Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


# How do we see fractures? Quantifying subjective bias in fracture data collection.

Billy J. Andrews[1*], Jennifer J. Roberts[1], Zoe K. Shipton[1], Sabina Bigi[2], Maria C. Tartarello[2], Gareth O. Johnson[1,3]

1 Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, G11XJ, Scotland
2 Department of Earth Science, Sapienza – University of Rome, P.le Aldo Moro, 5, 00185 Rome, Italy
3 School of GeoSciences, University of Edinburgh, Edinburgh, EH93FE, Scotland

*Correspondence to:* Billy J. Andrews (billy.andrews@strath.ac.uk)

**Abstract.** The characterisation of natural fracture networks using outcrop analogues is important in understanding sub-surface fluid flow and rock mass characteristics in fractured lithologies. It is well known from decision-sciences that subjective bias significantly impacts the way data is gathered and interpreted. This study investigates the impact of subjective bias on fracture data collected using four commonly used approaches (linear scanlines, circular scanlines, topology sampling and window sampling) both in the field and in workshops using field photographs. Considerable variability is observed between each participant's interpretation of the same scanline, and this variability is seen regardless of geological experience. Geologists appear to be either focussing on the detail or focussing on gathering larger volumes of data, and this innate personality trait affects the recorded fracture network attributes. As a result, fracture statistics derived from the field data and which are often used as inputs for geological models, can vary considerably between different geologists collecting data from the same scanline. Additionally, the personal bias of geologists collecting the data affects the size (minimum length of linear scanlines, radius of circular scanlines or area of a window sample) required of the scanline that is needed to collect a statistically representative amount of data. We suggest protocols to recognise, understand and limit the effect of subjective bias on fracture data biases during data collection.

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


## 1 Introduction

Natural fracture networks exert a strong control on the hydrogeological and mechanical properties of a rock mass, and are useful indicators of palaeostress directions. Geological models that depict the spatial distribution and nature of a fracture network rely on input data (either distributions or mean values) of fracture statistics to provide a geologically reasonable model of the subsurface. Models such as discrete fracture networks (DFNs) may be used for estimating up-scaled permeability (e.g. (Bigi et al., 2013; Min et al., 2004)) or for rock mechanics analysis (Harthong et al., 2012; Jing and Hudson, 2002), with applications, including understanding fluid flow in tight oil and gas reservoirs (Aydin, 2000) and hydrogeology (Comerford et al., 2018), and assessing rock strength for mine engineering (Mas Ivars et al., 2011). There are four fundamental methods of fracture data collection at outcrop analogues (summarised in sect. 2): linear scanlines; circular scanlines (Mauldon et al., 2001; Rohrbaugh et al., 2002)(Mauldon et al., 2001; Rohrbaugh et al., 2002); topology sampling (characterising node types); and tracing out the fracture network (window sampling). These methods are variably good at capturing the impact of orientation, censoring or truncation bias (Mauldon et al., 2001; Zeeb et al., 2013) and heterogeneity in the fracture network (Watkins et al., 2015)). We also argue that the methods also differ in the how susceptible they are to subjective uncertainties and the scale of these uncertainties.

Uncertainties in geological data can be broadly split into objective and subjective uncertainty (Tannert et al., 2007). Objective uncertainty (also called external, aleatory inherent, structural, random, or stochastic uncertainty) refers to more traditional concepts of uncertainty, such as precision or processing error in a technique or a dataset, and so can be represented through error bounds. Subjective uncertainty (also called epistemic, knowledge, functional, or internal uncertainty) arises from the mind, that is, stems from biases that affect how individuals perceive, gather and interpret geological data (Bond et al., 2015). Subjective uncertainty is common in geosciences where developing geological models typically relies on extrapolation of sparse data (Wood and Curtis, 2004), but it's magnitude and impact is difficult to quantify (Bond et al., 2015).

The collection of fracture attributes will be affected by subjective biases. Depending on the aims of the study (e.g. determining the connectivity and permeability of the fracture network; determining strength of a fractured rock mass; understanding paleostress conditions) these attributes could include the number of fracture sets, orientations, topology, trace lengths, degree of clustering, aperture and the intensity of the network (Jolly and Cosgrove, 2003; Lei et al., 2017; Watkins et al., 2015). For example, the scale of observation chosen by the user will impact the minimum fracture trace length recorded for the fracture network. Nixon et al. (2012) showed that when studying strike slip faults using bathymetry, an increase in resolution increases the recorded connectivity of the fault network. Areas of poor exposure (e.g. due to preferential erosion) requires the geologist to interpret how the fracture network connects, enhancing the scope for subjective uncertainty.

In this study we investigate the magnitude and source of subjective uncertainty in fracture data collected by linear scanlines; circular scanlines; fracture topology and window sampling. Fracture data were collected from Carboniferous rocks cropping out near Whitley Bay, Northumberland (UK) in two phases: (1) in the field where 7 participants collected fracture data directly from outcrop, and (2) two classroom workshops during which 29 participants with different levels of geological

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


training and expertise collected fracture data from field photographs. In both the field and classroom the participants collected fracture data individually and in small groups. We quantify and compare the scale of subjective uncertainty for each method we explore, and identify "problem areas" or factors which amplify the subjective uncertainty. We consider the effect of the variation on fracture statistics derived from the data collected, and propose a number of protocols to limit user bias in collaborative work.

## 2. Fracture data collection and analysis

Linear scanlines are a quick and relatively simple way of systematically collecting fracture data (Agosta et al., 2010; Bigi et al., 2015; Chesnaux et al., 2009; Guerriero et al., 2011; Ortega et al., 2006; Tóth, 2010). This method was developed in rock engineering for a quantitative description of discontinuities in rock masses (Priest, 1993), and then adopted to describe natural fracture networks (Becker and Gross, 1996; Van Dijk et al., 2000; Newman, 2005; Peacock and Sanderson, 2018). The method involves laying out a tape measure on the outcrop and measuring both the number (N) and the attributes of fractures which intersect the scanline (e.g. orientation, spacing, length above and below the scanline, aperture, type of terminations, filling or mineralization) (Priest, 1993; Priest and Hudson, 1981). To fully represent all the fracture sets occurring in a fracture network, multiple linear scanlines should be undertaken with different orientations, and the Terzaghi correction should be applied to reduce orientation bias (Mauldon and Mauldon, 1997; Terzaghi, 1965). The main purpose is to collect enough data to obtain a statistical distribution for each of the main fracture parameters rather than a mean value (Table 1). It has been suggested that over 225 fractures should be captured by one or more linear scanlines for the method to fully characterise a fracture network (Zeeb et al., 2013).

Circular scanlines provide estimates of fracture attributes based on the number of fractures intersecting a circular scanline, Nc, and the number of fracture trace endpoints, m, within a circular window (Mauldon et al., 2001; Rohrbaugh et al., 2002). The fracture density, intensity, and an estimate of mean trace length for the scanline can be calculated from the n and m values (Mauldon et al., 2001). To be statistically valid the number of fracture end points (m) should exceed 30 (Rohrbaugh et al., 2002), however, values between 20 and 30 can also be considered reliable (Procter and Sanderson, 2017). This rule defines the radius of the scanline as a function of fracture density and limits the use of the technique in areas of poor exposure and low-density fracture networks. A circular scanline is a maximum likelihood estimator (Lyman, 2003) and does not suffer from the same orientation biases observed in linear scanlines (Mauldon et al., 2001). Circular scanlines are ideal for rock masses with evenly distributed fracture attributes, but may need to be combined with other methods to give a true representation of the heterogeneity of the fracture network (Watkins et al., 2015).

Fracture topology describes a fault or fracture network as a series of branches and nodes (Manzocchi, 2002; Sanderson et al., 2018; Procter and Sanderson, 2017; Sanderson and Nixon, 2015; Laubach et al., 2018). A branch is a fracture trace with a node at each end that can be classified as terminating into rock at i-nodes (unconnected terminations), abutting against another fracture at a y-node, or crossing another branch at an x-node. Topology may be combined with circular scanlines by assessing

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


the nodes present within the circular window and using the sum of i- and y- nodes as the number of trace end points (m-value) in the circle (Procter and Sanderson, 2017). The ratio of node types is plotted on a triangular diagram (Manzocchi, 2002; Sanderson and Nixon, 2015).

Finally, window sampling is a technique all fractures within a given sample area (window) are traced out either by hand, or on a computer, and the resulting traces used to calculate the fracture statistics (Pahl, 1981; Priest, 1993; Wu and D. Pollard, 1995). This technique is often utilised to analyse remote-sampling data such as aerial photographs (Healy et al., 2017), Unmanned Arial Vehicle (UAV) images (Salvini et al., 2017), bathymetry (Nixon et al., 2012), or satellite imagery (Koike et al., 1998), as well as in outcrop studies (Belayneh et al., 2009). It has been suggested that a minimum of 110 fractures are sampled to be able to statistically describe the fracture network using window sampling (Zeeb et al., 2013).

Using the methods above fracture parameters can be collected which then enable calculation of key fracture statistics, for example, trace *length* (mean and distributions), fracture abundance (*Intensity* and *Density*), and connectivity (Summarised in Table 1).

Trace length, and trace length distribution are key fracture parameters for DFN simulations e.g. in simulating fracture-hosted fluid flow. Trace lengths may be measured directly with the linear scanlines, or estimated using the circular scanline method. Challenges in the determination of trace length for individual fractures include: the scale of observation used to collect the data (Zeeb et al., 2013); how fracture intersections are classified (Ortega and Marrett, 2000); and the fracture fill properties (Olson et al., 2009). Mean trace length is a commonly used fracture statistic and is useful where the fractures in a network are evenly distributed (Mauldon et al., 2001). However fracture modelling usually uses a statistical distribution representative of the fracture length population rather than the mean (Neuman, 1993). Trace length distribution, obtained from measuring individual fractures, should be used when investigating sub-surface fluid flow or characterising spatial variations in fracture trace length (Watkins et al., 2015). We investigate the impact of subjective bias on mean trace length (all methods) and the range of reported trace lengths for linear scanlines and window sampling and trace length distribution for window sampling.

The characterisation of fracture networks and comparison of techniques is greatly confounded by inconsistencies in terminology: for completeness we lay these out here. Because fractures may be sampled using techniques which are either 1-dimensional (scan-lines, boreholes), 2 dimensional (maps, surface exposure), or 3-dimensional (rock volumes), numerous different methodologies and terminology has arisen to characterise the abundance of fractures in a network. One of the most widely used method to characterise a network is to define the number of fractures (N) normalised to line length (L), sample area (A) or sample volume (V) depending on the dimension of sampling. In the literature this statistic is either termed *fracture intensity (I)* or *fracture frequency (f)* (Sanderson and Nixon, 2015). For linear scanlines *fracture spacing* can be regarded as the inverse of *fracture intensity* for a single set of sub-parallel fractures (Sanderson and Nixon, 2015). Fracture abundance within a network may also be expressed as the total trace length per unit area (Dershowitz and Einstein, 1988; Rohrbaugh et al., 2002). This statistic is either termed *fracture intensity* (Sanderson and Nixon, 2015) or *fracture density* (Nixon et al., 2012; Zeeb et al., 2013). One attempt to simplify the use of terms is to use the $P_{xy}$ terminology as defined by (Dershowitz and Einstein, 1988) where x denotes the dimension of the sampling region (1 = line, 2 = area, 3 = volume) and y donates the dimension of the feature (0 = number, 1 = length, 2 = area, 3 = volume). For the purposes of our study we use the term *fracture intensity* (I) to refer to number of fractures per line length (P10, for linear scanlines) or fracture length per unit area (P21, for circular scanlines), and we use *fracture density* for number of fractures per unit area (P20).

It is also important to understand how individual fractures relate to each other; particularly how the individual fractures connect, and hence contribute to the strength of, or fluid flow through, the rock mass. The number of connections on a fracture trace ($C_L$) is a commonly used measure of connectivity (e.g. Manzocchi, 2002). However, a fracture network consisting of only y and x nodes could have different $C_L$ values depending on the fracture intensity (Sanderson and Nixon, 2015). It has been suggested that it is better to either consider the average number of connections per branch ($C_B$) (Ortega and Marrett, 2000) or the proportion of connected nodes (Pc) (Sanderson and Nixon, 2015). In our study we use the proportion of connected nodes for circular scanline and window sampling. To measure connectivity in linear scanlines the proportion of fracture trace end points which are connected are considered.

## 3. Study methods

### 3.1. Study area

The field site is located in the Northumberland Basin, just north of Whitley Bay, NE England (Fig. 1). The Northumberland Basin is a 50 km wide, ENE-WSW trending half-graben formed during mid-late Carboniferous extensional reactivation of the underlying Iapetus Suture (Chadwick et al., 1995; Johnson, 1984). The stratigraphy consists of thinly (cm - dm) bedded sandstones, siltstones, shales, seat earth, and coals of the Middle Coal Measures (Westphalian B). At the field site the easily accessible and well exposed wave-cut platform clearly exhibits N-S striking faults and joint sets cross cutting E-W trending faults and joint sets, and populations of sub-vertical joints (>75° dip).

### 3.2 Fracture data collection procedure

Six linear scanlines were set up by laying out a tape measure on sandstone beds, both in map and cliff section (Fig. 1C). Participants were asked to identify for each fracture: a) the intersection distance along the tape and b) the length and termination (into rock, abutting against another fracture or not seen/obscured) of the fracture either side of the tape. Eight circular scanlines were drawn with chalk directly onto the sub-horizontal bedding planes of three separate, decimetre thick, medium grained sandstone beds (Fig. 1D). A N-arrow and NS/EW lines were drawn onto the circle to aid observation. participants counted the number of intersections with the circumference (Nc). Following the methodology of Procter and Sanderson (2017), participants were asked to identify the number of i-, y- and x- nodes within the circles. Finally, window sampling was conducted by tracing out the fracture networks on photographs of the circular scanlines in the workshops. Our study did not aim to collect sufficient fractures to represent the fracture network at the field site, and the tested scanlines were not designed to be statistically representative.

Fieldwork was undertaken by 7 participants (labelled A-G) in July 2018 with fracture data collected using field notebooks from 7 circular and 4 linear scanlines (Table 2). There was no particular guidance as to how the participants collected the scanline data, but no more than one person or one group collected fracture data from a scanline at any one time, so as to avoid influencing the data collected by other participants. For the same reason, participants did not annotate or disturb the rock or scanline. Orientation and aperture data were also measured in the field, but they are not included in this study because they generally are not included in circular scanline methods and cannot be measured from field photographs in the workshops. Three of the fieldwork participants also completed the workshop tasks (Participant C = Participant 8; Participant D = Participant 10; Participant G = Participant 11).

Workshop 1 (WS1) was held in September 2018 in Glasgow, with 11 participants (labelled P1-11). Workshop 2 (WS2) was held in October 2018 in Rome with 18 participants (P12-29). Participants were recruited from the authors' research groups (the Faults and Fluid Flow research group within the Centre for Ground Engineering & Energy Geosciences at the University of Strathclyde and the Tectonics and Fluid Chemistry Lab of Earth Science Dept. at Sapienza) as well as colleagues from their departments: participation was voluntary and all data were anonymised for analysis. Each 2-part workshop lasted 3 hours: in the first part, participants worked individually to complete 3 circular and 1 linear scanline, and in the second part, worked in small groups to complete 2 circular and 1 linear scanline (Table 2). Participants were provided with A3 colour photographs of the scanlines. WS1 participants were encouraged to annotate these with the observed fracture intersections and interpreted termination type, whereas WS2 participants were specifically asked to trace out the interpreted fracture network (i.e. to undertake window sampling). Both workshops enable us to investigate the impact of subjective bias, however, the fracture maps from WS2 enable us to examine the impact on window sampling along with investigating the root cause of differences in how participants classify nodes.

To examine the effect of geological experience on subjective uncertainty, participants were asked to indicate their level of geological training, familiarity with geological fieldwork, and their level of experience collecting fracture data (summarised in Table 3, questionnaire provided in Supplementary Information, S1). In the workshops, a small number of participants (Participants 2, 5, 24 and 28) consistently reported anomalously high Nc values compared to the node counts. Three of these participants (Participants 2, 5 and 28) had no formal geological training or experience in geological fieldwork and fracture data collection. It is possible that these participants only considered fractures that intersected the edge of the circle in their interpretation (neglecting fractures within the circle that do not intersect the circumference), introducing a different source of subjective error.

### 3.3 Post-workshop analysis

For the workshop data we digitised the interpreted fracture traces and node classification for all participants who traced the networks (see Table 2) using ArcGIS. Individual fracture trace lengths for all scanlines, and for linear scanlines the distance along the scanline that each fracture intersected were exported as 'Arcmap unit' lengths. These lengths were then scaled to the field to enable comparison of the fracture statistics. In some cases, the counts of Nc or node types reported by participants differed from the count indicated on the worksheet (see S7). In these cases, to be consistent with field data collection we take the value reported by the participant. Digitised networks from Circle 8 were used as a case example to (a) construct heat maps of point density for Nc, i, y-, x- nodes, and line density for fracture traces, and (b) identify areas within the circular scanline with greatest variability in fracture interpretation (trace, node type, termination etc.).

Fracture statistics, which were derived using the fracture parameters collected as described above, were then investigated for field and workshop participants. We report on the impact of subjective bias on the following fracture statistics; *fracture intensity (I), fracture density (d),* the *connectivity* of the network *(Pc & Pf), mean trace length (Tl),* and *trace length distributions (tl).* Statistics are calculated using the equations outlined in table 1.

In theory, each of the scanlines have a 'true' value for each of the fracture parameters (number and type of fracture intersections and terminations, i.e. Nc, Ni, Ny and Nx) and, consequently, the fracture statistics derived from these parameters (intensity, density, connectivity and mean trace length). In this paper we are not interested in defining that 'true' value, rather we wish to explore the ranges in reported values from different participants, showing the scale of subjective bias in the fracture data collected, and the factors that affect this range. In this study we wish to define the uncertainty, or level of variability, present in fracture data collection and the statistics which are derived from this data. We therefore report the range and mean for the fracture data collected. First, we present the effect of subjective bias on the 'raw' fracture parameters, and explore the consequent uncertainties in the derived fracture statistics in Sect. 4.

## 3. Results

### 3.1 Linear Scanlines

There is a reasonable amount of consistency between participants' reported number of fractures crossing each scanline, however the reported trace length data are much more variable (Table 4). This pattern is repeated between the field and workshops. The minimum reported trace length is consistent, for example participants in Line 6 range from 0.02 to 0.23 m (See S5). However, the maximum reported trace length is highly variable, e.g. for Line 6 it ranges from 0.25 to 0.72 m. It is clear that participants disagree in how each individual fracture terminates. For example, for one fracture intersecting Line 3, Participants G + F interpreted that after 8.0 meters the fracture terminated against another fracture, whereas Participants C + D felt that it terminated in an area of no exposure after 22.0 m (S5). No correlation was observed between the number of fractures intersecting the linear scanline and the range in trace lengths, either in the field or in the workshop.

The fracture traces interpreted on photographs in the workshops help us to understand the underlying controls on this subjective bias. We examined the fracture traces of Line 6 in detail and the interpreted fracture networks display considerable variability (Fig. 2). All participants identified two large fractures roughly 1/3 and 2/3 of the way along the scanline, however there were large differences in the way people interpreted the first third of the scanline: Participant 28 doesn't identify any fractures, whereas Participants 10 and 14 interpreted 3 and 10 fractures respectively. In this section the fractures are partly obscured by water and have a thin fracture trace. These 'hairline' fractures are also present in other parts of the scanline and in all cases contribute to areas of uncertainty. Another section that led to uncertainty was the feature trending at a low angle to the scanline half way along: only 14 out of the 29 participants interpreted this as a fracture. This trace is also the longest fracture that is reported on the scanline, with the other long fractures being censored by the edge of the picture. The main source of uncertainty in measuring linear scanlines on a photograph is therefore the decision of how a fracture terminates, and hence how long to report the fracture trace.

### 3.2 Circular Scanlines: Topological sampling and fracture mapping

The results of circular scanlines and topological scanlines have been reported together as participants defined all nodes within the circle in both the field and the workshops. For the circular scanlines the number of fracture terminations (m), although not explicitly discussed in this section, is equivalent to the total number of i- and y-nodes.

The number of fracture intersections with the edge of a circle (Nc) showed relatively low variability between participants in the field (Fig. 3). However, the variability in number of reported nodes is greater, and is largest for y-nodes. This pattern is repeated for the workshops (Fig. 4; Table 5). When high variability was observed in a particular topological parameter (e.g. y-nodes), it was not necessarily observed in the counts for other parameters (e.g. Nc) in the same circle. For example, the number of y-nodes interpreted in the field were highly variable for Circle 6 (7 to 27), even though this circle had the lowest range in values for Nc (6 to 9). In this case, clearly all the participants were observing almost the same fractures *intersecting* Circle 6 (i.e. subjective bias for Nc is low). Participants differed though in how they then observed and classified fractures *within* the circle, leading to a greater range in the number of fracture intersections. The consistent observation, is that subjective bias affects node counts more than Nc counts, but that the degree of variability is dependent on the circle.

No single circular scanline was particularly prone to subjective bias across all the fracture parameters studied. For example, compared to other circular scanlines, the range in data collected from Circle 5 is small for Nc, i and y-nodes, but is one of the most variable for x-nodes. In contrast, the range in data collected from Circle 3 is small for x- and y-nodes, but is one of the most variable for i-nodes and Nc (Table 5). The trends are seen in both field and workshop data.

Although individual circles displayed considerable variability between participants, participants often remained internally consistent (Fig. 3 and 4). For example, Participants A and C, or Participant 2, tended to report lower counts for all circles than Participant G, or Participant 13. That said, when Participants C and D repeated the data collection for the same scanline in the field, there was some discrepancy within the repeat data (Fig. 3), although this is far lower than the discrepancy between participants. The relative proportion of specific node classification (e.g. y-nodes) for both individual participants and groups also displayed consistency between circles (Fig. 5). For example, Participant 11 consistently recorded more y-nodes when compared to other participants, while 5 and 21 tended to record more i- and x- nodes.

In general, the scale of uncertainty (the range in reported values) in the workshop data is greater than field data as indicated by a wider range in reported values. Overall, the number of fractures reported was higher in the field data than the workshop data. For example, the reported number of fracture intersections in Circle 3 in the field (fig. 3) ranged from 19 (Participant C) to 30 (Participant B), whereas from the workshops ranged from 14 (Group 8) to 23 (Group 6) (Fig. 4). Similarly, the number of y-nodes is generally higher in the field and the rage in vales for each circle is less extreme – e.g. in the number of y-nodes Circle 5 ranged from 28 (Participant C) to 47 (Participant D) in the field (Fig. 3C), and from 4 (Participant, P2) to 41 (P13) in the workshops (Fig 4). It is possible that in the field participants can observe more fractures (e.g. the hairline fractures in fig 2) resulting in more consistency in their reported values.

### 3.3 Window sampling

For circles where window sampling was used the number of recorded fractures varied considerably (Table 6), with the maximum variation in range within the Circle 5 (13 to 56). The maximum trace length reported by all participants remained fairly consistent. However, considerable variability in trace length distributions was observed between participants (Fig. 6), with the number of small fractures recorded across all scanlines displaying the most variability. For example, the number of fractures below 0.2 m recorded for Circle 8 ranged from 7 to 41, which represents 36.8% and 75.9% of the reported fractures for both participants. While the number of small fractures recorded by participants varies between circles, whether a participant records a high or low relative percentage of small fractures remains consistent. For Circles 8, 5 and 1 Participant 3 consistently record a high percentage of small fractures and Participant 24 is consistently records a low percentage of small fractures (Fig. 6a). In short, Participants either consistently record the small fractures in a network, or consistently do not record the small fractures in a network. For trace lengths longer than around 15-20% of the diameter of the circle, the shape of the distributions remains consistent across all participants indicating that the larger scale fracture network is well classified.

#### 3.1.3 The effect of working in groups

No clear differences can be seen between data collected individually or as groups for either circular scanlines or window sampling (Table 6; Fig. 6b). Although the group circles have lower y-counts and higher mean trace length values, the differences are not enough to be confident that this is due to working in groups rather than differences in the fracture network. That said, groups generally reported more complex fracture networks with a higher reported number of small fractures. When working as groups that included a naturally detailed and naturally less detailed participant, the results tended to be more detailed: compare participants 2 and 11's recorded values when working individually or together as Group 3 (S7).

Overall there is less scatter observed in data collected as a group, however, due to the difference in the number of data points between individual and group scanlines it is not possible to know if this is an effect of a limited data set or a true impact of working in groups. Similarly, to when working individually, groups remain internally consistent in the number of small fractures recorded (Fig. 5b) and the relative percentage of recorded node types (Fig. 4b). For example, Group 12 reported consistently higher i-node counts compared to Group 7, who instead reported more y-nodes.

### 3.4 Areas of increased uncertainty: A case study using Circle 8

To highlight potential causes of differences in interpretation, Fig. 7d compares different interpretations of fracture trace and nodes in three particular problem areas from end-member participants 11, 18 and 21, who reported high, medium, and low node counts respectively. Area 1 is well exposed and contains several intersecting fractures. The nature of the connections was interpreted differently by each participant. Participant 21 interpreted only the major fractures coming into the junction, and depicted the fractures interesting in a star-like formation. Participant 18 interpreted a standard x-node, with a second larger fracture terminating against the NE-SW trending fracture (y-node), and also notes an E-W trending fracture linking the two major fractures and cutting the third (three x-nodes). Participant 11 differed from Participant 18 by interpreting the NE-SW fracture trace as being offset by the NW-SE fracture, such that the x-node interpreted by Participants 21 and 18, was instead interpreted as two y-nodes. Area 2 is a complex intersection of a number of NW-SE fractures with part of the photographed exposure obscured by shadow (a clear limitation of interpreting the scanline from photographs rather than in the field). Participant 21 did not interpret the fractures obscured by shadow, whereas Participant 18 did. Participant 11 depicts a number of smaller fractures which Participants 18 and 21 do not identify. Area 3 is an intersection of two large fractures which is obscured by a coarse sand infill. Both Participant 18 and11 interpreted the obscured connection as a simple x-node, whereas Participant 21 felt that the fracture bifurcated to frame the area of no exposure. Participant 18 and 21 interpreted the other fully exposed connections similarly (although Participant 21 does not depict a fracture to the south of the sand fill), whereas once again Participant 11 identifies several additional smaller and complicated fractures and fracture connections, particularly y-nodes. In each case it appears that participants effectively ‘self-censored’ their data according to their ‘preferred’ minimum trace length, and had different approaches to areas of shadow or obscured outcrop. The different geometry of the interpreted fracture intersections would result in significant differences in interpreted fracture development history.

When analysing the node classifications and interpreted trace lengths for all circles it was found that in many cases the fracture networks depicted or interpreted were not viable: in other words, there were undefined nodes or intersections which had a non-compatible number of branches entering the node (e.g. 4 nodes for a y-node or 5 for an x-node). Occurrences of these undefined or floating nodes were more common in WS1 than WS2, perhaps because WS2 participants were specifically asked to draw out the fracture network on their photographs.

### 3.5 The effect of working in groups

Large variability in the number of reported fractures in the field was also seen when linear scanlines were undertaken as pairs, for example for linear scanline 3 counts ranged from 21 (Participant C + D) up to 30 (Participant A + B). The groups are obviously made up of participants who have different ‘eye for detail’. When working individually Participants C and D both recorded low fracture counts, while Participant B recorded the joint highest. There is a suggestion in the data that when working as pairs, groups tended towards the more detailed member, for example Participant F recorded the lowest fracture count when working individually, however, in a group with Participant G recorded a higher than average fracture count.

Conversely, no clear differences can be seen between data collected individually or as groups for either circular scanlines or window sampling (Table 5; Fig. 5b). Although the group circles have lower y-counts and higher mean trace length values, the differences are not enough to be confident that this is due to working in groups rather than differences in the fracture network. That said, groups generally reported more complex fracture networks with a higher reported number of small fractures. However, similar to the linear scanlines, when working as groups that included a naturally detailed and naturally less detailed participant, the results tended to be more detailed: compare participants 2 and 11's recorded values when working individually or together as Group 3 (S7).

Overall there is less variability observed in data collected as a group (Fig. 6), and groups remain internally consistent in the number of small fractures recorded (Fig. 5b) and the relative percentage of recorded node types (Fig. 4b). For example, Group 12 reported consistently higher i-node counts compared to Group 7, who instead reported more y-nodes.

### 3.6 Time taken to collect data

Participants were broadly internally consistent in the time taken to complete their tasks (Fig. 3 & 8). For example, C and G tended to take longer than A or D in the field, and in workshop 2, Participant 29 consistently took longer than Participant 25.

In the field it generally took participants longer to count more nodes, however, the correlation is weak and dependent on both the circle and the participant (Fig. 8a). This was not the case in the workshops where no correlation between the time taken to record, or the variability in, the number of reported fractures was observed (Fig. 8a). Both the time taken and magnitude of variability was considerably higher in the workshops compared to the field. For example, Circle 5 took participants between 1 and 17 minutes in the workshop and 2 minutes 21 seconds to 4 minutes 26 seconds in the field.

Window sampling, which was undertaken in WS2, took longer than circular scanlines for the same circle in WS1, however, this difference is small. While it took 1.3 to 3.2 times as long to record Nc values, the time taken to undertake topological sampling within the circle is comparable for both individual (0.85 to 1.6) and group (0.95 to 1.05) circles. This shows that although circular scanlines are often suggested as a quick way of gathering fracture data, it does not take significantly longer to trace out the fracture network. This observation is important and suggests a similar amount of data could be collected using both methods.

### 3.7 Experience

The relationship between experience and the number of node counts has a large amount of scatter (Fig. 8b). Generally, participants with less experience undertaking geological field work or collecting fracture data counted fewer nodes than more experienced participants, however the trend is very weak. Perhaps counter-intuitively, experience does not reduce the time taken to collect fracture data (Fig. 8b). However, for node counts the fastest experts are still notably slower than the fastest inexperienced Participant. There is no indication that more experienced participants characterise more detail than those with less geological training or experience. It is possible that Participants with experience in fracture analysis will consider the

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


connections they observe, whereas beginners will draw the traces that they see without considering the implications of those connections (i.e. implied cross-cutting relationships).

#### 4. Effect of subjective bias on the derived fracture statistics

The variability in the collected fracture parameters will affect the derived fracture statistics in different ways. No particular equation for the calculated statistics (Table 1) has a relationship which makes that statistic sensitive to subjective bias in a particular fracture attribute. In order to identify which fracture statistics are most susceptible to subjective bias, we discuss and compare the results from all methods in terms of the relative ranges of values. The key observations and trends are summarised in Fig. 9. For linear scanlines no correlation between the number of observed fractures and fracture trace length was observed, e.g. Participants B and G both recorded 10 fractures intersecting Line 1, however, the derived mean trace lengths were 0.62 m and 0.25 m respectively. This is in contrast with window sampling, mean trace length decreases as fracture count increases ($R^2$ = 0.79 for Circle 8, see S8), and circular scanlines where mean trace length is a function of the number of fractures intersecting and terminating within a circle. Mean trace length derived from window sampling was consistently lower that that derived from circular scanlines of the same circle, for example mean trace length for Circle 5 derived from window sampling ranged from 0.19 to 0.46 m (S8).

*Fracture density,* which is calculated for circular scanlines and window sampling, was consistently the most variability statistics between participants. A higher value for fracture density was obtained using window sampling compared to circular scanlines for the same circle, however, the range in values are large in both. This is shown in Circle 8, where window sampling derived fracture density ranged from 22.9 to 68.8 F/A compared to 1.9 to 41.4 for circular scanlines.

Across all methods, *fracture intensity* has most consistently the smallest range in values, i.e. is the most certain statistic and displays the least variability for window sampling. If fracture spacing is used to calculate *fracture intensity* from linear scanlines considerable variability will be seen in the data. This derives from the fact that while the minimum fracture spacing is consistently small across al scanlines, the values for the maximum spacing varies considerably. The *connectivity* of the network (*percentage of connected fractures, Pf*) is highly variable for values derived from linear scanlines, however, the percentage of connected branches (*Pc*) is robust statistic for topological sampling.

The overall trends, presented in Table 7, suggest that although subjective bias impacts all data collection methods, window sampling generally displays less variability in derived statistics and appears to be least effected. *Fracture intensity* represents the most robust statistics, with *Mean trace length* and *fracture density* both displaying considerable variability between participants. The connectivity of the network was found to be robust for topological sampling, however, displayed considerable variability when derived from linear scanlines. In addition to the this we find that for workshop linear scanlines and window sampling, where fractures were traced out, it was possible to understand the causes of the variability seen in the fracture statistics.

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


## 5. Discussion

Subjective bias in fracture data collection has implications for the validity or reliability of the models that the data informs, such as the derived fluid flow parameters, rock strength characteristic or paleostress conditions. Here, we explore these implications, and, drawing on the participant's discussions following the workshop and field activities, explore potential reasons for the observed differences and trends.

### 4.1 The effect of user bias on scanline validity

As for all sampling, scanlines must contain enough datapoints to be statistically valid: the required number of datapoints depends on which aspects of the fracture network that are being investigated. However, our data demonstrate that in addition to the fracture network characteristics, the required scanline size (length of a linear scanline, circumference of a circular scanline or area of a window sample) is also dependent on the user collecting the data.

Our participants clearly observed different numbers of fractures in the same scanline (Table 6, Fig. 2), thereby affecting the size that a scanline would have needed to be to capture a representative sample of that network. As an example Zeeb et al. (2013) suggest that a minimum of 225 fractures are sampled for linear scanlines and 110 fractures for window sampling. If this is true, then for Line 3 where participants reported between 1.4 and 2.5 fractures per metre, the cumulative length of scanline would need to be between 90 and 160 m long. The number of fractures in Circle 5 reported for window sampling ranged from 13 to 56, which means between 2 and 9 circles of this size would need to be analysed to statistically represent the network. The innate variation between how participants view the fractures therefore results in significantly lengths of scanline or numbers of circles.

The appropriate circle *radius* is also dependent on the sampling characteristics of the person undertaking the work. For circular scanlines it is widely agreed that a minimum of 20-30 fracture terminations within a circle is appropriate to derive fracture statistics or undertake topological sampling, and the circle radius must be adjusted to capture enough fractures or fracture terminations (Procter and Sanderson, 2017; Rohrbaugh et al., 2002). Figure 10 shows the proportions of valid (capturing >30 terminations) and invalid (capturing <20 terminations) results for the circular scanlines. Out of the 29 participants that collected data from Circle 8 in the workshops, 12 identified over 30 fractures and so report valid results, another 8 collected over 20 fractures and their results are potentially valid, whereas 9 valid reported fewer than 20 fractures and so the statistics derived from their sample may not be unrepresentative. Since the number of fractures identified in the field is generally higher than in workshops, a greater proportion of field participants reported sufficient terminations within the circle to be statistically valid. For example, all field participants report valid data for Circle 4, whereas only 3 of the 9 groups in the workshops do.

In this work, the location and radius of all scanlines except C6 were selected by Participant G/11, who tended to be more detailed than other participants. This participant recorded enough terminations to class their data as valid for all circles studied. Therefore, this participant chose a circle radius appropriate to the level of detail to which *they* identify and characterise

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


fractures, but which is not appropriate for other less detailed observers. The effect is demonstrated in Fig. 11 which shows a synthetic fracture set which has been interpreted by a less detailed participant (A) and a more detailed participant (B). A statistically valid circular scanline (>30 fracture terminations) is drawn onto the interpreted network and the resulting differences in the fracture topology and the fracture statistics shown (Table C inset). For this example, for the scanline to be statistically valid, its radius must be 3 times larger for Participant A than Participant B.

How detailed a fracture network is interpreted to be therefore affects the derived fracture statistics. The fracture network interpreted by Participant B has more y-nodes, but similar counts of Nc, i-nodes and x-nodes (Fig. 11c). As a result, the connectivity of Participant B's network is higher (the other values - intensity, density and trace length - are very different because of the differences in circle radius for a valid sample). For our data, if invalid data are disqualified (i.e. removed from the dataset), the maximum trace length and density are more affected than the fracture intensity and connectivity. For example, the calculated maximum trace length for Circle 8 decreases from 2.88 to 0.92 m, and the maximum density for Circle 5 decreases from 46.5 to 12 f/A.

### 4.2 Bringing together data sets

It is clear from our data that participants have an inherent personality characteristic: end-members could be classified as either detail-oriented, e.g. picking up more small fractures and spending longer interpreting intersection geometry, or as focused on collecting a larger volume of data, e.g. covering larger areas and collecting more circles in the same amount of time. The presence and scale of subjective bias and potential influence on fracture data means that caution must be taken when comparing data collected by different geologists, or using different methods. Differences in interpretation could occur if large areas are mapped by different people, with certain areas being classed as having a higher/lower fracture density and trace length purely based on subjective bias. This is also relevant when either using, or comparing collected field data with the published literature.

### 4.2 Causes of subjective bias: fracture network characteristics

**Areas of limited exposure**, where participants are required to make an interpretation of how the fracture network connects. In the qualitative discussions following WS1, several participants reflected that where exposure was limited or obscured, they did not attempt to interpret where the fracture went nor the type of fracture intersection. Other participants, however, did interpret the network despite these difficulties, which increased the number of nodes and decreased the number of illogical 'floating' nodes. Clearly some felt it was most appropriate to interpret in the face of great uncertainty, so as not to discount the node(s), while others felt that this would be over interpreting. Both have sound reasoning, but will result in very different outcomes in terms of fluid flow or rock strength.

In some cases, uncertainties could easily be overcame in the field, for example where a fracture was obscured by shadow or seaweed. Some field participants described feeling for a fracture with fingers or pencils when obscured or extrapolating the exposed fracture traces, combined with other trends observed outside the unexposed area. Such 'exposure bias' is recognised when studying fault zones; by their nature, the fault rocks are preferentially obscured and therefore good continuous exposure of fault zones is very rare (Shipton et al., n.d.).

**The scale of observation:** In the workshops, participants were provided with a 2D 'birds eye' type view of the full circle. In the field, only the tallest geologists will be able to observe the full circle, with all others limited by to a smaller field of view. But, in the field, the participant can crouch down and get their eye in to the detail within a complex fracture network. This is most likely why for the same circular window, more nodes were counted in the field than in the workshop (Fig. 10). The impact of subjective bias on the required length of linear scanlines, radius of circular scanlines and area of sample windows will have particular consequences in areas of limited exposure, where a data volume orientated user may not be able to collect enough data to statistically represent the fracture network.

**What features counts?** Some participants focussed only on more pronounced fractures, ignoring, for example, smaller subsidiary fractures, closed or filled fractures, or thin 'hairline' fractures present in the scanline. This was particularly the case if there was a large or clear fracture. As one participant exclaimed "What do these tiny things matter if you have a massive fracture?". However, this was not what all participants thought, with other participants raising the importance of the spatial distribution of small fractures. In some cases, and particularly in the workshop, it was not clear whether faint or fine features represent surface processes such as wear, weathering, or salt precipitate from seawater. Similarly, following group work, some participants reflected that there were differences in whether the participant classed a jog in the fracture as a termination, or a slight side step of a continuous fracture. During WS1 discussions is was suggested that data may be able to be compared is a cut off was applied to the data, however, no consensus could be agreed upon with participants believing the resulting network would be flawed because (i) it would not be an accurate representation for flow and/or rock strength and (ii) more attention would be paid to simpler, larger, and more isolated structures that could have almost no flow or mechanical significance.

**Field vs photograph:** We suggest field work is preferable to field photographs due to being able to ground truth areas of uncertainty. As touched on already, in the field the participants can zoom in to more complex fractures, remove obstructing material, adjust so that something isn't in shadow, physically feel for the fracture, check if a feature rubs off, or if it is continuous into another plane of the outcrop. With the advent of digital image analysis techniques and UAV technology, it can seem preferable to perform digital fracture mapping, however, issues around hairline fractures, or potential weathering features, vegetation obscuring the network can be easily remedied in the field.

**Group work**: We find that working in groups increases the detail of observation, and so reduces the spread in results. The reported data tends towards the more detailed member of the group. A participant explained "[when we started working together] I very quickly …realised that [person X] cares about tiny features, so, together we cared about tiny features…but I was aware that if I was working on my own, I would have done it differently". When working together, they explain "I didn't find we were talking about 'does this fracture count?', instead we were discussing whether something was a Y-Y or an X, or where exactly a fracture goes or where it terminates and so on". In addition to the actual process of working together, group work might also be effective because, as one participant articulated "the very knowledge that you are working with someone changes your approach. You want to engage together and so you need to defend or explain your choice, which makes you more alert to what you are doing and why". The participants felt this slowed down the data collection process, however, this was only clearly observed in WS1, with the time taken for the group line and circles being comparable to those for individuals for WS2 (Fig. 3, Table 6). Not all groups discussed the interpretations together, instead opting to divvy up the window or scanline and work separately and combine results at the end, potentially losing some of the benefit of group work.

### 4.3 Causes of subjective bias: Differences in how people see and interpret fracture networks

While there is considerable amount of variability in the data sets, we find that individuals show internal consistency (Fig. 3 and 4). That is, individuals exhibited their personal characteristics or traits in the data that they interpret: they were either more detail focussed, or focussed on the volume of data collected. Thus variability in data that is collected by a single person may be reliably compared for changes in fracture statistics, however, care needs to be taken when comparing results from different participants. It is important to consider if you are working with a 'detailed' participant who will always likely pick up on small fractures and subtle features, or if you are working with a geologist who is more likely to miss these features but who may be able to collect data from a greater number of circles in the same amount of time.

It is interesting to consider why people are different but internally consistent. Probably, they consciously or subconsciously construct their own protocols around how the data should be collected, and what features should or should not be included. These will be shaped by:

(a) Practical and physical factors such as the quality of their eyesight, whether or not it is easy for them to repeatedly crouch down to get a closer view and stand up to move around, spatial co-ordination which will affect the ease with which they cover the scanline, and the time available to gather the data.

(b) Cognitive factors such as their tendency to over interpret, or what they think matters; if smaller or filled fractures are perceived not to matter, they won't be looking for them, and so won't necessarily notice them. In this way, the participant's mental model (Shipton et al n.d.) of the processes that they are measuring may in effect obscure or censor the network that they observe. The mental model may also be influenced also by the purpose of the survey and what kind of application the collected measures are intended for.

### 4.4 Reducing subjective bias

This paper highlights the contribution of personal bias in adding to uncertainties in the data that geologists collect. We encourage critique of the data collection process and potential uncertainties when collating or comparing fracture statistics from different field studies. Drawing on our results, we propose the following approaches to assess, reduce, and report the potential subjective bias in the data that geoscientists collect. These recommendations are not only relevant to field geoscientists, but also to modellers who use their data and geologists who use fracture sets to infer paleostress conditions:

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


- While all methods of collecting fracture data are susceptible to subjective bias, window sampling is the least effected. Further, the approach does not take much longer than topology sampling (<1.6 times as long individually and comparable as a group), thus, we recommend that where possible the window sampling approach is adopted. In addition to this the fracture network should be traced out regardless of which approach is adopted (circular, window, linear). This could be done either on a printed photograph/tablet or with chalk on the outcrop.
- When deciding the dimensions of your window or scanline, the minimum number of moderate-scale, obvious, fractures should be captured (i.e. when using a circular approach, the radius should capture 20-30 terminations of the major fracture sets), with the small fractures still recorded. Because all fractures larger than 10 to 15% of the circle radius are typically well defined, all data above this can be confidently compared between geologists with different fracture judgements.
- Consider what the data is for. For example, in fluid flow studies it is vital that information for all connected fractures are included in the data set, in which case, the location of small fractures that contribute to the network becomes key. However, our findings show that the inherent biases of the data gatherers determine if small fractures are collected. The spatial distribution, not just the relative percentage, of fracture terminations within a network should be assessed and recorded when reporting fracture statistics.
- If conducting collaborative fieldwork, whereby different individuals are collecting data from different areas, the team must first agree on a unified approach and classification system, the process of determining sample location and dimensions, and what to do when, e.g. a particular fracture intersection is obscured. It is important to characterise the way participants differentiate fracture terminations and distribution of reported trace lengths. The data collected at this site suggest that the position on the node-triangle is a factor of how detailed a user is, with detailed users tending to select a larger percentage of y-nodes compared to x- and i- nodes. This is likely due to the small fractures in this network being concentrated at fracture intersections and in the case where small fractures are isolated a detailed user would instead report a higher percentage of i-nodes.
- When working in groups, confer out loud the rationale or reason rather than divvying up tasks. This allows for the identification and reconciliation of differences in interpretation, thereby reducing the potential for subjective bias.
- Finally, but most importantly it is vital that we report our own biases and methods used to reduce bias in the field reports, to enable replicability and comparison of studies.

In principle, if these recommendations are followed, it would be possible to 'correct' networks captured by either 'more detailed' or 'more scanline' users to make them comparable. In the case where small fractures are vital for fluid flow or rock mass stability a large-scale fracture network collected by a less detailed user could be populated with small fractures typical of the network which are collected by a more detailed member of the group. Similarly, if small fractures are not important (e.g. if they are unconnected) then small fractures could be removed from the data sets of more detailed users and the large scale fracture networks compared. It is possible to assess the importance of the networks either through visually assessing the distribution of small fractures or comparing where different members of a group plot in node-triangle space. In a network

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


where small fractures are evenly distributed and not necessarily connected to the network then a detailed participant will identify a higher percentage of i-nodes then a less detailed user. In the case where small fractures are concentrated at fracture intersections and contributing to the network a detailed participant would instead notice a larger number of y-nodes. Combining this information with the trace length distributions enables an assessment to be made as to what type of correction is required to make different users data comparable.

## 5. Conclusions

This work has found that subjective bias has a considerable impact on fracture data collected from the same scanline by multiple participants. Although considerable variability is observed between participants, a degree of internal consistency in the number of fractures and node classifications is observed for each participant. This variation in reported fracture parameters affects the manner by which the data needs to be collected. A detailed user can collect data using a smaller radius circular scanline and require a smaller outcrop to characterise a network compared to a user who does not record as many small fractures. The number and trace length of fractures reported, and hence derived fracture statistics, have no correlation to the level of experience or time taken to complete the scanline. We suggest instead that the underlying control is the individual's personality or inherent level of detail.

The impact that the variability of reported fracture attributes has on derived fracture statistics can be large, with trace length and fracture density the most susceptible to subjective bias. When possible, it is important that fracture data are collected in the field, where the type of connections present can be assessed. Because the major fracture sets are captured by all participants it is important to record not just the number of terminations and individual trace lengths, but also where in the scanline/are the values recorded. We suggest that the network is always drawn, either onto printed field photos or using a tablet, and the trace length distribution and network topology are both reported and considered.

When working together, comparing collected fracture data to literature values, undertaking fracture analysis to understand the geological evolution of an area or using fracture data to populate DFN models, it is vital to have an appreciation of the level of detail used during fracture data collections. It is also important to understand the relative importance of small fractures to the network and the impact subject bias can have on capturing this. We show that the impact of subjective bias on fracture data collection can be extreme, however, using a number of recommendations provided in this study it should be possible to limit the effect for users of this data.

**Data availability**

The workshop documents and data collected as part of this study is available in the supplementary information.

**Supplement link (Will be included by copernicus)**

**Team List**

**Author contributions**

Initial discussions and planning of the paper was undertaken by all authors, with BA and JRR designing the workshops. The paper was prepared by BJA and JRR, with contributions from all authors.

**Competing interests**

The authors declare that they have no conflict of interest.

**Disclaimer**

**Special issue Statement (Will be included by Copernicus)**

**Acknowledgements**

This work was funded through Billy J. Andrews' PhD studentship, supported by the Environmental and Physical Sciences research council (EPSRC, award number EP/L016680/1), the ENOS project: H2020-EU.3.3.2.3. - Develop competitive and environmentally safe technologies for CO2 capture, transport, storage and re-use. Record Number: 664337, UKCCSRC, funded by the EPSRC (EP/K000446/1, EP/P026214/1) as part of the RCUK Energy Program and the Strathclyde Global Engagement Fund. The authors also wish to thank all participants who attended the workshops and took part in the fieldwork as part of this study.

## References

Agosta, F., Alessandroni, M., Antonellini, M., Tondi, E. and Giorgioni, M.: From fractures to flow: A field-based quantitative analysis of an outcropping carbonate reservoir, Tectonophysics, 490(3–4), 197–213, doi:10.1016/j.tecto.2010.05.005, 2010.

Becker, A. and Gross, M. R.: Mechanism for joint saturation in mechanically layered rocks: an example from southern Israel, Tectonophysics, 257(2–4), 223–237, doi:10.1016/0040-1951(95)00142-5, 1996.

Belayneh, M. W., Matthai, S. K., Blunt, M. J. and Rogers, S. F.: Comparison of deterministic with stochastic fracture models in water-flooding numerical simulations, Am. Assoc. Pet. Geol. Bull., 93(11), 1633–1648, doi:10.1306/07220909031, 2009.

Bigi, S., Battaglia, M., Alemanni, A., Lombardi, S., Campana, A., Borisova, E. and Loizzo, M.: CO2flow through a fractured rock volume: Insights from field data, 3D fractures representation and fluid flow modeling, Int. J. Greenh. Gas Control, 18, 183–199, doi:10.1016/j.ijggc.2013.07.011, 2013.

Bigi, S., Marchese, M., Meda, M., Nardon, S. and Franceschi, M.: Discrete fracture network of the Latemar carbonate platform, Ital. J. Geosci., 134(3), 474–494, doi:10.3301/IJG.2014.34, 2015.

Bond, C. E., Johnson, G. and Ellis, J. F.: Structural model creation: the impact of data type and creative space on geological reasoning and interpretation, Geol. Soc. London, Spec. Publ., 421(1), 83–97, doi:10.1144/SP421.4, 2015.

Chadwick, B. A., Holliday, D. W., Holloway, S., Hulbert, A. G. and Lawrence, D. J. D.: The structure and evolution of the Northumberland-Solway Basin and adjacent areas., London:HMSO, 1995.

Chesnaux, R., Allen, D. M. and Jenni, S.: Regional fracture network permeability using outcrop scale measurements, Eng. Geol., 108(3–4), 259–271, doi:10.1016/j.enggeo.2009.06.024, 2009.

Comerford, A., Fraser-Harris, A., Johnson, G. and McDermott, C. I.: Controls on geothermal heat recovery from a hot sedimentary aquifer in Guardbridge, Scotland: Field measurements, modelling and long term sustainability, Geothermics, 76(July), 125–140, doi:10.1016/j.geothermics.2018.07.004, 2018.

Dershowitz, W. S. and Einstein, H. .: Characterizing Rock Joint Geometry with Joint System Models, Rock Mech. Rock Eng., 21(1), 21–51, 1988.

Van Dijk, J. P., Bello, M., Toscano, C., Bersani, A. and Nardon, S.: Tectonic model and three-dimensional fracture network analysis of Monte Alpi (southern Apennines)., 2000.

Guerriero, V., Vitale, S., Ciarcia, S. and Mazzoli, S.: Improved statistical multi-scale analysis of fractured reservoir analogues, Tectonophysics, 504(1–4), 14–24, doi:10.1016/j.tecto.2011.01.003, 2011.

Harthong, B., Scholtès, L. and Donzé, F.: Strength characterization of rock masses, using a coupled DEM-DFN model, Geophys. J. Int., 191(2), 467–480, doi:10.1111/j.1365-246X.2012.05642.x, 2012.

Healy, D., Rizzo, R. E., Cornwell, D. G., Farrell, N. J. C., Watkins, H., Timms, N. E., Gomez-Rivas, E. and Smith, M.: FracPaQ: A MATLAB™ toolbox for the quantification of fracture patterns, J. Struct. Geol., 95, 1–16, doi:10.1016/j.jsg.2016.12.003, 2017.

Jing, L. and Hudson, J. A.: Numerical methods in rock mechanics, Int. J. Rock Mech. Min. Sci., 39(4), 409–427, doi:10.1016/S1365-1609(02)00065-5, 2002.

Johnson, G. A. L.: Subsidence and sedimentation in the Northumberland Trough, Proc. Yorksh. Geol. Soc., 45(71–83), 1984.

Jolly, R. and Cosgrove, J.: Geological evidence of patterns of fluid flow through fracture networks: examination using random realizations and connectivity analysis, Geol. Soc. London, Spec. Publ., 209(1), 177–186, 2003.

Koike, K., Nagano, S. and Kawaba, K.: Construction and analysis of interpreted fracture planes through combination of satellite-image derived lineaments and digital elevation model data, Comput. Geosci., 24(6), 573–583, doi:10.1016/S0098-3004(98)00021-1, 1998.

Laubach, S. E., Lamarche, J., Gauthier, B. D. M. and Dunne, W. M.: Editorial: Spatial arrangement of faults and opening-mode fractures, J. Struct. Geol., 108, 1, doi:10.1016/j.jsg.2018.01.008, 2018.

Lei, Q., Latham, J. P. and Tsang, C. F.: The use of discrete fracture networks for modelling coupled geomechanical and hydrological behaviour of fractured rocks, Comput. Geotech., 85, 151–176, doi:10.1016/j.compgeo.2016.12.024, 2017.

Lyman, G. J.: Stereological and Other Methods Applied to Rock Joint Size Estimation - Does Crofton's Theorem Apply?, Math. Geol., 35(1), 9–23, doi:10.1023/A:1022326927389, 2003.

Manzocchi, T.: The connectivity of two-dimensional networks of spatially correlated fractures, Water Resour. Res., 38(9), 1-1-1–20, doi:10.1029/2000WR000180, 2002.

Mas Ivars, D., Pierce, M. E., Darcel, C., Reyes-Montes, J., Potyondy, D. O., Paul Young, R. and Cundall, P. A.: The synthetic rock mass approach for jointed rock mass modelling, Int. J. Rock Mech. Min. Sci., 48(2), 219–244, doi:10.1016/j.ijrmms.2010.11.014, 2011.

Mauldon, M. and Mauldon, J. G.: Fracture sampling on a cylinder: From scanlines to boreholes and tunnels, Rock Mech. Rock Eng., 30(3), 129–144, doi:10.1007/BF01047389, 1997.

Mauldon, M., Dunne, W. M. and Rohrbaugh, M. B.: Circular scanlines and circular windows: New tools for characterizing the geometry of fracture traces, J. Struct. Geol., 23(2–3), 247–258, doi:10.1016/S0191-8141(00)00094-8, 2001.

Min, K. B., Jing, L. and Stephansson, O.: Determining the equivalent permeability tensor for fractured rock masses using a stochastic REV approach: Method and application to the field data from Sellafield, UK, Hydrogeol. J., 12(5), 497–510, doi:10.1007/s10040-004-0331-7, 2004.

Neuman, S. P.: Generalized scaling of permeabilities: Validation and effect of support scale, Geophys. Res. Lett., 21(5), 349–352, 1993.

Newman, M. E. J.: Power laws, Pareto distributions and Zipf's law, Contemp. Phys., 46(5), 323–351, doi:10.1080/00107510500052444, 2005.

Nixon, C. W., Sanderson, D. J. and Bull, J. M.: Analysis of a strike-slip fault network using high resolution multibeam bathymetry, offshore NW Devon U.K., Tectonophysics, 541–543, 69–80, doi:10.1016/j.tecto.2012.03.021, 2012.

Olson, J. E., Laubach, S. E. and Lander, R. H.: Natural fracture characterization in tight gas sandstones: Integrating mechanics and diagenesis, Am. Assoc. Pet. Geol. Bull., 93(11), 1535–1549, doi:10.1306/08110909100, 2009.

Ortega, O. and Marrett, R.: Prediction of macrofracture properties using microfracture information, Mesaverde Group sandstones, San Juan basin, New Mexico, J. Struct. Geol., 22(5), 571–588, doi:10.1016/S0191-8141(99)00186-8, 2000.

Ortega, O. J., Marrett, R. A. and Laubach, S. E.: A scale-independent approach to fracture intensity and average spacing measurement, Am. Assoc. Pet. Geol. Bull., 90(2), 193–208, doi:10.1306/08250505059, 2006.

Pahl, P. J.: Estimating the mean length of discontinuity traces, Int. J. Rock Mech. Min. Sci., 18(3), 221–228, doi:10.1016/0148-9062(81)90976-1, 1981.

Peacock, D. C. P. and Sanderson, D. J.: Structural analyses and fracture network characterisation: Seven pillars of wisdom, Earth-Science Rev., 184(June), 13–28, doi:10.1016/j.earscirev.2018.06.006, 2018.

Priest, S. D.: Discontinuity analysis for rock engineering, Springer Science & Business Media., 1993.

Priest, S. D. and Hudson, J. A.: Estimation of discontinuity spacing and trace length using scanline surveys, Int. J. Rock Mech. Min. Sci., 18(3), 183–197, doi:10.1016/0148-9062(81)90973-6, 1981.

Procter, A. and Sanderson, D. J.: Spatial and layer-controlled variability in fracture networks, J. Struct. Geol., 1–14, doi:10.1016/j.jsg.2017.07.008, 2017.

Rohrbaugh, J. B., Dunne, W. M. and Mauldon, M.: Estimating fracture trace intensity, density, and mean length using circular scan lines and windows, Am. Assoc. Pet. Geol. Bull., 86(12), 2089–2104, doi:10.1306/61EEDE0E-173E-11D7-8645000102C1865D, 2002.

Salvini, R., Mastrorocco, G., Seddaiu, M., Rossi, D. and Vanneschi, C.: The use of an unmanned aerial vehicle for fracture mapping within a marble quarry (Carrara, Italy): photogrammetry and discrete fracture network modelling, Geomatics, Nat. Hazards Risk, 8(1), 34–52, doi:10.1080/19475705.2016.1199053, 2017.

Sanderson, D. J. and Nixon, C. W.: The use of topology in fracture network characterization, J. Struct. Geol., 72, 55–66, doi:10.1016/j.jsg.2015.01.005, 2015.

Shipton, Z. K., Roberts, J. J., Comrie, E. L., Kremer, Y., Lunn, R. J. and Caine, J. S.: Fault Fictions: Cognitive biases in the conceptualization of fault zones, in Geological Society London Special Publication., edited by S. Ogilvie, J. L. Urai, S. Dee, R. W. Wilson, and W. Bailey., n.d.

Tannert, C., Elvers, H.-D. and Jandrig, B.: The Ethics of Uncertainty, EMBO Rep., 8(10), 892–896, doi:10.1038/sj.embor.7401072, 2007.

Terzaghi, R. D.: Sources of Error in Joint Surveys, Géotechnique, 15(3), 287–304, doi:10.1680/geot.1965.15.3.287, 1965.

Tóth, T. M.: Determination of geometric parameters of fracture networks using 1D data, J. Struct. Geol., 32(7), 878–885, doi:10.1016/j.jsg.2009.04.006, 2010.

Watkins, H., Bond, C. E., Healy, D. and Butler, R. W. H.: Appraisal of fracture sampling methods and a new workflow tocharacterise heterogeneous fracture networks at outcrop, J. Struct. Geol., 72(February), 67–82, doi:10.1016/j.jsg.2015.02.001, 2015.

Wood, R. and Curtis, A.: Geological prior information and its applications to geoscientific problems, Geol. Soc. London, Spec. Publ., 239(1), 1.2-14, doi:10.1144/GSL.SP.2004.239.01.01, 2004.

Wu, H. and D. Pollard, D.: An experimental study of the relationship between joint spacing and layer thickness, J. Struct.

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


Geol., 17(6), 887–905, doi:10.1016/0191-8141(94)00099-L, 1995.

Zeeb, C., Gomez-Rivas, E., Bons, P. D. and Blum, P.: Evaluation of Sampling methods for fracture network characterization using Outcrops, Am. Assoc. Pet. Geol. Bull., 97(9), 1545–1566, doi:10.1306/02131312042, 2013.

## Tables

| Fracture statistic | Notation | Definition (unit) | Input parameters and calculation | | |
|---|---|---|---|---|---|
| | | | Linear | Circular scanline | Window sampling |
| Density (D) | Areal (P20) | Number of fractures per unit area ($m^{-2}$) | - | $D = \frac{(Ni+Ny)}{2\pi r^2}$ | $D = \frac{N}{A}$ |
| Intensity (I) | Linear (P10) | Number of fractures per unit length ($m^{-1}$) | $I = \frac{Nc}{L} = \frac{1}{S}$ | - | - |
| | Areal (P21) | Fracture length per unit area (m x $m^{-2}$) | - | $I = \frac{Nc}{4r}$ | $I = \frac{\sum tl}{A}$ |
| Spacing (S) | Linear | Spacing between fractures (m) | $S = \frac{\sum s}{(N-1)} = \frac{1}{I}$ | - | - |
| Mean trace length (Tl) | Tl | Mean fracture length (m) | $Tl = \frac{\sum l}{N}$ | $Tl = \frac{Nc}{(Ni+Ny)} \times \frac{\pi r}{2}$ | $Tl = \frac{\sum l}{N}$ |
| Network topology | Topological sampling | Defining fracture nodes as I, y and x. | - | Yes | Yes |
| Connectivity | Using node topology (Pc) | Percentage of connected branches | - | $Pc = \frac{3Ny + 4Nx}{Ni + Ny + Nx}$ | $Pc = \frac{3Ny + 4Nx}{Ni + Ny + Nx}$ |
| | Using trace end classification (Pf) | Percentage of connected fractures | $Pf = \frac{F}{R + F} \times 100$ | - | - |
| Trace length distribution | Tl distribution (tl) | Distribution of individual fracture trace lengths | Yes | - | Yes |

**Table 1: Summary and definition of fracture statistics that can be derived from methods used in this work. Table adapted from Zeeb et al. (2013). Ni = number of i-nodes, Ny = number of y-nodes, Nx = number of x-nodes, r = radius of circular scanline, N = number of fractures, A = Area, Nc = number of fracture intersections with the scanline (either linear or circular), L = length of scanline, s = spacing between adjacent fracture traces on the scanline, tl = individual fracture trace length, F = fracture abuts against another fracture, R = fracture terminates into rock (n.b. some authors also distinguish stratabound fracture terminations), 'Yes' for trace length distribution & network topology indicates you can use that method to carry out the technique.**

| *Method* | | Field | | | Workshop | | | | Length or radius/m |
|---|---|---|---|---|---|---|---|---|---|
| | | Completed? | i | g | Completed? | i | g | Order | |
| *Circular* | C1 | ✓ | ✓ | | ✓ (WS1&2) | ✓ | | 3 | 1.0 |
| | C2 | ✓ | ✓ | | ✗ | | | | 1.0 |
| | C3 | ✓ | ✓ | | ✓(WS1&2) | | ✓ | 5 | 1.0 |
| | C4 | ✓ | ✓ | | ✓(WS1&2) | | ✓ | 4 | 1.0 |
| | C5 | ✓ | ✓ | | ✓(WS1&2) | ✓ | | 2 | 1.0 |
| | C6 | ✓ | ✓ | | ✗ | | | | 0.73 |
| | C7 | ✓ | ✓ | | ✗ | | | | 1.21 |
| | C8 | ✗ | | | ✓(WS1&2) | ✓ | | 1 | 0.5 |
| *Linear* | L1 | ✓ | ✓ | | ✗ | | | | 1.0 |
| | L2 | ✓ | | ✓ | ✗ | | | | 1.0 |
| | L3 | ✓ | | ✓ | ✗ | | | | 15.0 |
| | L4 | ✓ | | ✓ | ✗ | | | | 7.5 |
| | L5 | ✗ | | | ✓(WS1&2) | | ✓ | | 6.55 |
| | L6 | ✗ | | | ✓(WS1&2) | ✓ | | | 1.45 |
| *Window sampling* | C1 | | | | ✓P1,3,11 & WS2 | ✓ | | 3 | 0.5 |
| | C3 | | | | WS2 | | ✓ | 5 | 1.0 |
| | C4 | | | | WS2 | | ✓ | 4 | 1.0 |
| | C5 | | | | ✓P1,3,11 & WS2 | ✓ | | 2 | 0.5 |
| | C8 | | | | ✓P1,3,11 & WS2 | ✓ | | 1 | 0.5 |

**Table 2: Summary of circular (C) and linear (L) scanlines completed in the field and workshops (WS1 & WS2). Whether these were completed individually (i) or in groups (g) is noted. 'Order' refers to the order the scanlines were completed in the workshops. Four of the circular scanlines (C2,3,4,5) were completed both in the field and in the workshop, but none of the linear scanlines were completed in both, due to workshop time constraints. Window sampling, whereby participants drew out the interpreted fractures as well as completing topological sampling, was only completed by Participants 1, 3, 11 and all of Workshop 2 (WS2). The workbooks used in this study are supplied in the supplementary information (S3 & S4).**

| Group | N participants | Geological training | | | | | Familiarity with geological fieldwork | | | | | Familiarity with collecting fracture data | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | None | Low | Medium | High | (Other) | None | Low | Medium | High | (Other) | None | Low | Medium | High | (Other) |
| Field | 7 | 1 | 0 | 3 | 3 | 0 | 1 | 0 | 3 | 3 | 0 | 1 | 0 | 3 | 3 | 0 |
| WS1 | 11 | 2 | 2 | 3 | 2 | 2 | 2 | 1 | 5 | 1 | 2 | 3 | 2 | 5 | 1 | 0 |
| WS2 | 18 | 3 | 0 | 6 | 9 | 0 | 3 | 6 | 3 | 6 | 0 | 6 | 5 | 5 | 2 | 0 |

**Table 3. Summary of the level of geological training, and experience in geological fieldwork and fracture data collection, reported by field and workshop (WS) participants. Individual participant responses are provided in the Supplementary Information (S2).**

| Scanline | | Individual/ Group | n | Fracture count | | | Trace length (m) | | | Time (minutes) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Min | Max | Mean | Min | Max | Mean | Min | Max | Mean |
| L1 | Field | i | 6 | 3 | 10 | 6.8 | 0.03 | 2.22 | 0.55 | 5:32* | 9:00* | 7:16* |
| L2 | Field | G | 3 | 7 | 14 | 11 | 0.01 | 1.78 | 0.41 | - | - | - |
| L3 | Field | G | 3 | 21 | 38 | 28.3 | 0.04 | 23.08 | 1.2 | 10:00 | 13:00 | 11:00 |
| L4 | Field | G | 2 | 18 | 19 | 18.5 | 0.05 | 14.4 | 2.3 | - | - | - |
| L6 | WS1 | i | 11 | 10 | 23 | 15.6 | 0.02 | 0.61 | 0.21 | 2:17 | 8:40 | 4:56** |
| | WS2 | i | 18 | 9 | 25 | 18.8 | 0.03 | 0.72 | 0.24 | 1:51 | 24:00 | 7:42*** |
| L5 | WS1 | G | 5 | 22 | 31 | 24.8 | 0.12 | 2.72 | 0.86 | 5:57 | 9:35 | 7:34 |
| | WS2 | G | 7 | 15 | 28 | 21 | 0.14 | 2.43 | 0.96 | 5:00 | 13:00 | 8:23 |

**Table 4: Summary table of raw linear scanline results where i = individual, G = groups, n = number of participants/groups. *only two participants recorded time for this scanline **P10 did not record time taken to count nodes ***P23 did not trace fractures so only have spacing and time information.**

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


Solid Earth Discussions
Open Access
EGU

| | | | Nc | | | | Node Count | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Nc | | t (sec) | | i-node | | y-node | | x-node | | t (sec) | |
| | | i/g | Range | mean | Range | Mean | Range | mean | Range | mean | Range | mean | Range | Mean |
| C1 | Field | i | 15-21 | 17.5 | 19-42 | 29 | 0-3 | 1.0 | 12-21 | 17.7 | 6-14 | 8.5 | 137-230 | 178 |
| | WS1 | i | 14-23 | 18.1 | 36-99 | 59 | 0-12 | 2.1 | 1-38 | 19.7 | 4-11 | 7 | 119-447 | 247 |
| | WS2 | i | 11-25 | 18.1 | 15-295 | 111 | 0-6 | 1.7 | 4-34 | 18.2 | 4-14 | 7.8 | 82-1140 | 382 |
| C5 | Field | i | 14-19 | 15.7 | 14-43 | 26 | 4-8 | 5.8 | 28-47 | 35.7 | 2-8 | 4.2 | 127-245 | 171 |
| | WS1 | i | 7-18 | 12.4 | 20-120 | 52 | 3-14 | 6.4 | 4-34 | 19.8 | 1-6 | 2.3 | 150-1177 | 357 |
| | WS2 | i | 9-18 | 12.6 | 20-298 | 89 | 0-32 | 6.8 | 7-41 | 16.7 | 0-11 | 2.6 | 60-1050 | 377 |
| C8 | WS1 | i | 10-25 | 20.9 | 29-180 | 80 | 2-11 | 4.7 | 1-60 | 25.5 | 2-22 | 10.5 | 150-780 | 390 |
| | WS2 | i | 16-32 | 23.2 | 45-240 | 113 | 1-16 | 5.2 | 5-45 | 19.4 | 5-18 | 10.9 | 30-1440 | 627 |
| C4 | Field (i) | i | 12-20 | 14.8 | 24-50 | 39 | 5-19 | 13 | 20-34 | 28.8 | 0-4 | 0.8 | 147-215 | 178 |
| | WS1 | g | 11-18 | 14.8 | 60-330 | 132 | 7-19 | 11 | 6-27 | 14 | 1-4 | 2.8 | 324-521 | 405 |
| | WS2 | g | 10-18 | 14.4 | 64-323 | 177 | 5-23 | 8.9 | 5-27 | 12.3 | 0-3 | 1.6 | 115-720 | 347 |
| C3 | Field | i | 19-30 | 23 | 24-58 | 43 | 3-15 | 9.2 | 21-33 | 28.4 | 6-16 | 9.6 | 162-282 | 244 |
| | WS1 | g | 18-22 | 19.8 | 55-90 | 75 | 4-20 | 8.8 | 19-24 | 22.2 | 5-11 | 6.8 | 208-521 | 343 |
| | WS2 | g | 14-23 | 17.2 | 52-713 | 242 | 2- 54 | 13.9 | 11-22 | 17.9 | 3-10 | 4.9 | 143-600 | 329 |

**Table 5: Summary of fracture data and time taken for circular scanlines 1, 5 and 8, in the field and workshop, either working individually (i) or in groups (g). The data are presented in the order scanlines were completed in the workshops.**

| Circle | Number of participants | Number of fractures | | Trace length (m) | | |
|---|---|---|---|---|---|---|
| | | Range | Mean | Min | Max | Mean |
| 8 (i) | 20 | 18-54 | 34.4 | 0.01-0.10 | 0.70-0.98 | 0.27 |
| 5 (i) | 20 | 13-56 | 24.9 | 0.02-0.12 | 0.68-1.05 | 0.33 |
| 1 (i) | 20 | 9-40 | 23.6 | 0.01-0.40 | 0.67-1.03 | 0.37 |
| 4 (g) | 7 | 11-29 | 19.9 | 0.02-0.11 | 1.89-1.95 | 0.69 |
| 3 (g) | 7 | 18-50 | 27.4 | 0.04-0.22 | 1.82-2.01 | 0.61 |

**Table 6: Summary of fracture parameters reported for window sampling. Data is presented in the order the scanlines were undertaken within the workshops. (i) and (g) denote whether the scanline was undertaken individually or as a group.**

| Statistic | Circular Scanline – topology | Circular Scanline - Window | Linear Scanline |
|---|---|---|---|
| Intensity | Moderate spread, varies between circles. For Circle 1, 4 and 5 the calculated intensity from workshop and field data were very similar, however, the calculated intensity for Circle 3 was much lower in the workshop. In all cases ranges are greater when workshop data is used, particularly for Circles 1 and 5. | Limited spread between participants within circles. In all cases, apart from Circle 4, intensity calculated using window sampling is lower than that derived for node counting for a given circle. | Moderate spread, more for some Lines (Line 1, Line 6) than others (Lines 3 - 5, all low intensity, have small range). |
| Density / Spacing | Moderate to high spread. Generally higher spread observed in workshop data, particularly in Circle 1. Density calculated from workshop in all cases apart from Circle 1 is lower than when calculated from field data. | Moderate to high spread. Values consistently higher in workshop data when window sampling data is used compared to node counting, particularly Circle 8. Can be both comparable to field density (Circle 4) or considerably higher (Circle 1) | Moderate to large spread. Equally large range in workshops and field. Minimum fracture spacing was consistently small. |
| Mean trace length | Greatest spread. Particularly large spread in values derived from workshop data. How similar the range is between workshop and field data varies for different circles. | Moderate spread. The extremes in the ranges observed in mean trace length estimates are considerable lower than for node counting. Of all methods window sampling provides the smallest estimate for mean trace length. | Highly variable for most scanlines. Equally large range in workshops and field. Maximum reported trace lengths generally much larger than for other methods, due to the different scale of observation. |
| Connectivity | Least spread, both between circles, between methods, and settings (field vs workshop). | Not assessed separately from node classifications. | Highly variable for most scanlines. Equally large range in workshops and field. |

**Table 7: Summary of the broad trends in fracture statistics derived from the three methods we explored, presented in Fig. 9.**

**Figures**

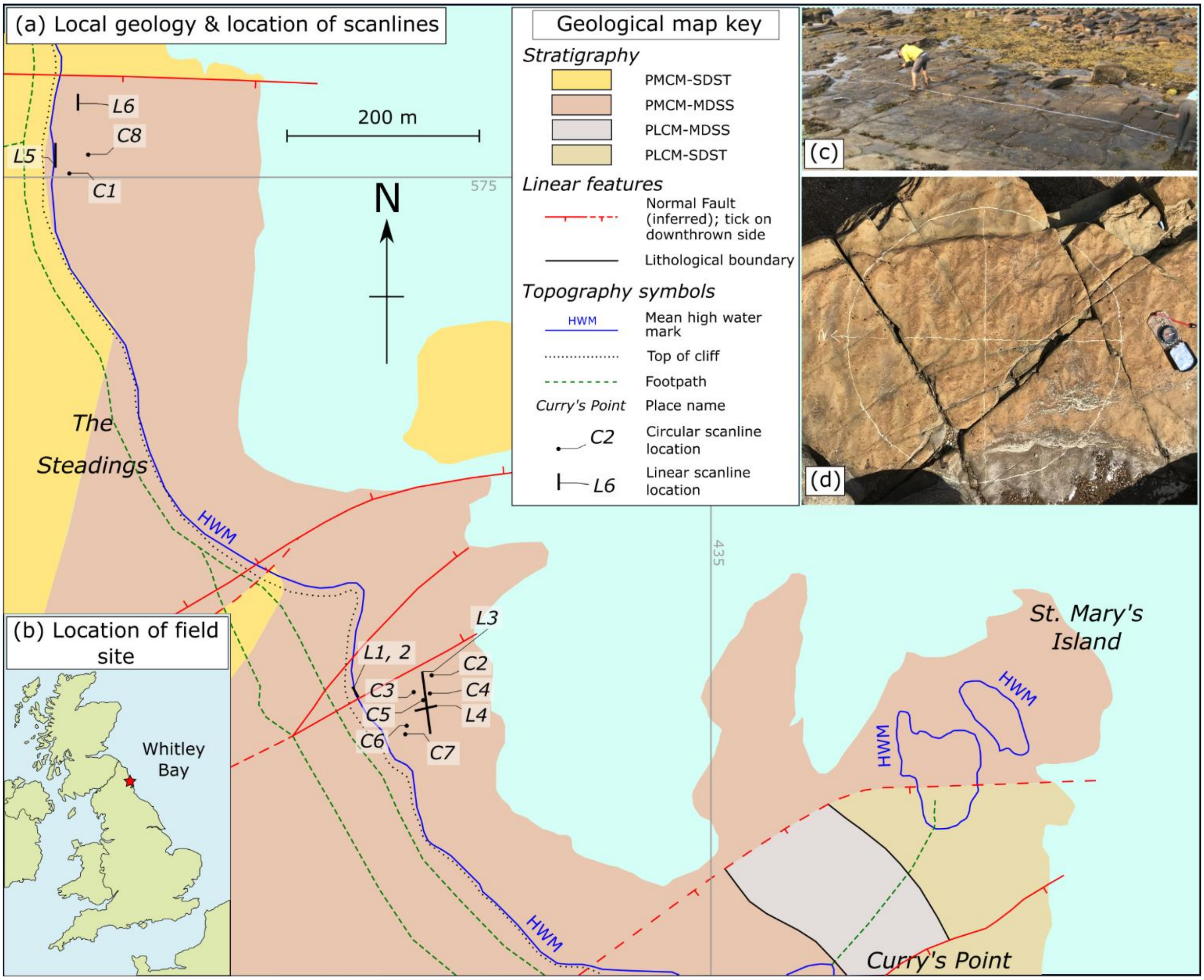
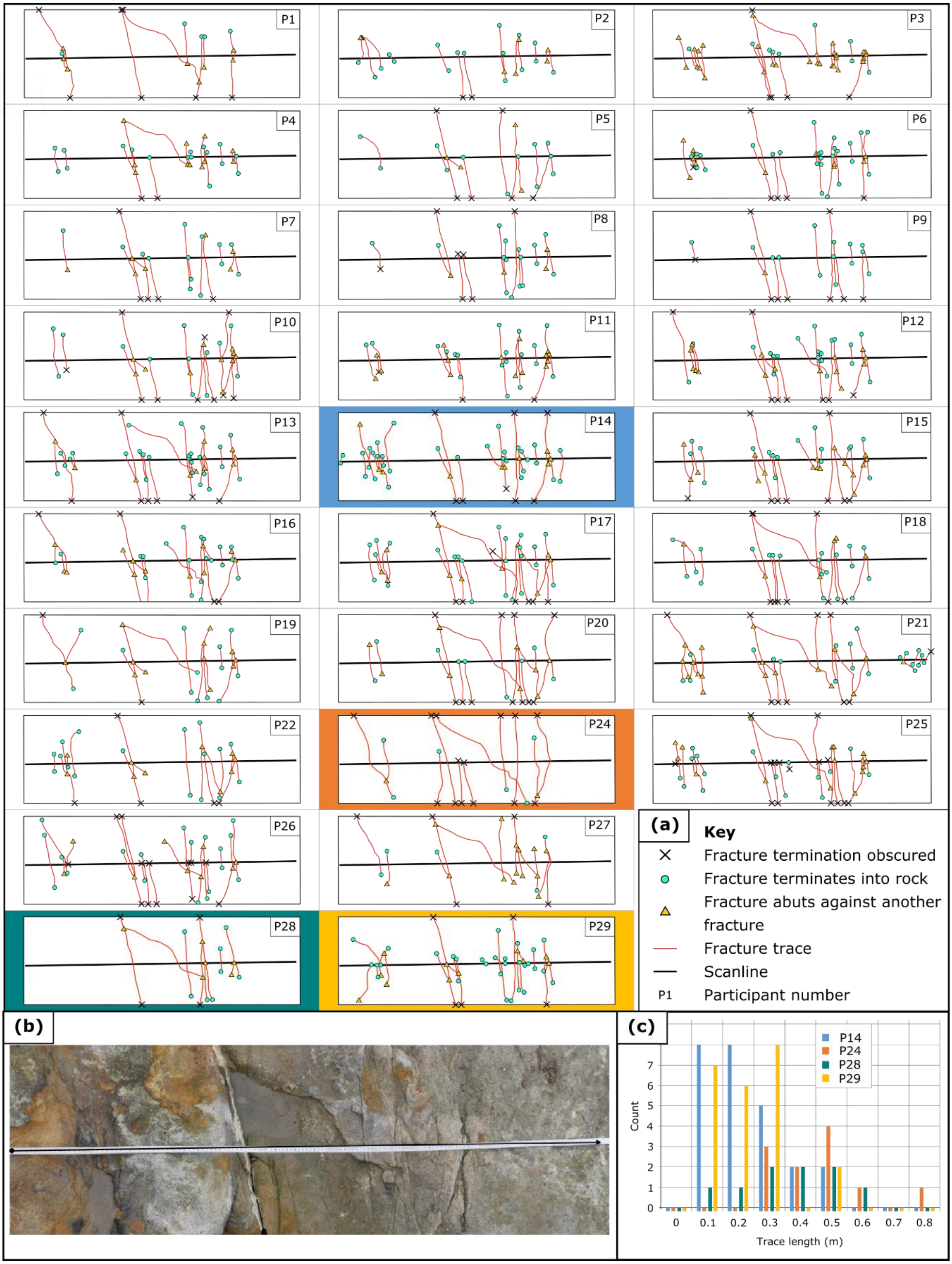



**Figure 1: Location map highlighting (a) the local geology and (b) the location of the study area, located near Whitley Bay, Northumberland (UK). Grid lines are annotated with UK national grid numbers. Field photographs of both linear (c) and circular (d) scanline methods are also shown (L3 [NZ34717545] and C8 [NZ34377609] respectively). The geological map is modified from Geological Map Data BGS © UKRI (2018), where stratigraphy is as follows: PLCM-SDST = Pennine Lower Coal Measures – Sandstone; PLCM-MDSS = Pennine Lower Coal Measures – Mudstone, siltstone and Sandstone; Pennine Middle Coal Measures – Sandstone; PLCM-MDSS = Pennine Middle Coal Measures – Mudstone, siltstone and Sandstone.**

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019



Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


**Figure 2: The interpreted fracture traces for Line 6 (length 1.45 m). (a) The digitised fracture networks for all workshop participants. (b) Field photograph of Line 6. (c) Fracture trace length histograms (bin = 0.1 m) for participants who recorded a low to high number of fractures. The corresponding digitised fracture trace is also highlighted in the appropriate colour. Key differences in the interpreted fracture networks are highlighted using participants who selected a low (Participant 28, 9 fractures), medium (Participant 10, 17 fractures) and high (Participant 14, 25 fractures) number of fractures.**

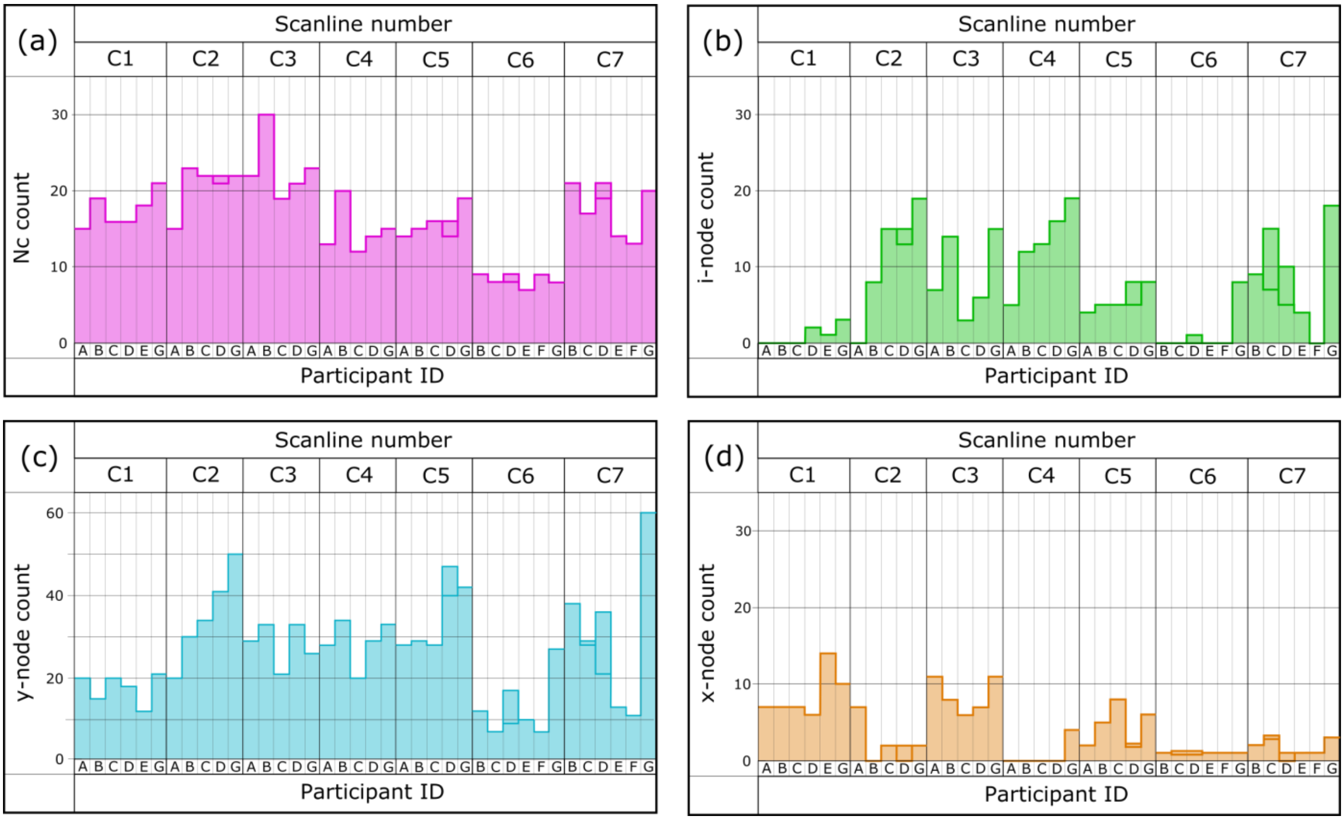



**Figure 3: Results of the fracture data from circular scanlines (C1-7) collected in the field by 7 participants (labelled A-G, though A, E and F did not complete all of the scanlines). (a) the number of fractures that intersected the circular scanlines (Nc). (b) fractures that terminated in rock (i-nodes). (c) fractures that terminated against another fracture (y-nodes). (d) fractures that intersect another fracture (x-nodes). Participants C and D repeated some of their measurements for selected circles and this is indicated by two bars in their column for that circle.**

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


Solid Earth Discussions
Open Access
EGU

(a)

| | | Count | | | | | | | | | | | | Time (s) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *Nc* | | | *i-node* | | | *y-node* | | | *x-node* | | | *Nc* | | | *Node* | | |
| | *P* | C8 | C5 | C1 | C8 | C5 | C1 | C8 | C5 | C1 | C8 | C5 | C1 | C8 | C5 | C1 | C8 | C5 | C1 |
| Workshop 1 | *1* | 23 | 11 | 21 | 4 | 9 | 2 | 22 | 24 | 32 | 12 | 2 | 7 | 78 | 70 | 68 | 540 | 324 | 337 |
| | *2* | 11 | 8 | 16 | 2 | 4 | 2 | 1 | 4 | 1 | 3 | 3 | 8 | 107 | 59 | 99 | 378 | 317 | 259 |
| | *3* | 24 | 14 | 20 | 5 | 12 | 1 | 60 | 34 | 38 | 10 | 2 | 7 | 46 | - | 72 | 460 | 1177 | 447 |
| | *4* | 22 | 12 | 16 | 6 | 3 | 0 | 28 | 17 | 18 | 9 | 1 | 7 | 106 | 53 | 50 | 602 | 333 | 119 |
| | *5* | 10 | 7 | 14 | 5 | 3 | 1 | 1 | 7 | 5 | 2 | 1 | 5 | 83 | 70 | 70 | 172 | 150 | 120 |
| | *6* | 25 | 14 | 23 | 4 | 5 | 12 | 27 | 26 | 29 | 12 | 1 | 11 | 52 | 32 | 51 | 312 | 330 | 416 |
| | *7* | 20 | 11 | 17 | 2 | 3 | 0 | 13 | 20 | 16 | 9 | 2 | 10 | 120 | 30 | 60 | 480 | 300 | 300 |
| | *8* | 25 | 16 | 19 | 6 | 7 | 3 | 26 | 12 | 20 | 15 | 6 | 8 | 36 | 28 | 38 | 150 | 150 | 211 |
| | *9* | 25 | 14 | 16 | 5 | 5 | 0 | 33 | 24 | 12 | 12 | 4 | 5 | 180 | 120 | 60 | 780 | 480 | 240 |
| | *10* | 21 | 12 | 18 | 2 | 5 | 0 | 23 | 19 | 19 | 22 | 1 | 4 | 29 | 20 | 49 | 171 | 186 | 141 |
| | *11* | 24 | 18 | 19 | 11 | 14 | 2 | 47 | 31 | 27 | 10 | 2 | 5 | 47 | 41 | 36 | 242 | 184 | 125 |
| Workshop 2 | *12* | 24 | 13 | 18 | 4 | 4 | 2 | 42 | 26 | 25 | 8 | 1 | 6 | 102 | 298 | 295 | 1200 | 235 | 290 |
| | *13* | 26 | 18 | 25 | 15 | 32 | 6 | 45 | 41 | 34 | 18 | 6 | 10 | 180 | 60 | 60 | 1380 | 900 | 540 |
| | *14* | 28 | 12 | 21 | 4 | 7 | 1 | 23 | 16 | 18 | 9 | 1 | 9 | 109 | 80 | 107 | 705 | 451 | 538 |
| | *15* | 25 | 16 | 22 | 2 | 5 | 1 | 31 | 32 | 34 | 16 | 5 | 8 | 129 | 64 | 80 | 864 | 737 | 528 |
| | *16* | 24 | 13 | 19 | 5 | 13 | 1 | 14 | 14 | 20 | 12 | 0 | 4 | 105 | 89 | 230 | 660 | 600 | 259 |
| | *17* | 19 | 11 | 20 | 3 | 6 | 2 | 20 | 15 | 13 | 13 | 7 | 14 | 94 | 58 | 48 | 622 | 310 | 509 |
| | *18* | 26 | 12 | 19 | 3 | 2 | 0 | 19 | 13 | 10 | 15 | 2 | 7 | 134 | 71 | 84 | 504 | 235 | 186 |
| | *19* | 22 | 9 | 20 | 4 | 4 | 3 | 26 | 14 | 32 | 8 | 1 | 4 | 210 | 112 | 176 | 598 | 350 | 430 |
| | *20* | 16 | 12 | 18 | 1 | 2 | 1 | 5 | 23 | 18 | 5 | 2 | 6 | 45 | 240 | 254 | 125 | 325 | 217 |
| | *21* | 25 | 14 | 22 | 4 | 7 | 4 | 6 | 10 | 12 | 18 | 11 | 8 | 55 | 33 | 45 | 295 | 237 | 289 |
| | *22* | 18 | 11 | 13 | 5 | 3 | 0 | 7 | 11 | 4 | 7 | 1 | 10 | 98 | 131 | 74 | 730 | 517 | 550 |
| | *23* | 25 | 11 | 15 | 16 | 11 | 6 | 8 | 9 | 10 | 6 | 2 | 7 | 120 | 60 | 120 | 300 | 120 | 540 |
| | *24* | 22 | 12 | 11 | 4 | 3 | 0 | 7 | 7 | 9 | 12 | 2 | 6 | 120 | 120 | 120 | 600 | 180 | 240 |
| | *25* | 23 | 12 | 16 | 2 | 2 | 0 | 18 | 11 | 21 | 6 | 0 | 6 | 70 | 20 | 40 | 240 | 60 | 180 |
| | *26* | 32 | 12 | 17 | 8 | 11 | 2 | 29 | 14 | 25 | 12 | 3 | 8 | 121 | 34 | 46 | 458 | 165 | 138 |
| | *27* | 20 | 12 | 15 | 4 | 2 | 0 | 21 | 18 | 12 | 9 | 0 | 7 | 52 | 25 | 32 | 527 | 252 | 213 |
| | *28* | 16 | 12 | 13 | 1 | 0 | 0 | 7 | 9 | 5 | 9 | 1 | 10 | 46 | 21 | 15 | 30 | 60 | 82 |
| | *29* | 27 | 14 | 21 | 8 | 9 | 1 | 22 | 18 | 25 | 13 | 1 | 10 | 240 | 90 | 180 | 1440 | 1050 | 1140 |

(b)

| | | Count | | | | | | | | Time (s) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *Nc* | | *i-node* | | *y-node* | | *x-node* | | *Nc* | | *Node* | |
| | *G* | C4 | C3 | C4 | C3 | C4 | C3 | C4 | C3 | C4 | C3 | C4 | C3 |
| Workshop 1 | *1* | 14 | 22 | 7 | 20 | 17 | 24 | 3 | 11 | 330 | 90 | 521 | 521 |
| | *2* | 13 | 18 | 11 | 4 | 11 | 19 | 1 | 6 | 62 | 82 | 324 | 208 |
| | *3* | 18 | - | 19 | - | 27 | - | 4 | - | 97 | - | 405 | - |
| | *4* | 11 | 18 | 9 | 5 | 9 | 23 | 2 | 5 | 60 | 73 | 357 | 332 |
| | *5* | 18 | 21 | 9 | 6 | 6 | 23 | 4 | 5 | 110 | 55 | 420 | 312 |
| Workshop 2 | *6* | 18 | 23 | 14 | 11 | 13 | 17 | 1 | 6 | 120 | 60 | 600 | 360 |
| | *7* | 18 | 18 | 5 | 3 | 27 | 22 | 3 | 3 | 129 | 129 | 720 | 600 |
| | *8* | 14 | 14 | 5 | 7 | 14 | 18 | 1 | 4 | 323 | 713 | 115 | 143 |
| | *9* | 12 | 16 | 5 | 16 | 11 | 22 | 1 | 3 | 184 | 389 | 445 | 168 |
| | *10* | 10 | 16 | 5 | 4 | 5 | 13 | 2 | 4 | 116 | 113 | 290 | 465 |
| | *11* | 12 | 18 | 5 | 2 | 8 | 11 | 0 | 10 | 300 | 240 | 120 | 360 |
| | *12* | 17 | 16 | 23 | 54 | 8 | 22 | 3 | 4 | 64 | 52 | 140 | 205 |

| Key | |
|---|---|
| Rank for Nc and node counts | Rank for Nc and node time |
| Lowest | Fastest |
| Medium | Medium |
| Highest | Slowest |

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


**Figure 4: Recorded fracture data (Nc, and node counts) and the time taken to undertake Nc and node counts for workshop (WS) participants (P) and groups (G). The data for each attribute has been colour-coded according to where the reported value for the parameter ranked for that circle. Data are presented in the order that they were completed in the workshop.**

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019

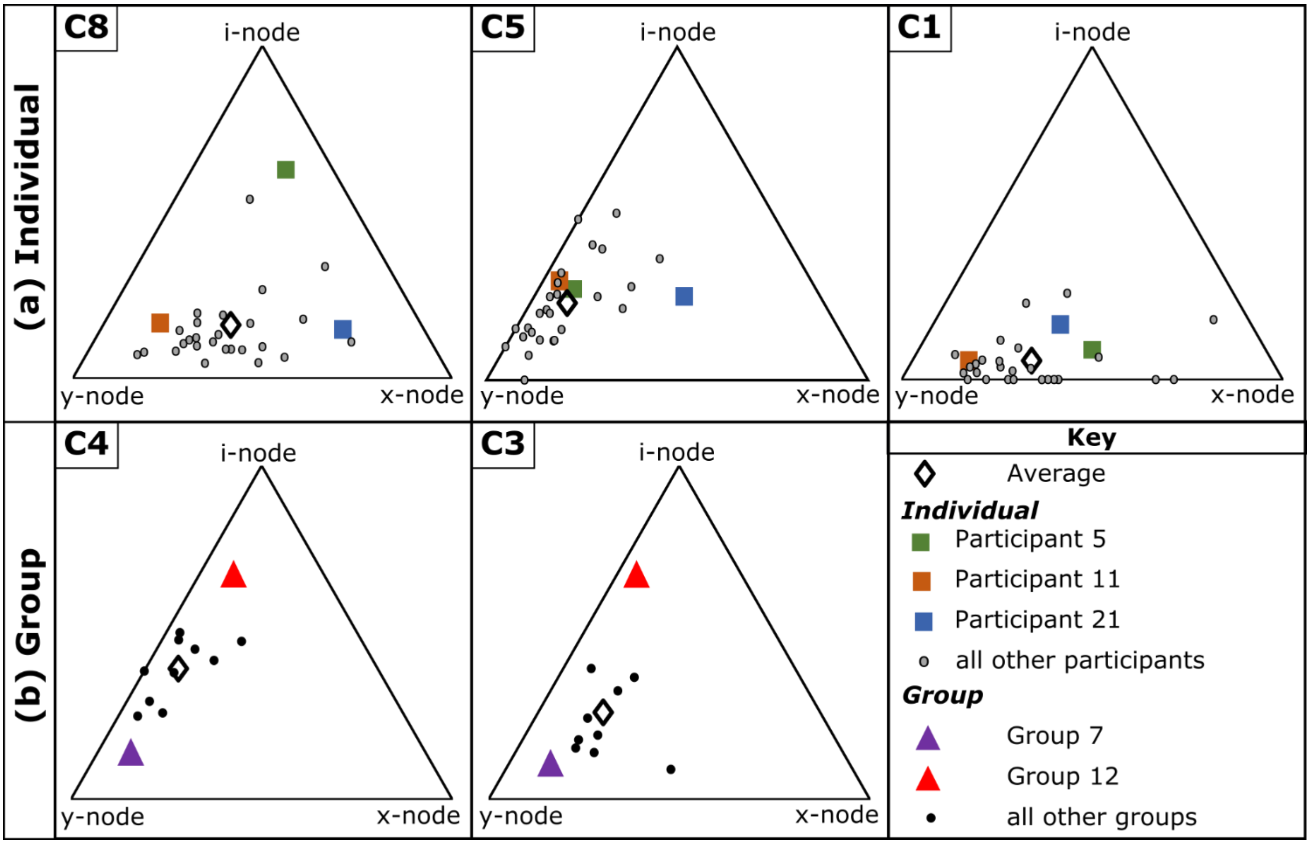




**Figure 5: Node triangles for workshop participants and groups. For individual circles (a), Participants 5, 21, and 11 were highlighted to show the consistency the way participants classified nodes. Participants were selected according the whether they reported a low (P5), medium (P21) or high (P11) node count. Similarly, for group circles (b) Groups 7 and 12 were highlighted as groups who recorded a high and low node count.**

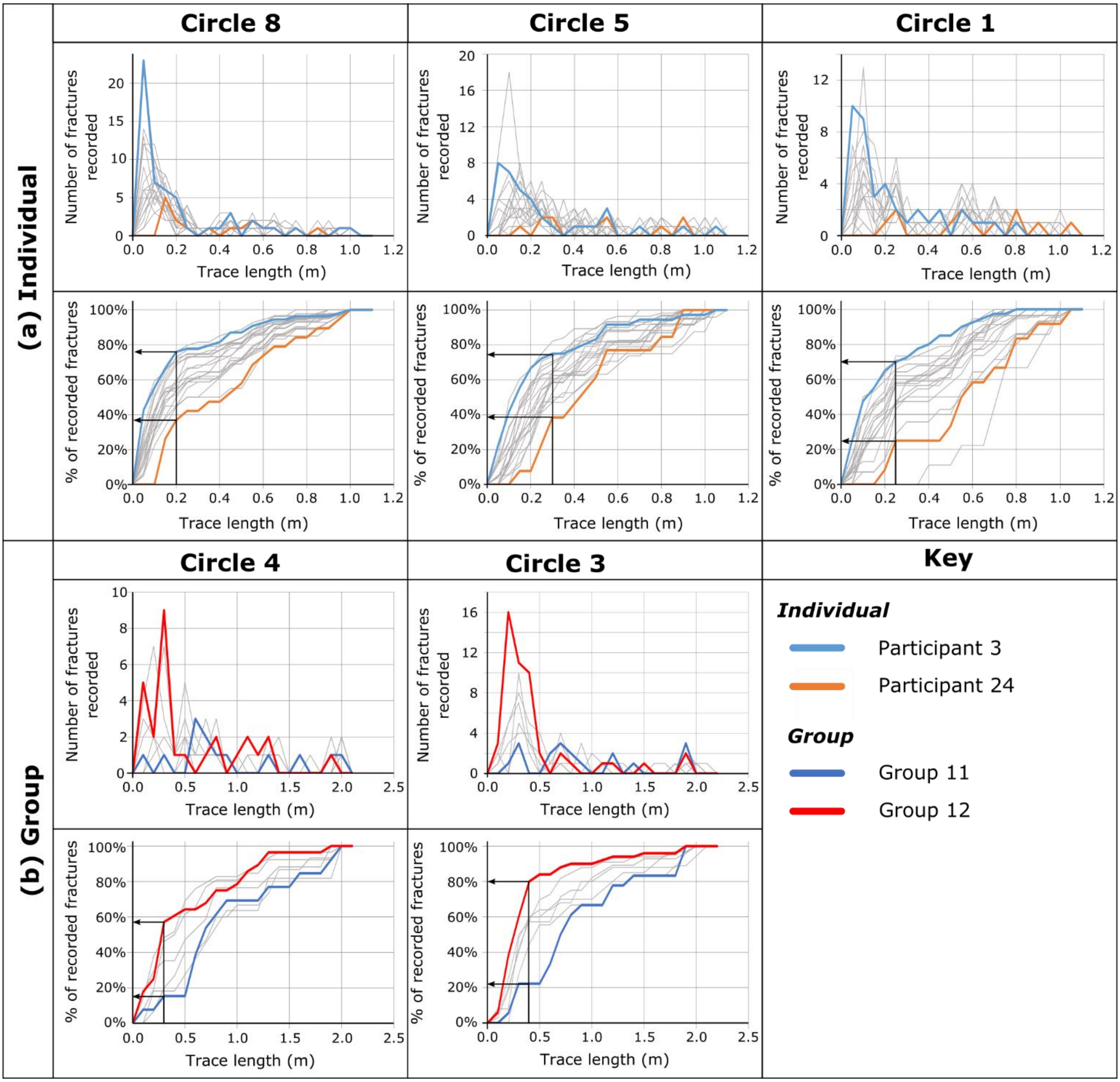
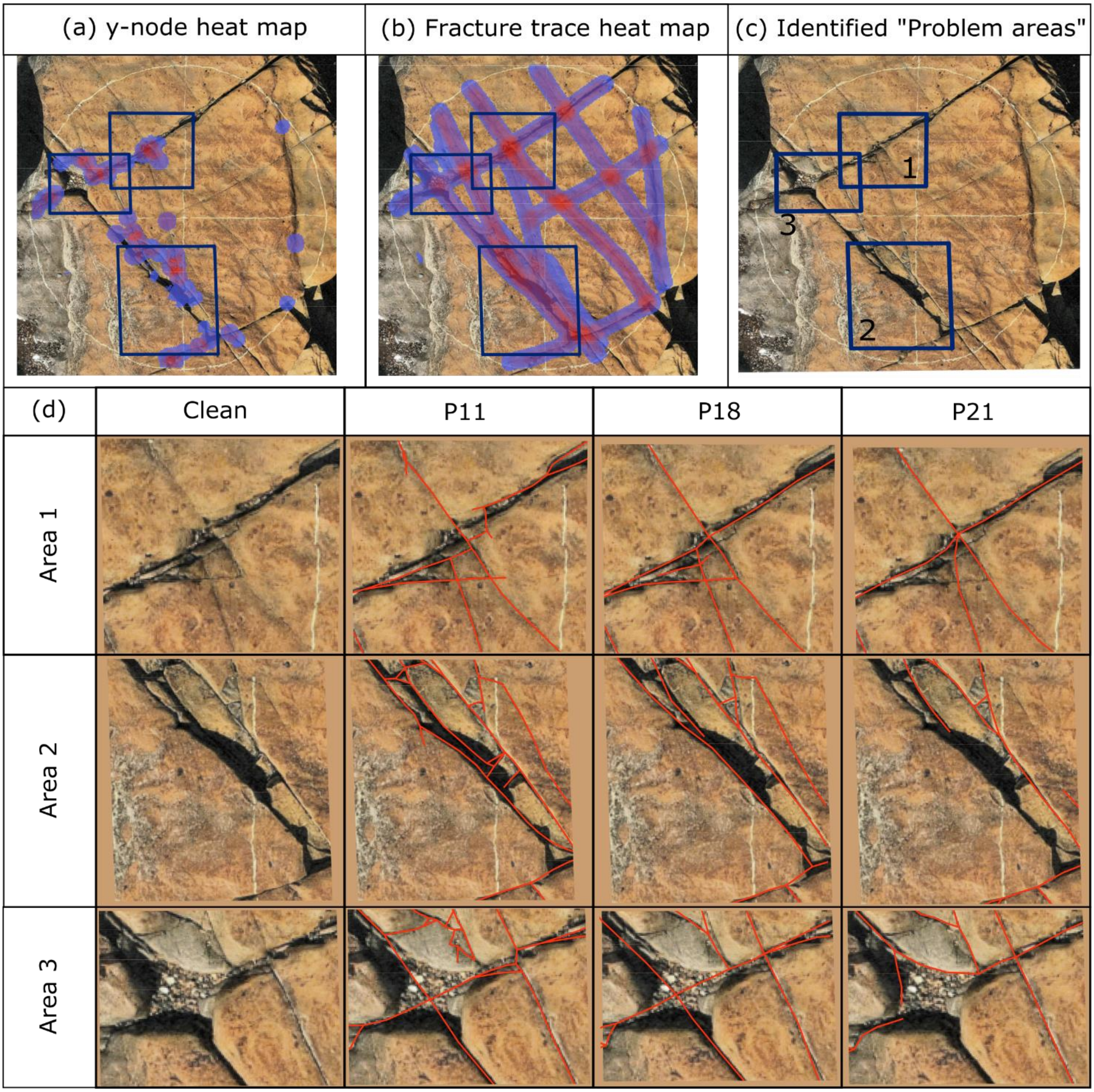



**Figure 6: Fracture trace length distributions for (a) individual and (b) group window sampling data. The results are presented as both histograms and normalised cumulative frequency curves of fracture trace length with bin widths of 0.05 m for individual and 0.1 m for group window sampling data. The range in the relative percentage of small fractures observed in the data is highlighted using Participants and groups who consistently observed a high and low percentage of small fractures (Participant's 3 and 24 and Groups 12 and 11 respectively).**

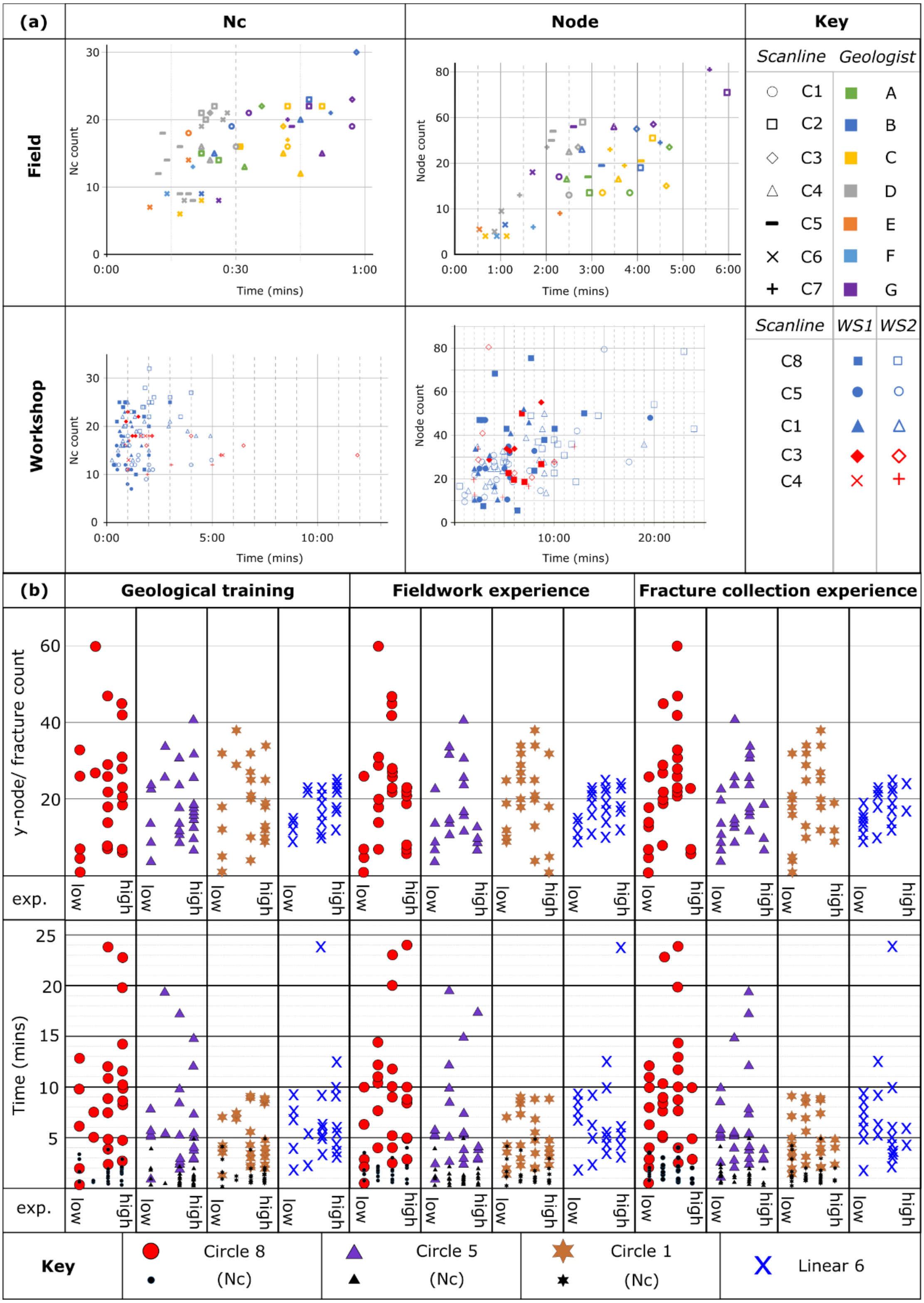
Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


Solid Earth
Discussions
Open Access

d




5 **Figure 7: A detailed study of the areas which cause increased uncertainty in Circle 8. The figure comprises of clean field photographs of Circle 8 with the (a) heat map of y-node point density, (b) heat map of fracture trace density and (c) areas identified as problem areas. In panel (d) the close up of areas 1, 2 and 3 along with the features recorded by Participants 11, 18 and 21 are shown. See text for full description.**


(a)
Nc
Node
Key
Field
Workshop
Nc count
Node count
Time (mins)
Scanline
Geologist
C1 C2 C3 C4 C5 C6 C7
A B C D E F G
WS1
WS2
C8 C5 C1 C3 C4
(b)
Geological training
Fieldwork experience
Fracture collection experience
y-node/ fracture count
Time (mins)
exp.
low
high
Key
Circle 8 (Nc)
Circle 5 (Nc)
Circle 1 (Nc)
Linear 6

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


Solid Earth
Discussions

**Figure 8: The impact of participant experience on the collection of fracture data. (a) The time taken in seconds to record fracture data (Nc and node counts) from circular scanlines both in the field and workshops. (b) The impact of experience on the recorded y-count and number of fractures in individual scanlines and the time taken to complete the workshop tasks.**

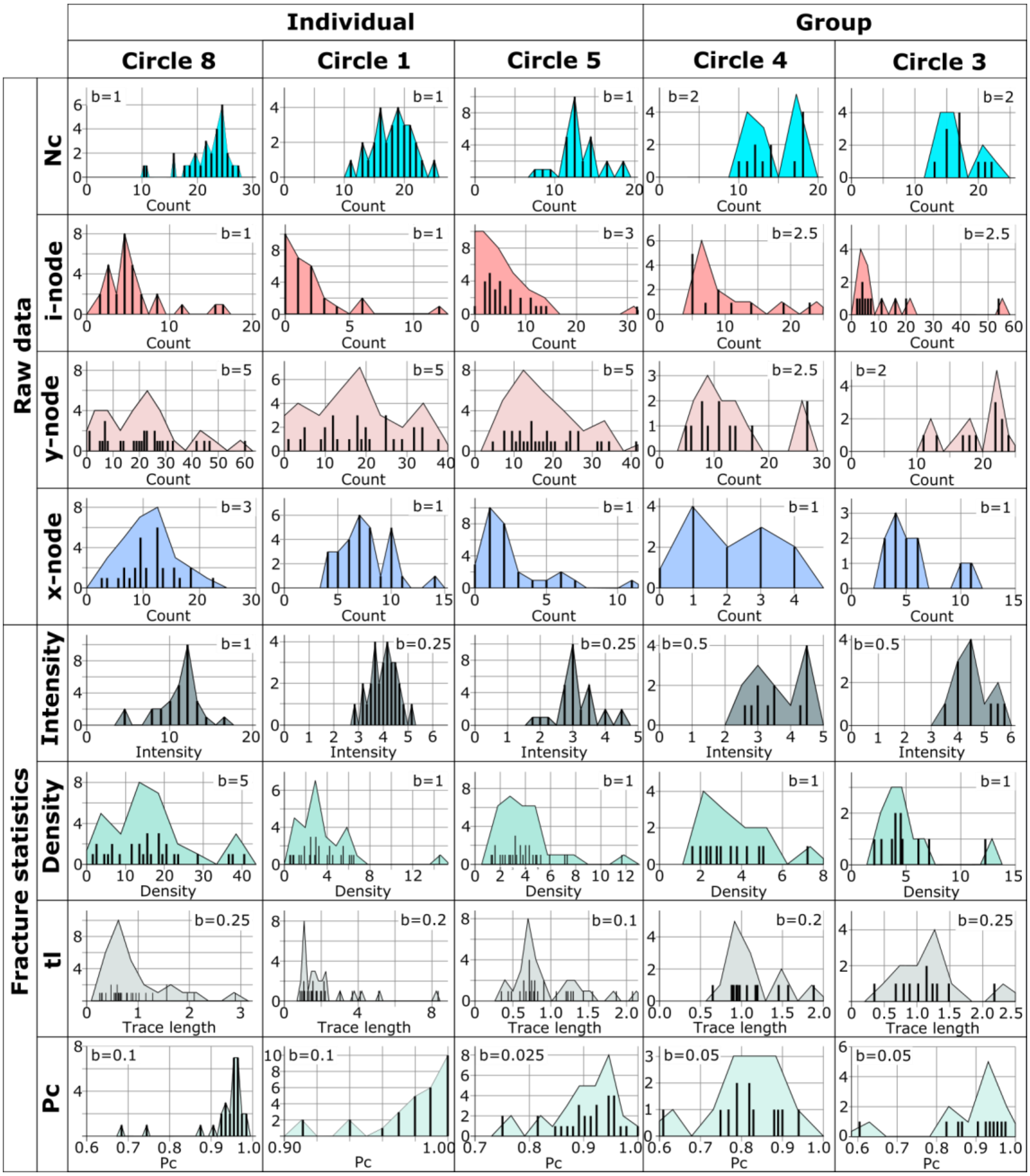



**Figure 9: Topological sampling results for individuals and groups for circular scanlines 1, 3, 4, 5 and 8. Each histogram reports the results for all workshop participants. The statistics have been derived from the data for each participant. Data is presented as both bar charts and shaded histograms with the bin width, b, indicated on the chart. In all cases the y-axis represents frequency and is scaled so the shape of the distributions can be assessed.**

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019

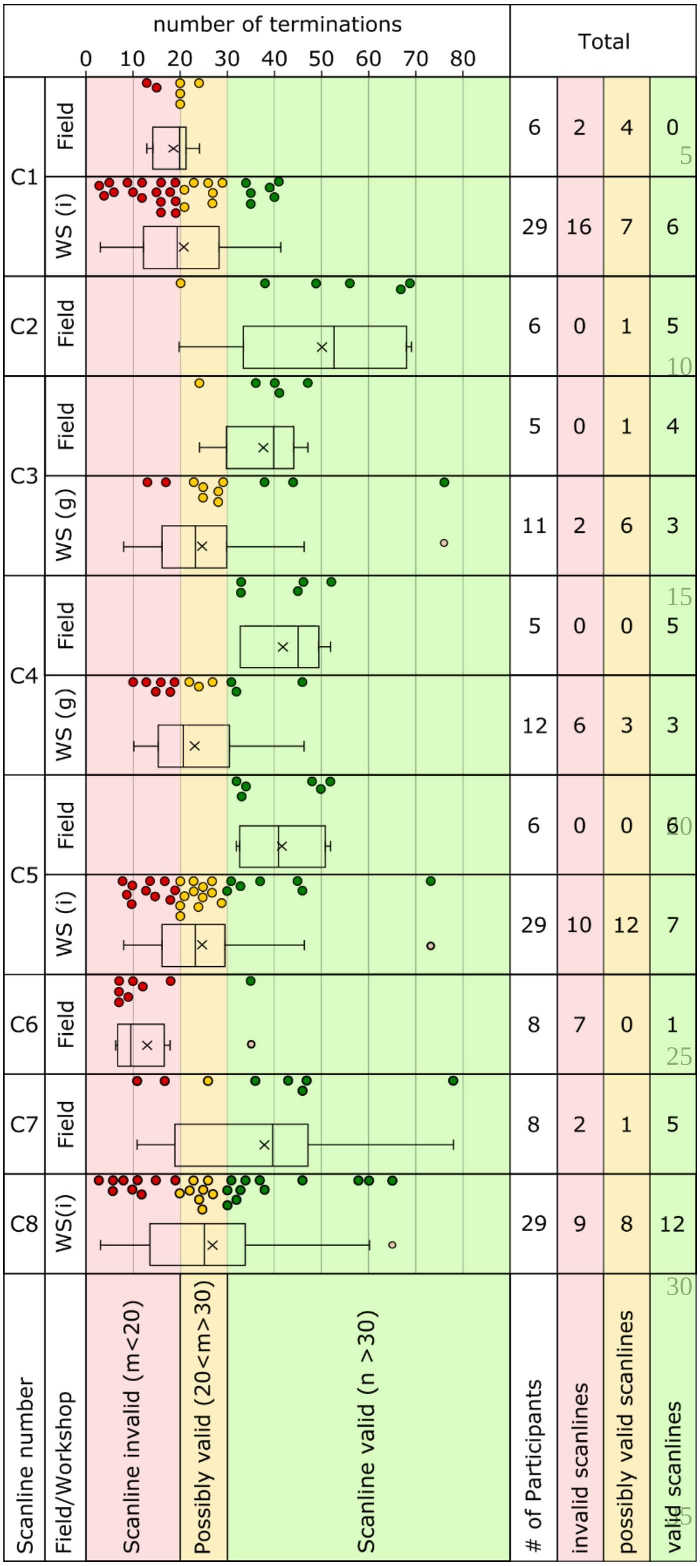
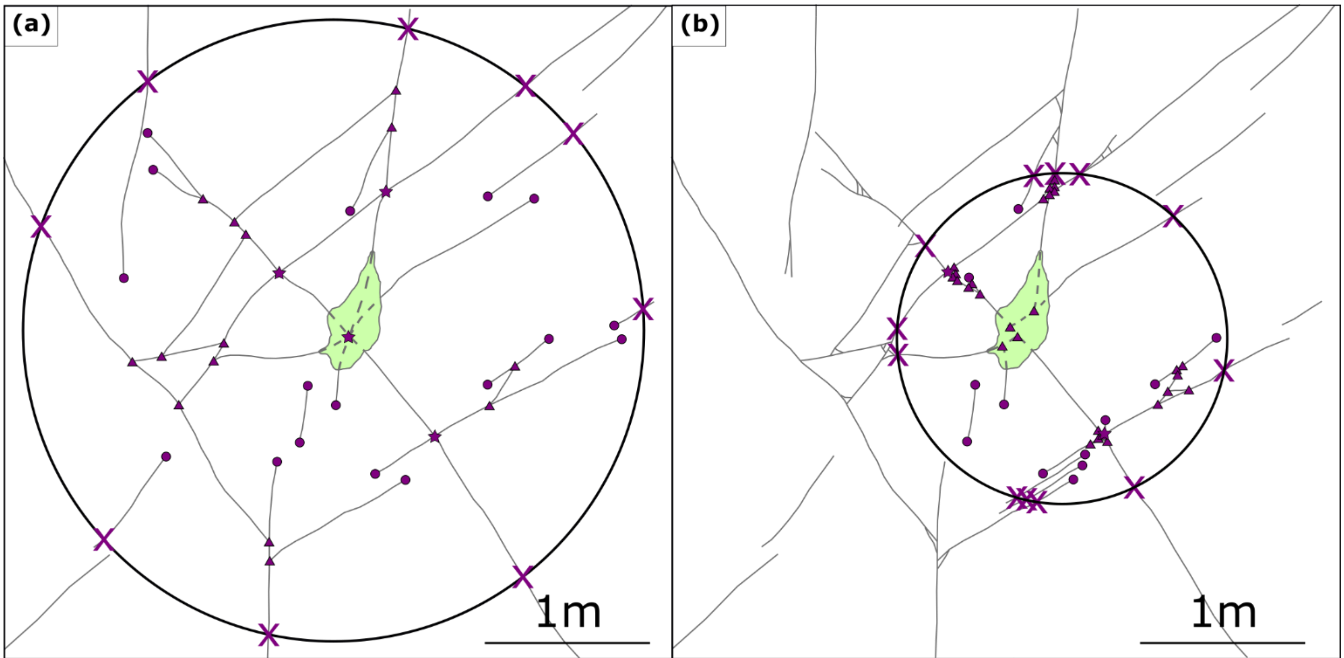



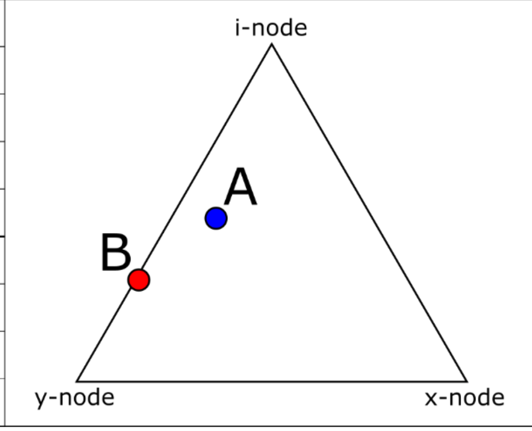

**Figure 10: The effect of subject bias on the validity of circular scanlines. The number of terminations recorded by individuals or groups is displayed for each circle and colour coded depending on where a valid (>30, green), possibly valid (20-30, yellow) or invalid (<20, red) number of terminations were recorded.**

Solid Earth Discuss., https://doi.org/10.5194/se-2018-135
Manuscript under review for journal Solid Earth
Discussion started: 8 January 2019


| (c) | Fracture statistic | A | B |
|---|---|---|---|
| | scanline radius | 3.75 | 1.00 |
| Raw data | Nc | 9 | 12 |
| | i-node | 17 | 12 |
| | y-node | 14 | 28 |
| | x-node | 4 | 2 |
| Statistics | Intensity ($fm^{-1}$) | 0.60 | 3.00 |
| | Density ($FA^{-1}$) | 0.35 | 6.40 |
| | TI (m) | 1.71 | 0.47 |
| | Pc | 0.77 | 0.88 |

**Node triangle**




**Key**

- fracture trace
- inferred fracture trace
- No exposure
- i-node
- y-node
- x-node
- Nc

**Figure 11: The impact of interpreter style on fracture statistics of a synthetic fracture network. (a) statistically valid topological sampling within a circular scanline for a fracture network which only considers the large scale fracture network. (b) statistically valid topological sampling within a circular scanline for the same large scale fracture network as (a), however, also capturing small scale fractures at fracture intersections. (c) The topology attributes (Nc, i-, y- and x-nodes), derived fracture statistics and node triangle of the different interpretations of the fracture network.**