

Comment for Dichiarante et al.

Title: Fracture attribute scaling and connectivity in the Devonian Orcadian Basin with implications for geologically equivalent sub-surface fractured reservoirs

This paper provides an extensive dataset investigating fracture properties over several orders of magnitude using 1D and 2D approaches. The authors provide evidence for power-law behavior over 8 orders of magnitude for fractures (faults, joints, and veins) in the Middle Devonian sandstones of Northern Scotland. The study represents one of the few multi-scale approaches to fracture characterization, and as such is of clear importance and relevance to the readers of *Solid Earth*. However, I have a few concerns that I would like to discuss, and in support of the open review system a number of minor points throughout the manuscript.

Major comments

1. The analysis of several sets and fracture types (fault, vein, and joints) within the same population.

Overall, I strongly agree that there needs to be more multi-scale analysis into fault and fracture attributes, and so far many researchers have remained in their 'safe space' (e.g. field geologist rarely assessing more than an outcrop scale). However, the underlying processes behind the formation of faults and joints are completely different, the mineralization can occur at several times during the structural evolution under different stress states. Additionally the underlying controls have been shown to change through time. For example, pre-existing structures can act to limit fault/joint propagation (e.g. Andrews et al., *in review*; Wilkins et al., 2001), and structural diagenesis change the nature of the mechanical stratigraphy (e.g. Laubach et al., 2010).

You provide a comprehensive structural evolution of the basin in section 3.1 and split the structures into three groups. While you state that you predominantly focus on G3 structure (L226-227), you also look at where G1 faults have been reactivated. Have these structures been assessed separately from the rest of the trend? It would be interesting to know whether the reactivation either caused reactivated G1 structures to have larger or shorter trace lengths. The reversal of the stress states from G1 to G3 will also likely to have caused the propping open of reactivated faults from G1. Has this had any effect on the aperture of these features.

With this in mind, along with the clear age relationships presented in the paper I find it odd that the fractures are not considered as sets. I think this would lead to some sub-trends falling out of the data and explain some of the plots that display trend changes.

I wonder about the viability of using 'kinematic aperture' when investigating veins, particularly when comparing this aperture with joints whereby the aperture could have been significantly altered by later tectonic events. I would also suggest that the kinematic aperture will always plot higher than you would expect for a given length due to the propping open of voids, or multiple crack-seal events.

2. Subjective bias during the digitization of fracture traces, with particular reference to the 2D analysis.

My key concern regarding the digitization, and later analysis, of the fracture networks relates to the 2D analysis presented in Figure 10-11. In a number of places (see attached figure for rough outlines for the circular window) it appears as if 'weathering' boundaries, or indeed the ruler for the circular window has influenced topological sampling. Several i-nodes may be observed both along the boundary of the sand and gravel to the east of Figure 10a, and along the N-S trend of the ruler. This

pattern is also observed in the density plots presented in Figure 10b. A less stark, but still clear effect, is the presence of weathered sections to the NE and SW of 10b.

I imagine that this is due to feature not being interpreted under areas of no-exposure, something that was shown to greatly increase the level of subjective bias during the collection of fracture data (Andrews et al., 2019). The knock on effect on the topology will be to introduce more i-nodes into the system and hence decrease the estimate of connectivity. If a directional connectivity is then considered, the effect of the ruler would produce a artifact and 'barrier' to flow in the NS direction when there are in fact many fractures that join the NS and NE-SW sets.

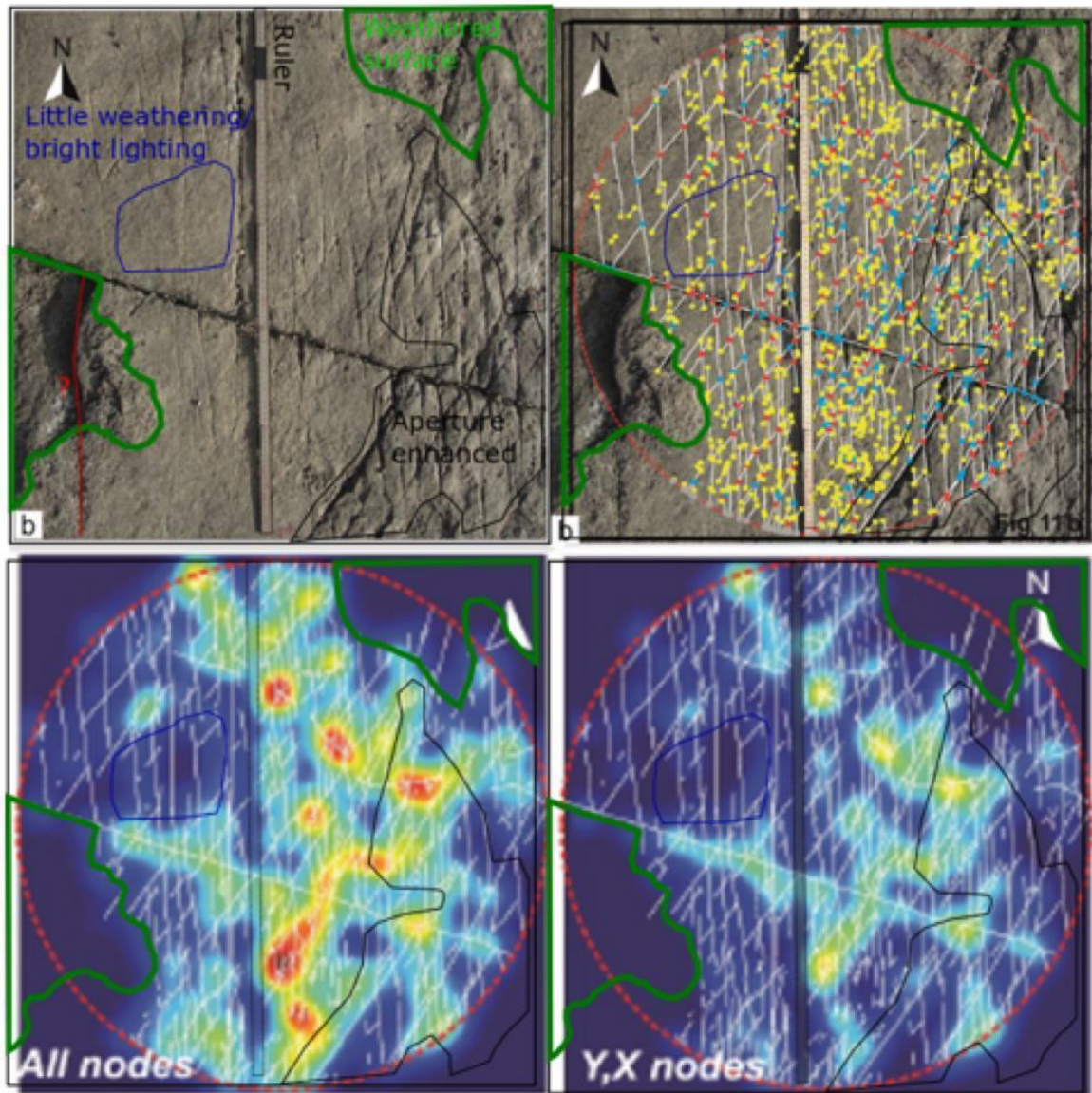


Figure: Areas of increased subjective bias due to exposure bias and the presence of the ruler in the field photograph.

3. Power-law vs log-normal

I found this one of the most fascinating things about this manuscript and felt it raised some very interesting points of discussion. Overall, I am not sure which way to fall on this argument and I am certain that a power-law would do a poor job of describing my data from the Carboniferous (coal

measures). For units which display little mechanical stratigraphy, or the mechanical stratigraphy has been 'lost' prior to deformation then I can see how power-law relationships could be very powerful. I wonder how many of the scale dependent mechanical boundaries (e.g. bedding vs facies) are hidden in the scatter of power-law relationships, particularly as it is often difficult to sample multiple scales of either at a single site. Therefore, could using the 'scatter' in a power-law relationship help us input some useful geological data into geo-models even if the subtleties are lost? Only similar studies in different lithologies and tectonics settings can hope answer my questions. How much scatter is 'ok' to produce a good enough geo-model?

Although I believe there are several points I don't agree with, particularly in the clumping of 'fractures' and the way in which some of the 2D data was digitized, the manuscript has presented an interesting dataset and judging by the discussion forum began an interesting debate. Overall, I think given a greater discussion of limitations this manuscript will make a useful contribution to the field and I look forward to reading the final version.

Line by line comments

L34-37: As an opening sentence I feel this opens the manuscript up to criticism, in particular clumping the faults, joints, and veins. As mentioned in my major comment 1, I have a fundamental issue with combining these datasets, and I think many others will question it too.

L56: What do you mean by '3D', do you mean truly 3D using such methods as xCT or 2.5D using outcrop photogrammetry or LiDAR?

L61-63: Strongly agree with the need for this work, however, I feel too many sub-sets have been clumped together, see major comment 1.

L71-71: In assessing the connectivity of a network is it worth considering the effect of the change in scale. Several authors have pointed out that the comparison on connectivity over many scales can lead to issues, and was one of the key drivers for Olsen et al., I believe in their 2004 paper, to assess aperture instead of trace length or connectivity. Subtle changes in observation also lead to differences in interpretation (Andrews et al., 2019; Peacock et al., 2019), with higher scatter observed at higher resolutions (Scheiber et al., 2015). Has the role of subjective bias on this dataset has been considered?

L79-82: I feel this sentence needs to acknowledge the limitation, otherwise it has the potential to lead to authors attempting to 'read-off' your scaling relationships without considering the geological implications.

L89: change 'sizes' to trace length. Also is all trace length bed-parallel? Or does the dip of bedding vary across the section? Do the authors have any inclination of the aspect ratio of faults, joints, and veins, and hence any controls of mechanical layering.

L98: with fracture orientation collected, and a good handle on the geological evolution, I am surprised that no analysis of sets was considered. If it made no difference to the scaling relationships then this is an important observation, however, I would be surprised!

L98-99: I think the textual and fill aspect of this paper is a nice aspect of the multi-scale analysis.

L101: With a number of the features being filled, I wonder the applicability of using the connectivity of the whole network? If we are looking to investigate flow surely it is best to only include those features which are conducive to flow, unless you feel features can be 'stimulated' for improved recovery.

L103: 'Kinematic aperture' – I have concerns on the applicability of using this for veins, particularly for veins that display multiple crack-seal events. This will always cause your feature to display an aperture higher than expected for a given trace length, as the feature will have been reactivated along, or through, the previous mineralised fill. Due to aperture being one of the key controls on DFN modelling, should we instead not be using the effective aperture?

L119-121: Log normal distributions can also be caused by pre-existing structures influencing the growth of later structures.

L134: I feel with any multi-scale analysis like this the effect of scale on the interpretation of features is important, and such the role of 'subjective bias' should be considered. A slew of papers have recently come out on the subject including Andrews et al., 2019; Peacock et al., 2019; Scheiber et al., 2015; Shipton et al., 2019; Gibson et al., 2019). What scale we make our observations has a fundamental effect on how we interpret a network.

L139: Does a power-law relationship fit all orientations?

L141-143: I think this is a KEY point of the paper and in my opinion the strongest part and as such should not be hidden away in the supplementary information. It provides a very useful practical example of how the fundamental work outlined by Healy et al., 2017 and Rizzo et al., 2017 can be applied to an extensive dataset. I think figures S2 and S3 should be incorporated into the main text, if you felt that would be too many figures then I don't think all examples from figure 7 need to be included, and 1-2 could be included as a combined figure with S2 and S3.

L155: What was a) the scale of resolution of all maps/photos and b) the 'digitisation' scale (see Major point 2).

L169: state why plotting it in a ternary diagram is useful, e.g. Mauldron et al., 2001.

L191: With clustering discussed, I wonder how you look at your connectivity and how this varies across the mapped areas at different scales. How does this effect other fracture attributes (i.e. in the high intensity zones is the trace length lower? In your smaller scale analysis how many samples do you think are a) away from faults and representative the 'background', b) close to faults and potentially representing a small but important sub-set of your dataset, and c) missing faults due to coastal erosion etc. obscuring areas of high fracture intensity (i.e. exposure bias, Shipton et al., 2019).

L233-235: you state that G3 structures represent the best analogue for the Clare field, however, on line 188 you state the areas formed under different tectonic settings. I wonder if the authors could comment on this?

L255: what scale were the lineaments mapped at?

L268-270: It would be useful to include the length of the scanline in the main text here to remind the reader on the scale of observation.

L272-273: was there any difference in fracture properties between the orthogonal scanlines?

L278-280: Interesting point, it's rare to see two sets that share fracture properties and I assume they formed where σ_1 and σ_2 were very similar? However, I would still like to see the trace length distributions split by orientation as this would back up your interpretation that they can be classified under the same population one set didn't affect the others growth.

L282: was it just Ni, Ny, and Nx that was connected, or whether it terminated into a stratigraphic layer?

L286-287: consider including the percentage of connected branches here, what is 'intermediate and 'high' connectivity?

L290: Although I like the use of microstructures in this way, I do wonder how representative the authors feel a sample taken within a fault zone is? Particularly with the local rotation of stress fields commonly observed here.

L297: was spacing measured orthogonal for all sets? Slightly confused on how spacing was extracted here. Was the Terzaghi correction undertaken for all features aligned obliquely to the scanline, and was spacing considered by sets?

L303: MLEs are the 1st results you present.. another reason why the supplementary information should be incorporated into the main text.

L309-312: I think this paragraph is unnecessary and misleading (your only centred due to censoring issues) and I think you should just reference the figure.

L322-325: This is a very interesting point and raises questions on what controls the mechanics of the rock mass at the time of faulting/jointing/veining. Another interpretation is that the scale at which these occur is variable, so the clear change you observe in the seminal mechanically layered sequence work is 'watered' down by a variability in the thickness of the layer. I was not totally convinced from the data that it had no effect, and also for the wave-cut platform examples it appears that the digitised network was on the bedding plane. This means that the mechanical stratigraphy will not have been encountered in that plane of observation.

L364: I find the fact you see this relationship over 4 orders of magnitude very interesting as I would expect the processes to change. Your data therefore suggests that this change is either hidden in the scatter or that there rocks display a continuous set of processes across these scales.

L367-370: the point of this paragraph is unclear, consider rewording or removing

L389: is this relationship in veins not due to the use of 'kinematic aperture'? with vein fill either 'propping up' the void space during periods of high fluid pressure such that aperture is not reduced with further tectonism, or where multiple crack-seal events occur the 'effective' aperture at the time of fluid flow is along, or cutting, the previous fill.

L396-L397: How does bedding vary across the section? How many structure cut across stratigraphic layering?

L429-431: The different number of nodes in each area makes this difficult to compare in the text, consider changing to a ratio.

L436: you state that ENE trending structures cross cut NNW trending structures, does this have no effect on your fracture attributes?

L439: How much of this change in connectivity do you think is due to the change in scale of observation? Nixon et al., 2012 provide a nice discussion on the effect of scale on connectivity.

L444-445: what is the fill of the ENE trending faults, and how will this effect the connectivity of the system? Does this cause an increase in fracture intensity, or is it only that the earlier fractures are offset by the later structures such that the number of nodes increases?

L465-467: Very interesting point!

Section 8.1, Paragraph 1: Please split this paragraph! It becomes unfocused contains several distinct points. Potentially on Line 483?

L473: I think the fact that boreholes often under sample steep features is worth mentioning in this manuscript.

L474-475: I remain to be fully convinced that these populations can't reliably be described using a log-normal distribution, however, can see the usefulness and power in using the Power-law. I wonder whether because many of the controls that underpin the log-normal distribution (e.g. mechanical stratigraphy) are scale dependent (e.g. formation, facies, bed, sub-bed, lamination), whether these are masked within a power-law relationships scatter. It may well be difficult to sample many scales of mechanical stratigraphy, so maybe the authors approach is the best we can do to inform our modelling.

L481-483: Some of the local slope variability also appears to be due to sampling bias due to the orientation of the scanline, extent of exposure, and orientation of the fracture.

L512: I would be fascinated to see how this translates to some of the datasets I've been working with in the Carboniferous. I suspect the strong mechanical stratigraphy will cause several scale-dependent log-normal populations depending what mechanical facies the fracture intersects (and offset for faults).

Section 8.2: I feel you are missing a key point by raising what is the norm for modelling in fracture controlled plays. The lack of geological input data means that even with several limitations, this work is highly valuable in particular in mature basins. I think the 'purpose' of this paper would be strengthened by weaving this into the introduction and discussion, and potentially highlighting how poor fractures are sampled in the sub-surface.

L542: very interesting! Do you have a handle on how much the 'effective' aperture is left after the propping material is discounted? This would provide a more reasonable figure to inform fracture modelling in these lithologies.

Discussion and Conclusions:

I feel that these sections could be slimmed down a bit and the paper loses a little focus. Overall, however, it is clear the importance of studies like this in forwarding our ability to predict fluid flow in fracture dominated plays.

Fig1: Overall, the figure could do with a little redrafting. the arrows to the censored fractures appear too much like fractures in panel (a), consider changing the colour? In panel (b) the title 'circular scanline window' should in fact be 'circular window' as no scanline analysis is shown (the scanline refers to the counting of 'n-nodes' and represents a 1D technique). The picture for Box counting could be increased to better illustrate the technique.

Fig2: consider adding 'u' nodes to the censored fractures on panel (d) and a branch triangle should be included.

Fig3: A zoomed out scale map should be included for non-UK based readers who may not be familiar with northern Scotland. The attribution on panels a) and c) is unclear and should be typed out as apposed to a screen shot. Panel C) misses it's panel label and a north arrow. In panel C it appears that there are two fracture, one of which is censored at a different trace-length compared to the

other due to the exposure and orientation of the coastline relative to the fracture sets. The weathering patterns on (d) appear to show that the linear scanline trends sub-parallel to a set of penetrative fractures that run sub-parallel to the post. If this is the case then this set will be drastically under sampled leading to a lower trace-length than the true population. The scale in e) and f) are unclear and f) misses its panel label.

Fig4: On panel j) IC and CC are switched over.

Fig6: Upon looking at these curves it appears many populations are sampling more than one trend. Using the example of the sub-regional length, the jump in the cumulative frequency curve just prior to 10^2 appears to match the exposure distance of one of the fracture sets in Figure 3c. I would be very interested to see the data by sets that underpins this work, as I suspect some of this variability is due to the sampling of multiple populations.

Fig7: I feel this figure would be best in the supplementary, with an example included as part of a figure in the main text that incorporates figures S2 and S3. Should fig7 stay, then the y-axis in panels a) and c) need labelling and the text in panel c) and potentially b) is too small to read clearly.

Fig8c: I would suggest labelling the aperture and trace length trends in the figure for clarity.

Fig9: This figure presents some very interesting data, and raises a few questions. 1) is the regression co-efficient a combination of both veins and 'other' (I assume joints and faults are grouped here), as there appears to be a large clustering of off-trend features in the other category. While I would agree that veins display a strong trend, I am less convinced by the joints and faults. Particularly when looking at the small aperture results (-5) which range roughly from log length -1.4-0.4 m. I agree that there has to be an upper bound, and that this will increase as length increases, however the lower bound is much harder to suggest. Is it a function of aperture can be larger for larger unfilled fractures, but large unfilled fractures don't need to have large apertures. This makes sense as open voids that are not propped can easily be closed during later tectonic events, and faults are known to be able to form with very thin fault cores. Is it really appropriate to combine both these datasets?

Fig10: a) are the fracture traces that are interpreted as terminating at what looks to be a wedge of cover truly i-nodes, or do they continue underneath? In fig. 12 you mention it is sand and gravel, if this is recent cover then I suggest that the interpretation boundary be extended around this and i-nodes not be interpreted here. The name of the headland just above 'Fig. 11a' is very small and difficult to read. What is the scale for rainbow depth scale? B) A number of potential 'subjective bias' issues are present in this interpretation as outlined in my major comment. Also how penetrative are these fractures? Do they extend to the base of the bedding plane? Brackets are missing after the number of nodes in the key.

Fig11: What is the scale for the node intensity? I-nodes also appear to be present along the edge of the ruler?

Fig12: Very interesting data in the topology plots! Great idea to look at different size circles like this and I would be interested to discuss what you found to be best to provide a) a representative connectivity and b) capture the heterogeneity of the system.

Fig14: Nice schematic, is the bedding sub-horizontal in the area?

Table2: you have split the data between joints and non-filled fractures, how were these differentiated? For your ranges you provide the minimum and maximum, however, this gives no reference as to whether these are outliers. I suggest including either mean (if normally distributed)

or median values, and potentially also a measurement of variability. This way it could be deduced whether you had a 'characteristic' length and aperture within a specific scanline. In the supplementary information it would also be good to see the spread of this data by sets.

References

- Andrews, B.J., Roberts, J.J., Shipton, Z.K., Bigi, S., Tartarello, M.C. and Johnson, G., 2019. How do we see fractures? Quantifying subjective bias in fracture data collection. *Solid Earth*, 10(2), pp.487-516.
- Andrews, B. J., Shipton, Z. K., Lord, R., and McKay, L. (accepted, in revision). The role of pre-existing jointing on damage zone evolution and faulting style of thin competent layers in mechanically stratified sequences: a case study from the Limestone Coal Formation at Spireslack Surface Coal Mine, *Solid Earth Discuss.*, <https://doi.org/10.5194/se-2019-202>.
- Laubach, S.E., Eichhubl, P., Hilgers, C. and Lander, R.H., 2010. Structural diagenesis. *Journal of Structural Geology*, 32(12), pp.1866-1872.
- M Mauldon, M., Dunne, W.M. and Rohrbaugh Jr, M.B., 2001. Circular scanlines and circular windows: new tools for characterizing the geometry of fracture traces. *Journal of Structural Geology*, 23(2-3), pp.247-258.
- Nixon, C.W., Sanderson, D.J. and Bull, J.M., 2012. Analysis of a strike-slip fault network using high resolution multibeam bathymetry, offshore NW Devon UK. *Tectonophysics*, 541, pp.69-80.
- Peacock, D.C., Sanderson, D.J., Bastesen, E., Rotevatn, A. and Storstein, T.H., 2019. Causes of bias and uncertainty in fracture network analysis. *J. Nor. Geol.*
- Scheiber, T., Fredin, O., Viola, G., Jarna, A., Gasser, D. and Łapińska-Viola, R., 2015. Manual extraction of bedrock lineaments from high-resolution LiDAR data: methodological bias and human perception. *GFF*, 137(4), pp.362-372.
- Shipton et al., 2019
- Shipton, Z.K., Roberts, J.J., Comrie, E.L., Kremer, Y., Lunn, R.J. and Caine, J.S., 2019. Fault fictions: systematic biases in the conceptualization of fault-zone architecture. *Geological Society, London, Special Publications*, 496.
- Wilkins, S.J., Gross, M.R., Wacker, M., Eyal, Y. and Engelder, T., 2001. Faulted joints: kinematics, displacement-length scaling relations and criteria for their identification. *Journal of Structural Geology*, 23(2-3), pp.315-327.
- Wilson, C.G., Bond, C.E. and Shipley, T.F., 2019. How can geologic decision making under uncertainty be improved?. *Solid earth*.