Causes of bias and uncertainty in fracture network analysis

David C.P. Peacock¹, David J. Sanderson², Eivind Bastesen³, Atle Rotevatn¹ & Tor H. Storstein³

¹Department of Earth Science, University of Bergen, Allégaten 41, 5007 Bergen, Norway.
²Ocean and Earth Sciences, University of Southampton, National Oceanography Centre, Southampton, SO14 3ZH, UK.
³NORCE Norwegian Research Centre AS, Nygårdsgaten 112, 5008 Bergen, Norway.

E-mail corresponding author (David Peacock): david.peacock@uib.no

Fault and fracture networks are analysed to determine the deformation history and to help with such applications as engineering geology and fluid-flow modelling. These analyses rely on quantifying such factors as length, frequency and connectivity. Measurements may, however, be influenced by a range of factors relating to resolution, geology, methods used and to the analyst(s). These factors mean that it can be difficult to obtain a single correct solution, with bias and uncertainty being introduced by different analysts, even for something as simple as counting the number of joint intersection points on a well-exposed bedding plane. These problems suggest there are significant issues in comparing databases, for example when using outcrop analogue data to model subsurface data. Our recommendation is that analysts and modellers should be aware of the potential pitfalls in their measurements of structures and, therefore, be more cautious with resultant analyses and models. We suggest that analysts assess their results by testing the reproducibility. Simple ways of doing this include: (1) checking for change in measurements (e.g., fracture frequencies) during the course of a study; (2) remeasuring part of the fracture network to check if the same results are obtained, and; (3) get one or more other analysts to blind-test the fracture network.

Keywords: fracture networks, measurement, resolution, reproducibility, critical evaluation.

Received 21. December 2018 / Accepted 2. March 2019 / Published online 24. April 2019

Introduction

It is natural to have confidence in one’s own measurements, with the perception being that fieldwork and other analyses give clear and unambiguous data. Here, we discuss some of the biases and uncertainties associated with quantitative structural analysis, with particular reference to the study of fracture networks. We give advice about how to critically assess such datasets, especially by testing the reproducibility of the results (e.g., Sander et al., 1997).

The accurate characterisation of parameters such as fracture frequency and connectivity is important in developing flow models for hydrocarbon (e.g., Mohaghegh, 2013; Sun et al., 2016), geothermal (e.g., Pruess & Narasimhan, 1982) and groundwater (e.g., Nueman, 2005) reservoirs, and for predicting fluid flow for contaminant transport in rock (e.g., Bear et al., 1993). A wide range of techniques are used to characterise fracture networks, with topology being of particular current interest (e.g., Diegel, 1986; Huseby et al., 1997; Manzocchi, 2002; Sanderson & Nixon, 2015; Thiele et al., 2016a, 2016b). Topology analyses 2D fracture networks as a system of nodes and branches, where the branches represent the fractures and the nodes represent the tips or intersections of the fractures. One of the basic measurements in topology is the number of nodes (e.g., Andresen et al., 2013), in this case the number of intersection points between fractures. Topology emphasises the relationships between two
or more individual structures, including crossing and abutting relationships of faults and joints, so is useful in determining connectivity of the network, age relationships and likely fluid flow (Sanderson & Nixon, 2015; Peacock et al., 2017, 2018). Note that the terminology used here follows that of Peacock et al. (2016).

To illustrate some of the problems involved in collecting accurate fracture data, and therefore in using fracture data for modelling, we focus on three field examples. The first is the case of joints exposed on Liassic limestone beds on the coast of Somerset, UK (51°11′29″N, 3°14′13″W). The area shows normal faults with throws of up to hundreds of metres, reverse-reactivation of many of these faults, and later strike-slip faults (e.g., Whittaker & Green, 1983; Willems et al., 1997; Peacock & Sanderson, 1999). The joints are approximately perpendicular to bedding. The joints post-date faulting (Peacock, 2001) and were probably created under waning Alpine Orogenic stress (Rawnsley et al., 1998), meaning they have a relatively simple tectonic history. The other two field examples involve analysis of fracture systems in the crystalline rocks of western Norway on the islands of Froya (63°42′59″N, 8°39′07″E; e.g., Storstein, 2017) and Sotra (60°16′27″N, 4°57′39″E; e.g., Fossen, 1998). These crystalline rocks consist of Precambrian gneiss deformed during Caledonian orogenesis and by various phases of post-Caledonian brittle deformation (e.g., Fossen, 1998).

In this paper we discuss some of the sources of bias and uncertainty encountered while measuring fracture networks, how these can be critically evaluated, and the implications for the resulting interpretations and models. We focus on 2D data from outcrops. The aim is to get analysts to think about the problems and limitations of their structural datasets, and to encourage them to think of solutions to these problems. We divide problems encountered into those related to data resolution, geology, methods, and those relating to the analysts themselves. Note that overlap commonly occurs between these factors. The main recommendation is that analysts must be aware of the limitations of their data, so they are able to assess the veracity of the data, and, hence, to properly understand the resulting interpretations and models. We attempt to give a scientific account of the problems involved in measuring fracture networks, but we also give some anecdotal examples and our perceptions of the problems we have experienced in such analyses, hoping that such information will be useful to others.

Reproducibility is a key aspect of science (e.g., Sander, 2007; McNutt et al., 2016). If two analysts cannot obtain similar results from the same rocks, it suggests that any modelling based on such analyses will contain significant uncertainties and potential problems. We suggest that any measurements of fracture networks should be tested for reproducibility, e.g., by seeing how measurements change during the course of a study or by blind testing by other analysts.

Simple examples of counting errors

Selection criteria
We gave 40 students a photograph of 1 m² of a jointed bedding plane (Fig. 1A) and asked them to count the intersection points between the joints. Their results show a considerable spread (Fig. 1B). This simple exercise suggests that it is difficult or impossible to find a unique or ‘correct’ answer to something as simple as the number of points on an exposure surface.

The result shown in Fig. 1B can be attributed to various factors, all related to how the intersection points were selected and counted. In subsequent sections, we analyse factors that may affect the selection and measurement of fractures, but the task of counting involves two sources of variation (or error): (1) the selection of intersections, and (2) the counting of these features. We discuss a number of factors affecting the selection of features associated with fractures, but we start with the counting.

Precision
To test whether the spread of results shown in Fig. 1B was related to the selection and interpretation of the joint intersections or just counting errors, we also gave the students a figure with 213 randomly distributed points (Fig. 1C), which we asked them to count. We chose 213 as this was a number close to the number of intersections in Fig. 1A obtained by the first author of this paper. It also represents a precise integer as opposed to a number such as 210 or 200, which could be a count rounded to two or one significant digit, respectively.

The results of the count by students are shown in Fig. 1D, illustrating that, even in this unambiguous example, people can make mistakes and obtain different results from each other. Of the 32 students who completed the task, 20 (62.5%) got the correct answer, and 28 (87.5%) got within ± 2 of the correct answer, which might, for many purposes, be considered a tolerable level of precision (± 1% of correct value). Of those who made minor counting errors, most undercounted, i.e., missed at least one point, with only 4 (12.5%) double counting a least one point.

Fig. 1C was plotted from a table of data derived using a random number function in a simple computer spreadsheet. The computer could be asked to count these data and should produce the same (correct) number on each occasion. The difference in the performance between people and computer can be considered as the precision of the analyst or the counting device.
Factors affecting fracture surveys

Resolution and data factors

Data from various sources are used for fracture network analysis. Outcrop data may come from wave-cut platforms (e.g., Ryan et al., 2000), cliffs (e.g., Lore et al., 2001), quarries (e.g., Salvini et al., 2017) or tunnels (e.g., Kulatilake et al., 2003), and may be analysed in the field (e.g., Min et al., 2004), from field photographs (e.g., Ryan et al., 2000), aerial photographs (e.g., Mabee et al., 1994) or from satellite images (e.g., Zeeb et al., 2010). Fracture data from wells may be obtained from core or borehole images (e.g., Genter et al., 1997). Subsurface data may also be obtained from seismic surveys (e.g., Tod et al.,

Figure 1. Examples of counting from apparently simple datasets. (A) Photograph of a 1 m² area of a Liassic limestone bedding plane at East Quantoxhead, Somerset, UK. (B) Graph showing the numbers of intersection points of the joints shown in (A), counted by 40 students. The results show a considerable scatter (mean = 187.45, standard deviation = 24.8). (C) Graph of 213 randomly distributed points. (D) Graph showing the numbers of randomly distributed points shown in (C), counted by 32 students. Whereas 20 students correctly counted the correct number of points, 12 counted incorrectly.
These data can be collected from 1D (including well data and core scans from outcrops; e.g., Terzaghi, 1965), in 2D (including outcrop mapping; e.g., Gillespie et al., 1993), and in 3D (including seismic or open cast mine data; e.g., Watterson et al., 1996; Pickering et al., 1997). Each of these datasets is subject to a range of sampling problems, including those relating to scaling relationships of the fracture networks (e.g., Watterson et al., 1996). Various authors have attempted to integrate fracture data obtained from surveys of different resolutions (e.g., Odling, 1997; Odling et al., 1999; Bour et al., 2002; Jones et al., 2009; Strijker et al., 2012).

Fracture lengths are important in characterising the network (e.g., Singhal & Gupta, 2010), and fracture trace lengths are used here as an example of resolution and data problems, which include the following:

- Fractures commonly consist of many interacting and linked segments (Fig. 2), and this can occur at all scales (e.g., Tchalenko, 1970; Bonnet et al., 2001). What appears to be a single fracture when viewed from a distance may be seen to be a fracture zone consisting of a series of smaller segments when viewed in more detail. Some information about the scaling relationships of this segmentation (e.g., Peacock, 2003; Long & Imber, 2011) is therefore needed for accurate fracture (and therefore flow) models. Different people will handle segmentation in different ways. For example, where one person interprets a single vein, another person may interpret a series of linked vein segments (Fig. 3).

- It can be difficult to find the tips of fractures, because fracture aperture decreases below resolution near their tips. For example, it is difficult to locate fault tips from seismic data because even high-resolution seismic data may only detect faults with throws greater than a few metres. The low-displacement tips will not be observed, which influences the measurement of fault lengths and connectivity (Pickering et al., 1997). Watterson et al. (1996) use coal mine plans to show how interpretations of fault lengths and connectivity depend on the resolution of faults (Fig. 4). Peacock (2007) shows how fault networks can show different patterns at different scales of observation. Whatever scaling relationships the fractures (or faults) obey, the implication is that there are some fractures that are too big or small to be sampled at the scale of observation (e.g., La Pointe et al., 1993).

- A fracture geometry that we have encountered in Somerset has caused us problems in measuring fracture networks. We name these tetrahedron

Figure 2. Examples of fracture networks. (A) and (B) Examples of joints on Liassic limestone bedding planes, Somerset, UK (A) Single tip-to-plane linkage, giving one intersection point (node). (B) Double tip-to-plane linkage, giving two intersection points. (C) Fracture corridor (e.g., Cilona et al., 2016) ~0.5 m wide, consisting of a zone of closely-spaced joints (Precambrian gneisses, Søreide, Bergen, Norway). The 'background' joints are more widely-spaced than in the fracture corridor. (D) Zone of deformation bands that have grown, linked and merged to form a zone with high connectivity and frequencies (Miocene carbonates, Malta).
structures because they form four-sided blocks bounded by three joints and a bedding plane (Fig. 5). These tetrahedron structures occur where a later fracture abuts an earlier fracture to form a node at an angle of less than about 80°, but splits closer to the earlier fracture to form a tetrahedron-shaped (four-sided) block of rock. Both veins and joints commonly show these tetrahedron structures. The result is that what initially appears to be a single node is actually three closely-spaced nodes. Whereas we do not have a model to explain tetrahedron structures, we recognise them as a common structural geometry that influences the fracture network characteristics, especially the counting of intersections and, hence, evaluation of the topology of the network.

Data issues from outcrops include uneven measurement surfaces (e.g., non-planar outcrops), incomplete exposure (e.g., vegetation and water cover), and areas of light and shadows on photographs. The changes in such factors with time, due to weathering and erosion of an exposure, mean that there will be reproducibility problems for different analysts studying the same outcrop in the same ways, but at different times.

To address these issues, fracture sampling needs to have defined resolution limits (e.g., sampling only fractures of a certain length), and the resolution of the imaging techniques used must be assessed (e.g., Odling, 1997). This will help with upscaling the geometry of the fracture network to the scales that are of interest to reservoir models. Upscaling is a key issue in using fracture data to build reservoir models, for example how aspects of a fracture network measured in wells or outcrops can be realistically added to the blocks (commonly with dimensions of tens or hundreds of metres) used in a reservoir model (e.g., Warren & Root, 1963). Numerous attempts have been made to upscale fracture data (e.g., Sahimi, 2000; Painter & Cvetkovic, 2005; Lei et al., 2015), but this problem has not been fully resolved.

**Geological factors**

Different lithologies and rock units can have different fracture frequencies (cf. joints in coal and granite) and
Figure 4. Maps of faults from the Yorkshire coalfield at different resolutions (from Watterson et al., 1996, fig. 6). (A) Map showing the fault traces with throws of more than 1 m. (B) Map showing faults with throws of more than 40 m. Notice that the different maps show different fault patterns, frequency and connectivity (Peacock, 2007).
different textures and fabrics (e.g., Ladeira & Price, 1981; Laubach et al., 2018; Procter & Sanderson, 2018). This means that measurements of fracture networks that include more than one lithology should be divided on the basis of those lithologies. Such division may be difficult for some datasets, such as from satellite images, the results for which may be an average for several different lithologies.

Fractures tend not to be perfectly planar, but commonly show bends and steps along their traces at many scales (e.g., Tchalenko, 1970). Such irregularities can make it difficult to map fracture traces, and similar problems are caused by the way fracture traces are exposed on uneven topographic surfaces or non-planar cliffs.

There has been a recent tendency for analysts to just measure fractures in an outcrop, core, etc., without stating whether these fractures are joints, veins, faults, etc. This tendency has been criticised by Manda & Horsman (2015) and by Peacock et al. (2016) because these different fracture types may form at different times and have different implications for kinematics, mechanics and fluid flow. It may be difficult to identify fracture type in the field, and can be impossible from remotely-sensed data (e.g., satellite images). Foliation, bedding and other non-fracture discontinuities may also make fracture identification problematic.

Analytical techniques that overcome such geological problems must be used when collecting fracture network data. For example, fracture distributions in different lithologies must be measured and differences understood. Knowledge of fracture geometries should be incorporated in models, such as measurements of the scaling relationships of fracture trace lengths and scaling of steps (e.g., Peacock, 2003). It is also vital that any analysis of ‘fractures’ in rock exposures and core define what types of fractures are being studied. It is wrong to dismiss faults, veins, joints, etc. as ‘fractures’ (e.g., Manda & Horsman, 2015).

Exposure conditions
Exposure conditions may influence the ability to measure a fracture network. We observe fracture corridors (e.g., Questiaux et al., 2010) on Sotra that contain many closely-spaced fractures when observed on coastal exposures. Inland, however, they are typically preferentially eroded and vegetated, so tend to be mapped as single lineaments. The measurements of fracture intensity and of the number of fracture intersections are therefore strongly influenced by the quality of the exposure. As a result, an individual well-exposed fracture may be represented similarly to a fracture corridor in inland areas.

Weathering and discoloration around fractures, and partial cover by soil, vegetation or water on an outcrop can also cause measurement problems. For example, weathering of bedding planes in Somerset can make identification of joints ambiguous (Fig. 6). Even well-exposed rocks can give ambiguous results if weathered, so the same outcrop may give different results in different seasons or through time as it weathers.

Studies of fracture networks in outcrops are generally confined to areas with high-quality exposure (e.g., Watkins et al., 2015), which may not be representative. For example areas of higher fracture intensity may be eroded and obscured by vegetation. Analyses of fracture
data from aerial photographs and satellite images tend to include areas of poor or no exposure, so some form of extrapolation across poorly-exposed areas has to be made using data from the well-exposed areas. As with data factors, such geological factors may be tackled using stochastic modelling (e.g., Meyer & Einstein, 2002; Doonechaly & Rahman, 2012; Zambrano et al., 2016).

It is possible that problems of missing data, as come from areas of no exposure, may be tackled using an understanding of the variability in the data.

**Method-related factors**

There are potential method-related problems created by the analyst’s decision about what features to include in an analysis, and these commonly relate to resolution issues. For example, *hairline fractures* are commonly reported from core, even though it may not be clear what type of fractures these are (e.g., Wennberg et al., 2013) and such small features are at the limits of measurability. Remote-sensing analyses of lineaments are also prone to the personal judgement of an analyst. Similarly, some seismic interpreters use time-slices to identify potential faults that do not displace reflectors in section view, whereas other seismic interpreters will not do this.

A range of questions also arise when digitising fracture traces to develop network models (Fig. 7), including:

- Should a fracture be ‘snapped’ to another fracture that it appears to intersect?
- Should the entire trace length of a fracture be digitised as a single feature, or should it be divided into segments, for example on the basis of nodes? In the latter case, how should branches be handled?
- Is it possible to be prescriptive enough to cover all possible scenarios with a set of procedures (Fig. 7)? Even then, what biases and artefacts do these procedures introduce to the dataset?

Various tools exist to help with digitising fracture networks (e.g., routines in ArcGIS; Nyberg et al., 2018), but care is needed. For example, some algorithms may ‘snap’ fractures at intersection points when traces approach but do not necessarily meet.

Whereas it is standard practice for scientists to specify the methods they have used in data collection and analyses, some factors may not be obvious to report but will still have a significant effect. For example, we have observed that swapping between different computers, with different mouse devices, keyboards and monitors with different sizes and resolutions, can influence the speed, precision and accuracy of sampling. Bond et al. (2015) show that methods used affect the results of seismic and borehole analyses, while Alcalde et al. (2017b) show that framing bias can be caused by the way seismic images are presented.
Psychological issues. Various psychological issues may influence the ability or willingness of an analyst to measure geological structures. Different analysts may have different attention spans, or the concentration of an individual analyst may vary during the course of a study, or even through the day. Fractures around the limits of observation, or that are closely spaced, tend to be particularly prone to such variations and biases between analysts. The willingness of the analyst to extrapolate (e.g., through areas of poor exposure) influences fracture trace mapping, particularly influencing interpretation of fracture lengths. Similarly, one analyst may identify a single throughgoing fracture while another analyst will characterise the same feature as a series of stepping segments (Fig. 3). This is also true with subsurface data, where extrapolation below seismic resolution may decide whether two faults intersect or not (e.g., Pickering et al., 1997; Rotevatn & Fossen, 2011).

Analyst-related factors

Scientists strive to be honest and fair reporters of truth. There are, however, factors that may over-ride such high ideals. We may be prone to unconscious biases and can make mistakes. The following is an anecdotal list of some of the problems we have encountered or suspect can occur.

Biological issues. Different analysts may have different quality eyesight, so can have different abilities to see structures, thereby producing different results. Fig. 8 shows how different numbers of fractures were identified during the course of a study, which the analyst (Storstein) attributes to the development of repetitive stress injury during digitisation. This reduced sampling speed, forcing him to focus on an area for a longer time, leading to the identification of more fractures. Another common change in conditions during data collection is changes in weather during fieldwork. While changing weather may not reduce the number and accuracy of measurements, it will tend to decrease the speed of measuring.

Experience issues. Different analysts will have different experiences, and the experience of an individual analyst may change during the course of a study, which can influence results (e.g., Bond et al., 2015; Alcalde et al., 2017a). One of the authors (Storstein) found that he tended to measure more fractures from satellite images and aerial photographs as his experience increased (Fig. 8). The style of interpretation also changed, and he found that, instead of viewing just each individual fracture, it became progressively easier to identify fractures as part of a larger structure. Knowledge of the
Figure 8. (A) Map of lineaments mapped on Frøya (Storstein, 2017). The areas were mapped in sequence from area 1 to area 6. (B) Graph showing the numbers of lineaments per unit area for the different areas digitised. The areas were mapped in progressively greater detail as the analysis proceeded.
software and different focuses of different analysts may influence the results. For example, an analyst interested in faults will produce a very different map of an outcrop than an analyst interested in joints.

The above factors vary from person-to-person and over time, so a simple step to reduce their effects is to use the same analyst, operating under similar conditions, throughout a study. If this is not possible, changes in the quality of the interpretation during the course of a study can be tested using subsets of data obtained from different analysts at different times (as shown in Fig. 8).

In industry, and particularly amongst those providing routine services, a common strategy is to set up detailed workflows for the collection, recording and analysis of data, and to train all analysts to use this standardised approach. This often leads to a ‘house style’, which can help when comparing studies carried out at different times and by different analysts. The disadvantage of this is that the workflows can ‘pigeon-hole’ data collection, with observations that need to be treated separately being combined.

Discussion

Researchers must be sceptical about their own data and aware of any biases in their analysis.

Comparison between different datasets

The problems encountered in measuring fracture networks and other geological structures suggest there are difficulties in comparing different datasets, which are each influenced to different degrees by a range of resolution, geological, method-related and analyst issues. In most cases it should be possible to provide metadata on the types of fracture analysed, the host rock, resolution limits, exposure characteristics, measurement methods and the name of the analyst, all of which could provide useful information in comparison of different datasets. Maps of outcrops made from satellite images, aerial photographs and photographs taken from the ground have different resolutions. To some extent, these resolution limits can be assessed and expressed quantitatively. For example, it may be possible to specify a lower limit to the length of fractures included in the sample (Fig. 4; e.g., Pickering et al., 1997). Problems are exacerbated when comparing between fracture data collected in 1D (e.g., scanlines, boreholes), 2D (e.g., outcrop mapping) and 3D (e.g., seismic data).

There are critical implications of, and potential problems in, the comparison of results from different datasets. For example, observer bias and different methods with respect to data collection of fault and fracture data may infect the global database of fault and fracture statistics. Take, for example, data on fault displacements vs. fault lengths collected by different analysts, in different ways, in different locations, for different reasons and for different fault types (for a review, see Kim & Sanderson, 2005). The resultant best-line fits through such clouds of data will be unreliable.

Sampling strategies and experimental design

Many of the issues of bias and uncertainty in fracture data can be reduced, if not eliminated, by the careful design of experiments in which the aim is explicitly expressed by a testable hypothesis, preferably null hypotheses. For example, if we seek to evaluate the fracture intensity in two rock units or locations, we need to compare the variability between the locations with that within the locations. If we recognise the potential sources of bias and uncertainty in our procedures, we can then devise a sampling plan to minimise these effects. We might choose to adopt similar methods, specify similar exposure conditions and use the same investigators operating under similar conditions. Then by observing similar fracture types, under specified resolution limits and noting any other geological features, we should obtain a consistent dataset to allow testing of our hypothesis, with the use of appropriate statistical methods. A recent example of this approach is given by Procter & Sanderson (2018), who use a carefully controlled and explicitly specified sampling plan to evaluate the joint intensity in different limestone layers, with a view to assessing the spatial variability of these joints.

Testing results using the principle of reproducibility

The range of issues discussed in section 3 suggests that it can be difficult to quantify, and therefore to combat, biases in fracture network measurements. It appears that many geologists do not question their data, so our main aim is to highlight these issues and to make people more aware of problems, even if we cannot provide ready solutions. While it may be difficult to combat these sources of error and bias, it is at least possible to estimate what cumulative effects they may have on the data using the principle of reproducibility (e.g., Burns et al., 1976). We recommend that analysts assess their results using three simple tests of reproducibility (Table 1):

1. The data can be tested for consistency, for example to see if there has been a change in fracture frequencies through the course of a study (e.g., Fig. 8).
2. The analyst could try measuring part of the fracture network again to check if they come up with the same results.
Table 1. Some of the questions that should be asked about fracture network datasets to enable critical assessment, and thereby allow the validity of the data (and of subsequent models) to be established. Also shown are three simple ways for testing the reproducibility of the data.

| General questions | Are the data from a reliable source (e.g., published in a reputable peer-reviewed journal)? |
|-------------------|------------------------------------------------------------------------------------------|
|                   | Has any attempt been made to test the validity of the data, or the sensitivity of data to variability? |
|                   | How aware are the authors or analysts of any problems in their analysis? |
| Questions about resolution and data | What is the effective resolution of the dataset? |
|                   | Has a scaling relationship been calculated? |
|                   | How reliable is the upscaling of the geometries? |
| Questions about exposure conditions | What was the quality of the exposure, core, etc., when the survey was undertaken? |
|                   | Did the quality of the exposure change between surveys? |
| Questions about geology | Have the fractures been divided into different types, e.g., faults, veins and joints? |
| Questions about methods | How carefully have the methods used been described? |
|                   | Were reliable or appropriate methods used? |
|                   | What is the size of the database? |
|                   | Have different datasets been merged? |
|                   | Have the relevant data been collected (e.g., apertures and connectivity for fluid-flow modelling)? |
| Questions about analysts | Are the analysts reliable and what is their background? |
|                   | How many analysts were involved in data collection? |
|                   | Why was the analysis undertaken? |
| Tests for reproducibility | Check for changes in consistency during a study |
|                   | Re-measure part of the fracture network to see if the same results are obtained |
|                   | Blind test by one or more other analysts |
3. Ideally, the fracture data can be measured again by one or more other analysts in a blind test, with the results compared (e.g., Fig. 1).

If there is not consistency in the results, it suggests problems in reproducibility. One way of dealing with this would be to reanalyse the data until consistency can be maintained. Another way would be to use the lack of consistency to give a measure of the confidence in the data, with the confidence range being used as an input into modelling.

**Measurement uncertainty as input to models**

Structural geologists commonly supply data about fracture networks to modellers to develop models for fluid flow within fractured reservoirs (e.g., Spence et al., 2014). It is incumbent on geologists to inform the modellers about the biases and uncertainties in the data they have collected, and it is incumbent on the modellers to challenge the geologists about the veracity of the datasets they have supplied. The modellers must be able to make informed decisions about the effects of biases and uncertainties on their models. Fracture network uncertainties may be one of several problems and inaccuracies in the modelling that uses these measurements, including uncertainties in the effects of fracture apertures, fracture roughness, permeability along the fractures, stress state in the reservoir, fluid pressure gradients, fluid viscosity, and the validity of the upscaling model. Uncertainties in fracture networks should therefore be modelled in the same way as other variables and uncertainties (e.g., Caers, 2011), with the aim of producing a realistic spread of possible results from models.

**Use of outcrop analogues**

The data problems discussed above suggests using outcrop analogues can be misleading. Additional problems exist in using outcrop analogues as inputs into sub-surface models. In particular, rocks exposed at the Earth’s surface have a different geological history than rocks in the sub-surface (e.g., Sanderson, 2015), and are commonly much more heavily jointed because of uplift, stress-release and cooling (e.g., Rawnsley et al., 1998). Here we discuss what can be gained by using outcrop analogues.

Although joints are usually more common in outcrops than in the sub-surface, some structures may be equally developed, including faults and veins. Whereas joint networks may be more developed in surface exposures, they may give a ‘maximum’ case of joint frequencies in the sub-surface. Outcrops may also demonstrate particular chronologies and network patterns, and different joint frequencies in different lithologies. They may also indicate spatial and temporal relationships with other structures (e.g., Sanderson, 2015), such as faults (e.g., Peacock, 2001) and folds (e.g., Silliphant et al., 2002; Hanks et al., 2004).

**Conclusions**

Structural analysis requires the accurate measurement of structural features, such as lengths and connectivity within fracture networks. These are important for the analysis and modelling of those networks, including the development of realistic reservoir simulation models. There are, however, a range of issues that reduce the precision and accuracy of measurements of structures. Factors include: (1) resolution and data (e.g., measuring fractures below the resolution of the survey); (2) geological (e.g., zones of fracture clustering); (3) exposure conditions (e.g., degree of cover and accessibility); (4) methods used (e.g., how portions of a fracture, or linked fractures, are digitised); (5) analyst (e.g., different experiences or interests of different analysts).

These factors mean there can be significant differences in measurements of similar fracture networks because of the data used, the geological conditions, the analytical methods, who has undertaken the measurements, and why and when the measuring took place. There may therefore not be a single ‘correct’ answer for a particular fracture network. These various factors suggest there may be significant issues in comparing data about fracture networks obtained from different datasets, obtained in different ways, at different times by different people for different reasons. For example, aerial photographs and field observations are prone to different degrees to biases related to resolution, geology, method-related and analyst issues.

The range of likely sources of bias and error in fracture network measurement means that it will be difficult to assess, and therefore to counteract, each effect. An important first step is for analysts to be aware of these errors and to have some scepticism of their own methods and results. Steps can be taken to check the validity of the results using the principle of reproducibility. This can be done using three simple methods: (1) test the dataset for consistency for changes during the course of a study (e.g., Fig. 8); (2) remeasure part of the fracture network to see if the same results are obtained, and; (3) the dataset can be blind tested by one or more other analysts (Table 1). A lack of reproducibility can be addressed by reanalysing the data until consistency is obtained, or by assessing confidence in the results.

Analysts and modellers need to be aware of measurement problems, and therefore be cautious when using structural data in modelling. There are practical steps that should be taken in any such analysis, including...
References

Alcalde, J., Bond, C.E., Johnson, G., Butler, R.W.H., Cooper, M.A. & Ellis, J.F. 2017a: The importance of structural model availability on seismic interpretation. *Journal of Structural Geology* 97, 161–171. https://doi.org/10.1016/j.jsg.2017.03.003.

Alcalde, J., Bond, C.E. & Randle, C.H. 2017b: Framing bias: the effect of figure presentation on seismic interpretation. *Interpretation* 5, T591–T605. https://doi.org/10.1190/INT-2017-0083.1.

Andresen, C.A., Hansen, A., Le Goc, R., Davy, P. & Hope, S.M. 2013: Topology of fracture networks. *Frontiers in Physics* 1. https://doi.org/10.3389/fphy.2013.00007.

Bear, J., Tsang, T.F. & de Marsily, G. 1993: *Flow and Contaminant Transport in Fractured Rock*. Academic Press, London, 560 pp. https://doi.org/10.1016/B978-0-12-083980-3.50005-X.

Bond, C.E., Johnson, G. & Ellis, J.F. 2015: Structural model creation: the impact of data type and creative space on geological reasoning and interpretation. In Richards, F.L., Richardson, N.J., Rippington, S.J., Wilson, R.W. & Bond, C.E. (eds.): *Industrial Structural Geology: Principles, Techniques and Integration*, Geological Society of London, Special Publication 421, pp. 83–97. https://doi.org/10.1144/SP421.4.

Bonnet, E., Bour, O., Odling, N.E., Davy, P., Main, I., Cowie, P. & Berkowitz, B. 2001: Scaling of fracture systems in geological media. *Reviews of Geophysics* 39, 347–383. https://doi.org/10.1029/1999RG000074.

Bour, O., Davy, P., Darcel, C. & Odling, N. 2002: A statistical scaling model for fracture network geometry, with validation on a multiscale mapping of a joint network (Hornelen Basin, Norway). *Journal of Geophysical Research Atmospheres* 107, ETG 4–1–ETG 4–12.

Burns, K.L., Sheperd, J. & Berman, M. 1976: Reproducibility of geological lineaments and other discrete features interpreted from imagery: measurement by a coefficient of association. *Remote Sensing of Environment* 5, 267–301. https://doi.org/10.1016/0034-4257(76)90055-9.

Caers, J. 2011: *Modeling Uncertainty in the Earth Sciences*. Wiley-Blackwell, Hoboken, New Jersey, 229 pp. https://doi.org/10.1002/9781119995920.

Gilon, A., Aydin, A., Likerman, J., Parker, B. & Cherry, J. 2016: Structural and statistical characterization of joints and multi-scale faults in an alternating sandstone and shale turbidite sequence at the Santa Susana Field Laboratory: implications for their effects on groundwater flow and contaminant transport. *Journal of Structural Geology* 85, 95–114. https://doi.org/10.1016/j.jsg.2016.02.003.

Diegel, F.A. 1986: Topological constraints on imbricate thrust networks, examples from the Mountain City window, Tennessee, U.S.A. *Journal of Structural Geology* 8, 269–279. https://doi.org/10.1016/0191-8146(86)90048-9.

Doonechaly, N.G. & Rahman, S.S. 2012: 3D hybrid tectono-stochastic modeling of naturally fractured reservoir: application of finite element method and stochastic simulation technique. *Tectonophysics* 541–543, 43–56. https://doi.org/10.1016/j.tecto.2012.03.012.

Fossen, H. 1998: Advances in understanding the post-Caledonian structural evolution of the Bergen area, West Norway. *Norsk Geologisk Tidsskrift* 78, 33–46.

Genter, A., Castaing, C., Dezayes, C., Tenzer, H., Traineau, H. & Villemin, T. 1997: Comparative analysis of direct (core) and indirect (borehole imaging tools) collection of fracture data in the Hot Dry Rock Soultz reservoir (France). *Journal of Geophysical Research* 102, 15419–15431. https://doi.org/10.1029/97JB00626.

Gillespie, P.A., Howard, C.B., Walsh, J.J. & Watterson, J. 1993: Measurement and characterization of spatial distributions of fractures. *Tectonophysics* 226, 113–141. https://doi.org/10.1016/0040-1951(93)90114-Y.

Hanks, C.L., Wallace, WK., Atkinson, P.K., Brinton, J., Bui, T., Jensen, J. & Lorenz, J. 2004: Character, relative age and implications of fractures and other mesoscopic structures associated with detachment folds: an example from the Lisburne Group of the northeastern Brooks Range, Alaska. *Bulletin of Canadian Petroleum Geology* 52, 121–138. https://doi.org/10.2113/52.2.121.

Huseby, O., Thovert, J.F. & Adler, P.M. 1997: Geometry and topology of fracture systems. *Journal of Physics A: Mathematical and General* 30, 1415–1444. https://doi.org/10.1088/0305-4470/30/5/012.

Jones, R.R., McCaffrey, K.J.W., Clegg, P., Wilson, R.W., Holliman, N.S., Holdsworth, R.E., Imber, J. & Waggott, S. 2009: Integration of regional to outcrop digital data: 3D visualisation of multi-scale geological models. *Computers and Geosciences* 35, 4–18. https://doi.org/10.1016/j.cageo.2007.09.087.

Kim, Y.S. & Sanderson, D.J. 2005: The relationship between displacement and length of faults: a review. *Earth-Science Reviews* 68, 317–334. https://doi.org/10.1016/j.earscirev.2004.06.003.

Kulatilake, P.H.S.W., Um, J.G., Wang, M., Escandon, R.F. & Narvaz, J. 2003: Stochastic fracture geometry modeling in 3-D including validations for a part of Arrowhead East Tunnel, California, USA. *Engineering Geology* 70, 131–155. https://doi.org/10.1016/S0013-7952(03)00087-5.

Ladeira, F.L. & Price, N.J. 1981: Relationship between fracture spacing and bed thickness. *Journal of Structural Geology* 3, 179–183. https://doi.org/10.1016/0191-8111(81)90013-4.

La Pointe, P.R., Wallmann, P.G. & Dershowitz, W.S. 1993: Stochastic estimation of fracture size through simulated sampling. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts* 30, 1611–1617. https://doi.org/10.1016/0148-9622(93)90165-A.

Laubach, S.E., Lamarche, J., Gauthier, B.D.M., Dunne, W.M. & Sanderson, D.J. 2018: Spatial arrangement of faults and opening-mode fractures. *Journal of Structural Geology* 108, 2–15. https://doi.org/10.1016/j.jsg.2017.08.008.

Lei, Q., Latham, J.P., Tsang, C.F., Xiang, J., La, P. 2015: A new approach to upscaling fracture network models while preserving geostatistical and geomechanical characteristics. *Journal of Geophysical Research* 120, 4874–4897. https://doi.org/10.1002/2014JB011736.

Long, J.J. & Imber, J. 2011: Geologic controls on fault relay zone scaling. *Journal of Structural Geology* 33, 1790–1800. https://doi.org/10.1016/j.jsg.2011.09.011.
Spence, G.H., Couples, G.D., Bevan, T.G., Aguilera, R., Cosgrove, J.W., Daniel, J.M. & Redfern, J. 2014: Advances in the study of naturally fractured hydrocarbon reservoirs: a broad integrated interdisciplinary applied topic. In Spence, G.H., Redfern, J., Aguilera, R., Bevan, T.G., Cosgrove, J.W., Couples, G.D. & Daniel, J.M. (eds.): Advances in the Study of Fractured Reservoirs, Geological Society, London, Special Publications 374, pp. 1–22. https://doi.org/10.1144/SP374.19.

Storstein, T.H. 2017: Onshore-offshore comparative study of basement geology and structure onshore/offshore Trøndelag. Unpublished MSc thesis, University of Bergen, 266 pp.

Strijker, G., Bertotti, G. & Luthi, S.M. 2012: Multi-scale fracture network analysis from an outcrop analogue: a case study from the Cambro-Ordovician clastic succession in Petra, Jordan. Marine and Petroleum Geology 38, 104–116. https://doi.org/10.1016/j.marpetgeo.2012.07.003.

Sun, J., Gamboa, E.S., Schechter, D. & Rui, Z. 2016: An integrated workflow for characterization and simulation of complex fracture networks utilizing microseismic and horizontal core data. Journal of Natural Gas Science and Engineering 34, 1347–1360. https://doi.org/10.1016/j.jngse.2016.08.024.

Tchalenko, J.S. 1970: Similarities between shear zones of different magnitudes. Bulletin of the Geological Society of America 81, 1625–1640. https://doi.org/10.1130/0016-7606(1970)81[1625:SBSZOD]2.0.CO;2.

Terrzaghi, R.D. 1965: Sources of errors in joint surveys. Geotechnique 15, 287–304. https://doi.org/10.1680/geot.1965.15.3.287.

Thiele, S.T., Jessell, M.W., Lindsay, M., Ogarko, V., Wellmann, J.F. & Pakyuz-Charrier, E. 2016a: The topology of geology 1: topological analysis. Journal of Structural Geology 91, 27–38. https://doi.org/10.1016/j.jsg.2016.08.009.

Thiele, S.T., Jessell, M.W., Lindsay, M., Wellmann, J.F. & Pakyuz-Charrier, E. 2016b: The topology of geology 1: topological uncertainty. Journal of Structural Geology 91, 74–87. https://doi.org/10.1016/j.jsg.2016.08.010.

Tod, S., Taylor, B., Johnston, R. & Allen, T. 2007: Fracture prediction from wide-azimuth land seismic data in SE Algeria. The Leading Edge 26, 1154–1160. https://doi.org/10.1190/1.2780786.

Warren, J.E. & Root, P.J. 1963: The behavior of naturally fractured reservoirs. Society of Petroleum Engineers Journal 3, 245–255. https://doi.org/10.2118/426-PA.

Watkins, H., Bond, C.E., Healy, D. & Butler, R.H.W. 2015: Appraisal of fracture sampling methods and a new workflow to characterise heterogeneous fracture networks at outcrop. Journal of Structural Geology 72, 67–82. https://doi.org/10.1016/j.jsg.2015.02.001.

Watterson, J., Walsh, J.J., Gillespie, P.A. & Easton, S. 1996: Scaling systematics of fault sizes on a large-scale range fault map. Journal of Structural Geology 18, 199–214. https://doi.org/10.1016/S0191-8141(96)00045-9.

Wennberg, O.P., Casini, G., Jahanpanah, A., Lapponi, F., Ineson, B., Wall, B.G. & Gillespie, P. 2013: Deformation bands in chalk, examples from the Shetland Group of the Oseberg Field, North Sea, Norway. Journal of Structural Geology 56, 103–117. https://doi.org/10.1016/j.jsg.2013.09.005.

Whittaker, A. & Green, G.W. 1983: Geology of the Country Around Weston-super-Mare. Memoir of the Geological Survey of Great Britain, Sheet 279 and parts of 263 and 295, scale: 1:50,000.

Willetts, E.J.M., Peacock, D.C.P. & Aydin, A. 1997: Nucleation and growth of strike-slip faults in limestone. Journal of Structural Geology 19, 1461–1477. https://doi.org/10.1016/S0191-8141(97)00056-4.

Zambrano, M., Tondi, E., Korneva, I., Panza, E., Agosta, F., Janiseck, J.M. & Giorgioni, M. 2016: Fracture properties analysis and discrete fracture network modelling of faulted tight limestones, Murge Plateau, Italy. Italian Journal of Geosciences 135, 55–67. https://doi.org/10.3301/IJG.2014.42.