = 0.3$ and $n_s = 3$, the exponent is about 2.1; If $p = 0.1$ and $n_s = 3$, the exponent is about 3.1; If $p = 0.01$ and $n_s = 3$, the exponent is about 5.9. Therefore, a larger exponent may suggest a smaller coalescence probability.

These derivations can explain why fracture lengths of different scales follow a power-law distribution and provide a possible range for the exponent $[1, \infty]$. The exponent depends on the coalescence probability and the number of coalesced fractures, but it has to be larger than 1. Bonnet et al. (2001) provided a compilation of power-law exponents for fracture length distributions of different outcrop maps. Only one map has their exponent equal to 0.9 (table 2 in Bonnet et al., 2001). A possible reason is their length measurement is inaccurate, and the fitting is not perfect because the outcrop map shows a km-scale fault system.

Appendix B. Cumulative length distribution

Furthermore, we can derive the cumulative length distribution by an integration of the length distribution.

$$c_N \sim \int_{l_{min}}^{l_N} l^{-a} dl = (1-a) \{l_N^{1-a} - l_{min}^{1-a}\} \quad (\text{B.1})$$

therefore:

$$c_N \sim l_N^{1-a} \quad (\text{B.2})$$

where $a = \frac{\ln(p)}{\ln(n_s)} - 1$

From collected length data, we can see that the cumulative length distribution does not follow a single power-law distribution but multiple power-law distributions. The exponents become larger when fracture lengths increase. The larger exponent means a smaller coalescence probability. Therefore, the coalescence probability, p , should be a function of the fracture length:

$$p = p(l, \text{others}) \quad (\text{B.3})$$

From collected data, we know $\frac{\partial p}{\partial l} < 0$.

References

- Adler, P.M., Thovert, J.-F., 1999. *Fractures and Fracture Networks*, vol. 15. Springer Science & Business Media.
- Ahmadi, H., Pekkan, E., 2021. Fault-based geological lineaments extraction using remote sensing and GIS—A review. *Geosciences* 11 (5), 183.
- Akara, M.-E.M., Reeves, D.M., Parashar, R., 2021. Impact of horizontal spatial clustering in two-dimensional fracture networks on solute transport. *J. Hydrol.* 603, 127055.
- Barton, C.C., 1995. *Fractals in the Earth Sciences*. Springer.
- Barton, C.C., Hsieh, P.A., 1989. Physical and hydrologic-flow properties of fractures. In: 28th International Geological Congress Field Trip Guidebook, vol. 385, p. 36.
- Barton, C.A., Zoback, M.D., Moos, D., 1995. Fluid flow along potentially active faults in crystalline rock. *Geology* 23 (8), 683–686.
- Becker, I., Koehler, B., Waldvogel, M., Jelinek, W., Hilgers, C., 2018. Comparing fracture statistics from outcrop and reservoir data using conventional manual and t-LiDAR derived scanlines in Ca2 carbonates from the Southern permian basin, Germany. *Mar. Pet. Geol.* 95, 228–245.
- Berkowitz, B., 1995. Analysis of fracture network connectivity using percolation theory. *Math. Geol.* 27 (4), 467–483.
- Berkowitz, B., 2002. Characterizing flow and transport in fractured geological media: A review. *Adv. Water Resour.* 25 (8–12), 861–884.
- Bertrand, L., Géraud, Y., Le Garzic, E., Place, J., Diraison, M., Walter, B., Haffen, S., 2015. A multiscale analysis of a fracture pattern in granite: A case study of the tamariu granite, Catalunya, Spain. *J. Struct. Geol.* 78, 52–66.
- Bisdom, K., 2016. Burial-related fracturing in sub-horizontal and folded reservoirs: Geometry, geomechanics and impact on permeability.
- Bonnet, E., Bour, O., Odling, N.E., Davy, P., Main, I., Cowie, P., Berkowitz, B., 2001. Scaling of fracture systems in geological media. *Rev. Geophys.* 39 (3), 347–383.
- Bour, O., Davy, P., 1997. Connectivity of random fault networks following a power law fault length distribution. *Water Resour. Res.* 33 (7), 1567–1583.
- Cartwright, J.A., Trudgill, B.D., Mansfield, C.S., 1995. Fault growth by segment linkage: An explanation for scatter in maximum displacement and trace length data from the canyonlands grabens of SE Utah. *J. Struct. Geol.* 17 (9), 1319–1326.
- Chae, B.G., Ichikawa, Y., Jeong, G.C., Seo, Y.S., Kim, B.C., 2004. Roughness measurement of rock discontinuities using a confocal laser scanning microscope and the Fourier spectral analysis. *Eng. Geol.* 72 (3–4), 181–199.
- Chen, Z., Wang, M., 2017. Pore-scale modeling of hydromechanical coupled mechanics in hydrofracturing process. *J. Geophys. Res. Solid Earth* 122 (5), 3410–3429.
- Cladouhos, T.T., Marrett, R., 1996. Are fault growth and linkage models consistent with power-law distributions of fault lengths? *J. Struct. Geol.* 18 (2–3), 281–293.
- Cook, A.M., Myer, L.R., Cook, N.G.W., Doyle, F.M., 1990. The effects of tortuosity on flow through a natural fracture.
- Coulomb, C.A., 1773. *Essai sur une application des regles de maximis et minimis a quelques problemes de statique relatifs a l'architecture*. Mem. Div. Sav. Acad.
- Darcel, C., Bour, O., Davy, P., De Dreuzy, J.R., 2003. Connectivity properties of two-dimensional fracture networks with stochastic fractal correlation. *Water Resour. Res.* 39 (10).
- Davy, P., 1993. On the frequency-length distribution of the san andreas fault system. *J. Geophys. Res. Solid Earth* 98 (B7), 12141–12151.
- Davy, P., Le Goc, R., Darcel, C., 2013. A model of fracture nucleation, growth and arrest, and consequences for fracture density and scaling. *J. Geophys. Res. Solid Earth* 118 (4), 1393–1407.
- Dershowitz, W.S., Einstein, H.H., 1988. Characterizing rock joint geometry with joint system models. *Rock Mech. Rock Eng.* 21 (1), 21–51.
- Dershowitz, W.S., Herda, H.H., et al., 1992. Interpretation of fracture spacing and intensity. In: The 33th US Symposium on Rock Mechanics. USRMS, American Rock Mechanics Association.
- Duffy, O.B., Nixon, C.W., Bell, R.E., Jackson, C.A.-L., Gawthorpe, R.L., Sanderson, D.J., Whipp, P.S., 2017. The topology of evolving rift fault networks: Single-phase vs multi-phase rifts. *J. Struct. Geol.* 96, 192–202.
- Feng, S., Wang, H., Cui, Y., Ye, Y., Liu, Y., Li, X., Wang, H., Yang, R., 2020. Fractal discrete fracture network model for the analysis of radon migration in fractured media. *Comput. Geotech.* 128, 103810.
- Gertsch, L.S., 1995. Three-dimensional fracture network models from laboratory-scale rock samples. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* 7 (32), 322A.
- Gillespie, P.A., Howard, C.B., Walsh, J.J., Watterson, J., 1993. Measurement and characterisation of spatial distributions of fractures. *Tectonophysics* 226 (1–4), 113–141.
- Healy, D., Rizzo, R.E., Cornwell, D.G., Farrell, N.J., Watkins, H., Timms, N.E., Gomez-Rivas, E., Smith, M., 2017. FracPaQ: A MATLAB™ toolbox for the quantification of fracture patterns. *J. Struct. Geol.* 95, 1–16.
- Holland, M., Saxena, N., Urai, J.L., 2009. Evolution of fractures in a highly dynamic thermal, hydraulic, and mechanical system-(II) remote sensing fracture analysis, Jabal Shams, Oman Mountains. *GeoArabia* 14 (3), 163–194.
- Huang, L., Su, X., Tang, H., 2020. Optimal selection of estimator for obtaining an accurate three-dimensional rock fracture orientation distribution. *Eng. Geol.* 270, 105575.
- Im, K., Elsworth, D., Fang, Y., 2018. The influence of preslip sealing on the permeability evolution of fractures and faults. *Geophys. Res. Lett.* 45 (1), 166–175.
- Ito, T., Zoback, M.D., 2000. Fracture permeability and in situ stress to 7 km depth in the KTB scientific drillhole. *Geophys. Res. Lett.* 27 (7), 1045–1048.
- Jafari, A., 2011. *Permeability Estimation of Fracture Networks* (Ph.D. thesis). University of Alberta.
- Kemeny, J., Post, R., 2003. Estimating three-dimensional rock discontinuity orientation from digital images of fracture traces. *Comput. Geosci.* 29 (1), 65–77.
- Kim, H.M., Inoue, J., 2003. Analytical approach for anisotropic permeability through a single rough rock joint under shear deformation. *J. Geophys. Res. Solid Earth* 108 (B8).
- Koike, K., Nagano, S., Ohmi, M., 1995. Lineament analysis of satellite images using a segment tracing algorithm (STA). *Comput. Geosci.* 21 (9), 1091–1104.
- Laubach, S.E., 1988. Subsurface fractures and their relationship to stress history in East Texas Basin sandstone. *Tectonophysics* 156 (1–2), 37–49.
- Lei, Q., Latham, J.-P., Tsang, C.-F., 2017. The use of discrete fracture networks for modelling coupled geomechanical and hydrological behaviour of fractured rocks. *Comput. Geotech.* 85, 151–176.
- Li, X., Liu, J., Gong, W., Xu, Y., Bowa, V.M., 2022. A discrete fracture network based modeling scheme for analyzing the stability of highly fractured rock slope. *Comput. Geotech.* 141, 104558.
- Li, X., Zhang, D., Li, S., 2015. A multi-continuum multiple flow mechanism simulator for unconventional oil and gas recovery. *J. Nat. Gas Sci. Eng.* 26, 652–669.
- Makarov, P.V., 2007. Evolutionary nature of structure formation in lithospheric material: Universal principle for fractality of solids. *Russ. Geol. Geophys.* 48 (7), 558–574.
- Mattila, J., Follin, S., 2019. Does in situ state of stress affect fracture flow in crystalline settings? *J. Geophys. Res. Solid Earth* 124 (5), 5241–5253.
- Nur, A., 1982. The origin of tensile fracture lineaments. *J. Struct. Geol.* 4 (1), 31–40.
- Odling, N.E., 1997. Scaling and connectivity of joint systems in sandstones from western Norway. *J. Struct. Geol.* 19 (10), 1257–1271.
- Odling, N., Gillespie, P., Bourguin, B., Castaing, C., Chiles, J., Christensen, N., Fillion, E., Genter, A., Olsen, C., Thrane, L., et al., 1999. Variations in fracture system geometry and their implications for fluid flow in fractures hydrocarbon reservoirs. *Pet. Geosci.* 5 (4), 373–384.
- Olson, J.E., 2007. Fracture aperture, length and pattern geometry development under biaxial loading: A numerical study with applications to natural, cross-jointed systems. *Geol. Soc. Lond. Spec. Publ.* 289 (1), 123–142.
- Otsuki, K., Dilov, T., 2005. Evolution of hierarchical self-similar geometry of experimental fault zones: Implications for seismic nucleation and earthquake size. *J. Geophys. Res. Solid Earth* 110 (B3).
- Pickering, G., Bull, J.M., Sanderson, D.J., 1995. Sampling power-law distributions. *Tectonophysics* 248 (1–2), 1–20.

- Pogacnik, J., Elsworth, D., O'Sullivan, M., O'Sullivan, J., 2016. A damage mechanics approach to the simulation of hydraulic fracturing/shearing around a geothermal injection well. *Comput. Geotech.* 71, 338–351.
- Prabhakaran, R., Urai, J.L., Bertotti, G., Weismüller, C., Smeulders, D.M., 2021. Large-scale natural fracture network patterns: Insights from automated mapping in the lilstock (Bristol channel) limestone outcrops. *J. Struct. Geol.* 150, 104405.
- Priest, S.D., Hudson, J.A., 1976. Discontinuity spacings in rock. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* 13 (5), 135–148.
- Priest, S.D., Hudson, J.A., 1981. Estimation of discontinuity spacing and trace length using scanline surveys. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* 18 (3), 183–197.
- Prioul, R., Jocker, J., 2009. Fracture characterization at multiple scales using borehole images, sonic logs, and walkaround vertical seismic profile. *AAPG Bull.* 93 (11), 1503–1516.
- Reisenhofer, R., 2014. the Complex Shearlet Transform and Applications To Image Quality Assessment (Master's Thesis). Technical University of Berlin.
- Riley, M.S., 2005. Fracture trace length and number distributions from fracture mapping. *J. Geophys. Res. Solid Earth* 110 (B8).
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. [arXiv:1505.04597](https://arxiv.org/abs/1505.04597).
- Rouleau, A., Gale, J.E., 1985. Statistical characterization of the fracture system in the stripa granite, Sweden. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* 22 (6), 353–367.
- Sanderson, D.J., Nixon, C.W., 2015. The use of topology in fracture network characterization. *J. Struct. Geol.* 72, 55–66.
- Sano, O., Ito, I., Terada, M., 1981. Influence of strain rate on dilatancy and strength of oshima granite under uniaxial compression. *J. Geophys. Res. Solid Earth* 86 (B10), 9299–9311.
- Segall, P., Pollard, D.D., 1983a. Joint formation in granitic rock of the Sierra Nevada. *Geol. Soc. Am. Bull.* 94 (5), 563–575.
- Segall, P., Pollard, D.D., 1983b. Nucleation and growth of strike slip faults in granite. *J. Geophys. Res. Solid Earth* 88 (B1), 555–568.
- Song, J.-J., Lee, C.-I., Seto, M., 2001. Stability analysis of rock blocks around a tunnel using a statistical joint modeling technique. *Tunn. Undergr. Space Technol.* 16 (4), 341–351.
- Sornette, A., Davy, P., Sornette, D., 1993. Fault growth in brittle-ductile experiments and the mechanics of continental collisions. *J. Geophys. Res. Solid Earth* 98 (B7), 12111–12139.
- Sornette, D., Sornette, A., 1999. General theory of the modified gutenbergrichter law for large seismic moments. *Bull. Seismol. Soc. Am.* 89 (4), 1121–1130.
- Thiele, S.T., Grose, L., Samsu, A., Mickethwaite, S., Vollgger, S.A., Cruden, A.R., 2017. Rapid, semi-automatic fracture and contact mapping for point clouds, images and geophysical data. *Solid Earth* 8 (6), 1241–1253.
- Ukar, E., Laubach, S.E., Hooker, J.N., 2019. Outcrops as guides to subsurface natural fractures: Example from the Nikanassin formation tight-gas sandstone, Grande Cache, Alberta foothills, Canada. *Mar. Pet. Geol.* 103, 255–275.
- Wang, J., Howarth, P.J., 1990. Use of the hough transform in automated lineament. *IEEE Trans. Geosci. Remote Sens.* 28 (4), 561–567.
- Watkins, H., Bond, C.E., Healy, D., Butler, R.W., 2015. Appraisal of fracture sampling methods and a new workflow to characterise heterogeneous fracture networks at outcrop. *J. Struct. Geol.* 72, 67–82.
- Weismüller, C., Prabhakaran, R., Passchier, M., Urai, J.L., Bertotti, G., Reicherter, K., 2020. Mapping the fracture network in the lilstock pavement, Bristol channel, UK: Manual versus automatic. *Solid Earth* 11 (5), 1773–1802.
- Williams, J.H., Johnson, C.D., 2004. Acoustic and optical borehole-wall imaging for fractured-rock aquifer studies. *J. Appl. Geophys.* 55 (1–2), 151–159.
- Wyller, F.A., 2019. Spatio-Temporal Development of a Joint Network and Its Properties: a Case Study from Lilstock, UK (Master's thesis). The University of Bergen.
- Zhu, W., He, X., Patzek, T.W., 2021a. Insights into the coupled effects of fracture geometry and sealing on stimulated reservoir volume in shales. In: 3rd International Discrete Fracture Network Engineering Conference. OnePetro.
- Zhu, W., Khirevich, S., Patzek, T.W., 2018. Percolation properties of stochastic fracture networks in 2D and outcrop fracture maps. In: 80th EAGE Conference and Exhibition 2018, vol. 2018. (1), European Association of Geoscientists & Engineers, pp. 1–5.
- Zhu, W., Khirevich, S., Patzek, T., 2020. Fracture recognition with U-net and pixel-based automatic fracture detection. In: Fourth Naturally Fractured Reservoir Workshop, vol. 2020. (1), European Association of Geoscientists & Engineers, pp. 1–5.
- Zhu, W., Khirevich, S., Patzek, T.W., 2021c. Impact of fracture geometry and topology on the connectivity and flow properties of stochastic fracture networks. *Water Resour. Res.* 57 (7), e2020WR028652.
- Zhu, W., Khirevich, S., Patzek, T.W., 2022b. HatchFrac: A fast open-source DFN modeling software. *Comput. Geotech.* 150, 104917.
- Zhu, W., Lei, G., He, X., Yang, Y., Santoso, R.K., Wang, M., 2022a. Can we infer the percolation status of 3D fractured media from 2D outcrops? *Eng. Geol.* 106648.
- Zhu, W., Yalcin, B., Khirevich, S., Patzek, T.W., 2021d. Fault traces: Generation of fault segments and estimation of their fractal dimension. *Lithosphere* 2021 (Special 4), 4991604.