Reviewer 1 General comments and suggestions in red

| Comment and suggestion | Reply/Action |
|---|---|
| The main claim of the paper is that it is possible to combine patterns for all types of fractures (opening-mode fractures and faults) imaged from micro-scale to regional scale to find meaningful size scaling patterns. Another claim is that such broad scale scaling observations can be used by making projections to other scales of interest to get input for or to provide information for 'realistic' reservoir models and as input for fluid flow simulations, etc.<br>There is room for improvement in how clearly these claims are stated, how they are related to previous work, and how well they are defended and supported. | We have revised the manuscript extensively taking into account the reviewers comments with the aim of clarifying the claim and how our data support that claim. We are also now much clearer about the limitations of our study. |
| 1(a)The Introduction should have a clearer statement of claims.<br>The paragraph should be broken up to separate the inventory part from a revised and augmented section that explicitly spells out the claims; text that could start out 'Here we show that. . .'. A clear statement of the claims is essential.<br>1(b) These claims also need to match the Conclusions. Neither of these conditions are currently met. The text doesn't make the claims clear.<br>1(c) And the first conclusion (line 640) that the outcropping rocks are a 'direct analogue' is not a conclusion at all. This point was merely asserted in the text without much back up. The comments in line 604-605 seem to point just to a similarity in spacing values. But if this is a major conclusion it needs to be signaled more clearly and the evidence needs to be presented more effectively.<br>It may be easier to just assume that the outcrops may be pretty good analogs and present the evidence for this without making it a major conclusion (but explain what you are doing). | a) We have completely rewritten the introduction to make the claim clear as the reviewer requested.<br><br>b) The conclusions have been rewritten to match the claim<br><br>c) we have removed the first conclusion and followed the reviewer's recommendation. |

| | |
|---|---|
| 2(a) The discussion of previous work is not adequate. The account needs to be more complete and more nuanced.<br>Some of the relevant references are cited, but the scope of this previous work is not clear from the presentation. A more informed and complete account is needed.<br>2(b) Also missing are some of findings from previous work that bear on the main claims of this MS. For example, large aperture size distribution data sets for opening-mode fractures have been collected from a wide range of sandstones. (for example, Hooker et al. 2014) with the aim of predicting the average spacing/intensity of open fractures in reservoirs, and some of these predictions have been tested with horizontal cores or outcrop analogs. This previous work needs to be accounted for more explicitly. And providing a comparison of the results in the current study to the findings of Hooker et al. (2014) seems like an obvious step for putting the current work into context. It should also be addressed in the paper that extensive size-scaling investigations show that some sandstones do not show wide fracture size scaling ranges (see next item).<br><br>The limitations of scaling that have been found need to be acknowledged. Some tests like Hooker et al., 2009 show that, of example, microfracture aperture size distributions can be projected over several orders of magnitude to accurately predict intensity at sizes where fractures can impact production. But other studies, for example Laubach et al. 2016, show that in some sandstones, fractures have a narrow (characteristic) aperture size range and accurate projections from populations of small aperture sizes to large are impossible. This raises a concern that is directly related to the claims of the MS, since these observations imply that some fracture patterns do not scale (they don't have scale invariant properties; they can't be projected to or from larger or smaller sizes). What about circumstances where there is evidence of narrow fracture attribute size ranges? The evidence of the literature seems to be saying that some fractures patterns 'scale' but others do not. Taking a for instance from within the area represented by MS figure 3, Laubach et al. 2014 J. Struct. Geol. showed that two adjacent sandstones, influenced by faults or the same population as described in this MS, have drastically differing fracture attributes (size, spacing, porosity preservation). According to the proposition in this MS, these attributes should be predictable by the MS's regional scaling relation. The test case should be discussed. It's hard to see how the regional could get this right, since these contrasting patterns are on the same. scale. But the differences between the two sandstones are just those that would affect reservoir behavior. The MS claims on this topic need to be reconsidered or at least more completely explored in this light. | a) We have added much of the literature that the reviewer has provided. The introduction is now more complete and nuanced.<br><br>b) we have included the Hooker et al (2014) findings in our introduction and in our discussion.<br><br>We have added discussion of this. We are not saying we have found a general law – that all fractures within sandstones show a wide range of scaling but we are saying that in this instance – the Group 3 structures in the Devonian rocks do show scale invariant properties and this is a good analogue for Clair. |
| 4. A related problem is in the description of studies that examine fractures having a wide range of sizes. The contrast between 'given scales' versus 'multiscale' is problematic, since 'given scale' seems to imply a narrow size range, but some of the studies cited under 'given scale' cover three or four orders of magnitude in scale. Maybe this is just an oversight. The types of structures analyzed and the size ranges analyzed need to be accurately portrayed. Moreover, since the outcrop structures in the reservoir sandstone analog in this MS seem mainly to be opening-mode fractures, the MS should pay closer attention to the previous work on scaling of opening-mode fractures in sandstone. It's surprising that there is no explicit comparison with the compendium of data in Hooker et al. (2014) for example. Or any discussion of the problems with collecting reproducible length data in sandstones outlined by Ortega and Marrett (which is in the reference list). | We have addressed this issue in the new introduction to the paper |

| | |
|---|---|
| 5. Is the distinction in this MS between 'given scales' versus 'multiscale' between data sets where the structures are clearly genetically related and of the same type, versus mixed populations of opening-mode fractures and faults that may not be related? The text I think could be read this way although this isn't stated explicitly. This part of the MS may be the most problematic. As noted in the comments below keyed to lines in the text, it is not always clear what kind of structure is being compared or projected. This needs to be corrected. Partly this problem in the text comes from using the general term 'fracture' to mean either opening-mode fracture or fault. This usage is stated right at the outset. But it leads to problems, confusing and obscuring the argument. The case is being made in the MS seemingly that, for example, patterns of faults visible on seismic can be used to predict the size distributions and connectivity of opening-mode fractures at the reservoir/outcrop scale. This is a very considerable claim (I'm dubious). But the claim should at least be made explicitly and defended openly. | We have explicitly addressed this issue in the revised paper in the following ways 1)We make it clear that the attribute analysis is focused on the Group 3 structures. These include both faults and opening mode strcutures where we observe them in outcrop. At our Dounreay location, faults with metre-scale contained the same mineralisation as opening mode fractures. They clearly contributed to the flow in the subsurface. 2) In the discussion we make it clear that the assumption that the extent to which the scaling of fracture aperture attribute to the regional scale strcutures needs to be tested. |
| 6. The claim that multiscale analysis can be useful for informing geological models has been supported by examples from the literature (these should be noted) but **the claim that, for example, regional lineaments and seismically detected fault trace patterns can be used to predict meaningful fracture attributes at the grid block** or smaller scale seems to me to be a bridge too far. If this is the claim, then a more convincing case is needed to support it. | We have discussed more example of multiscale analysis and we discuss the application of this more carefully in the the discussion section |
| 6b An obvious concern is the projections in figure 8. This figure seems to be saying that aperture and length can be predicted to within two orders of magnitude. What are the error bars on that already really wide prediction? How could such a prediction be used? The authors need to explain how to be useful, 'predictions' can span orders of magnitude (compare the prediction of Hooker et al. 2009 with the two orders of magnitude of size range in the projections of figure 8). Core and outcrop analog data show that fracture patterns at the core and outcrop scale can vary considerably in ways that directly impact fluid flow. As noted above, with adequate samples where microfracture populations are present some of these attributes can be accurately projected over three or four orders of magnitude to predict the attributes of large fractures. But these are cases where the small and | We have not calculated error bars on this prediction – we wish to limit ourselves here to showing the concept of how length scaling constrained over 8 orders of magnitude could be used to estimate apertures. |

| | |
|---|---|
| large fractures are growing and interacting together in a specific rock type. Contemporaneous, interacting fractures are the ones likely to develop power-law size distributions (Cladouhos, Marrett 1996). | In Figure 8 we plot schematically illustrate all the datasets to show the range in intensities from different exposures to illustrate the inherent variabilities. We have modified the discussion of this plot in the revised version and schematically added some low and high strain power law lines to Figure 8c. The C & M 1996 paper is interesting and we have included it. |
| 7. The referencing of certain points needs to be made more complete or more accurate. I've flagged instances in the following detailed notes. As it stands now, I don't think the MS properly represents or credits previous work. | We have added many of the references the reviewer has suggested. |
| 8. There are a number of places in the text where reorganization is needed. The Introduction could be clearer. Some of the material in the Discussion looks more like observations/results. I've flagged some of these issues in the detailed line comments that follow. Improving the overall presentation will increase the impact of the paper. | We have reorganised and rewritten the manuscript extensively in line with the comments form the reviewer. |
| 9. Meaning I've flagged some areas in the text where meaning is unclear. | We have tried to improve the meaning where it has been flagged. |
| My opinion. This broad statement about fault and opening-mode fracture size scaling is true to an extent. Marrett et al. (1999) documented power-law scaling across 3 to almost 5 orders of magnitude regardless of rock type or movement mode. This was the study that established that such systematic relations exist and that extrapolation from one scale to some other scale of interest was a feasible approach. It's a surprising omission to leave this paper out. One thing that Marrett et al. did not do was to mix opening-mode fracture and fault data sets. Doing so requires some defending. It's ok to make the general point that some faults and some opening-mode fracture populations show scaling patterns (although subsequent work shows that some populations do not scale in this way). But it is problematic to lump them all together as 'fractures' if in your description and discussion you let the reader lose track of which kind of structure you are talking about. You are making the claim that it doesn't matter which type of 'fracture' is analysed - that's fine if you can defend it but it's not convincing if you just use the all-purpose word 'fracture' in a way that makes it hard for the reader to assess the | We thank the reviewer for reminding us about the Marrett et al 1999 study which we now include.<br><br>See above reply to point No. 5 |

| | |
|---|---|
| strength of your claim. For example, in lines (602-612) it's hard to tell which type of structure you mean. Marrett, R., Ortega, O. J., & Kelsey, C. M. (1999). Extent of power-law scaling for natural fractures in rock. Geology, 27(9), 799-802. | |

Reviewer 1 – Line by line comments (numbers refer to original manuscript)

| |
|---|
| 34-37 We have rewritten the introduction taking into account this feedback and providing a greatly expanded number of references |
| 37. We have added the de Dreuzy reference and included this general point |
| 38. We have revised this section |
| We have clarified this in the rewrite |
| 42. We have clarified what we mean here in the rewrite |
| 43. This is a good point that we omitted and now include specifically and cite the Laubach et al 2019 paper. |
| 44. Our study was completed before the JSG Special issue was published but we have now added the result a spatial correlation analysis to the manuscript as plots in the supplementary file and discuss the results. |
| 48. We have rewritten this section and removed any impression that opening mode fractures can be detected on seismic. Interestingly, elsewhere we have published work on opening mode fissure structures in basement rocks that potentially are expressed in seismic attribute maps of basement highs (the Lancaster field). We do not mention it here as the lithology is different. |
| 50. We agree with this point and in our re-writing of the introduction have taken these comments into account. |
| 52. This has been corrected |
| 54-63. This has all been rewritten |
| 56. We have included the Ukar et al study in a new section 2.2 which discusses the validity of the Orcadian basin analogue for Clair |
| 57. We have revised the introduction so this point np longer applies. |
| 62. We have included these studies in our introduction |
| 67. We have added a new section 2.2 which discusses this aspect directly |
| This is a valid point that we acknowledge in our discussion |
| 75. We have clarified this |
| 76. We have clarified that we do mean aperture and not fault width for these structures. |
| 77. We have rephrased this sentence |
| 80. We have edited this sentence |
| 89. We have added the reference to Marrett et al 2018 and changed 'spacing' to 'spatial arrangement' |
| 103. No. we used a feeler gauge which is used in engineering to measure widths as small as 0.02 mm. We used this tool in conjunction with a hand-lens in the field or on a hand-sample. We have modified the text by adding 'an engineering feeler gauge in conjunction with a hand lens' to make this clearer. |
| 113. We accept this point and have made it clear in this paragraph that spacing attribute is more likely to show log-normal distribution whereas size attributes are more likely to show power-law scaling. |
| 125. We have removed 'although unknown' from this sentence to avoid making this inference |
| 127-132. We have edited this section carefully to make the point suggested here about limits to power-law behaviour.<br>We have also made it clear in the following section that power-law scaling should not be assumed |
| 132. This was an error – we have corrected this to 'over several orders of magnitude length scale'. We moved this sentence as part of revisions in response to the previous comments |
| 151. We recorded fractures that are related to the Group 3 set – this is made clear in the 'Geological Setting' section |
| 156. We have changed this to 'box-counting dimension' – which is the plot that we are referring to here. |
| 157. Correct – we have added 'of connection types' for clarity |
| 165-175. We agree and have added 'fracture trace' before connectivity to make this clear. |

| |
|---|
| 188. We have added a a new section (2.2) justifying the use of the Orcadian basin as an analogue for the Clair reservoirs and included the Ukar et al criteria. |
| 210. The point here is that these fracture fills show the structures formed in the subsurface and are not related to exhumation of the analogue. We have clarified this point in the text. |
| 233. Yes and we have moved text that described the evidence for this in the discussion to this part as we realize now it is important primary evidence |
| 256-261. We added an estimate of the difference this makes to the number of fractures we were able to record as follows. |
| 279. We modified the text to say ….'they mutually cross-cut each other which enabled us to infer that they were active during the same geological event' to make it clear what our evidence was. |
| 291. We added text to explain this and acknowledge the limitation. 'At the scale of a thin section, only samples from fault zones contain enough fractures to produce a statistically significant sample. We thus recognise that the results at this scale are representative of fracture intensities within fault zones and provide an upper limit relative to background. At the scale of a thin section, only samples from fault zones contain enough fractures to produce a statistically significant sample' |
| 285-299. See above comment |
| 314. We have analysed the spatial arrangement using the Marrett et al (2018) method and have added a short discussion of the results and the plots to the supplementary file. |
| 350. We have added a citation to this paper |
| 361. Corrected by removing 'coefficient' |
| 368. Edited this paragraph for clarity. |
| 370 Corrected – removed this sentence in rewrite |
| 380-389. This is now discussed in the last paragraph in section 8.1 |
| 394. We have added the evidence that the features in the bathymetry are the same as those in the adjacent coastline. We also added some context for the Brims Ness photo location as well. |
| 412. We have added a citation to the Questiaux et al references and now report the clustering recorded by the spatial arrangement analysis. |
| 459-463. We have added a sentence in this section which explains the significance of the box counting dimension and referenced Hirata (1989) |
| 465-467. We agree the box counting doesn't tell you anything about the spatial arrangement and as noted above we have performed the spatial correlation analysis and reported the results. |
| 467. As Walsh & Watterson (1993) discussed 'A fracture pattern incorporates many different attributes such as orientation distribution, size population and fracture trace geometry'. There is no simple relationship between the box counting dimension and the size distributions |
| 471-500. We have taken the reviewer's advice an broken this section into 3 paragraphs and state the supporting evidence for the assertions more clearly. |
| 475. We discuss truncation and censoring in Section 3.1.1 and now make it clear again here what the effect can be and that our MLE methods can help to reduce the uncertainty in fitting power-laws to datasets that have a somewhat limited scale range. |
| 481. We have clarified that the intensity variations and slope variability could be due to samples taken in slightly different contexts (e.g. inside or outside a damage zone). |
| 481. We have not been able to access the Hooker et al (2014) data as these are not available publicly. We have added in the discussion that our results are in agreement with Hooker et al (2014) |
| 483. We have clarified what we mean by 'reducing the influence of individual datasets' by rewriting this section. We have justified plotting the data from different scales because we are confident they ARE genetically related as we have explained in section 3. |
| 489. We have corrected the reference to Gillespie et al 1999. We have reported in section 5.1, the results from the Spatial correlation analysis recommended by the reviewer. (we note that most of this work was done prior to the publication of this volume). |

| |
|---|
| 496. We have clarified what we mean here and changes made in preceding paragraph discussed our result in context of Hooker et al (2014). |
| 502. We have given citations to a number of previous studies in a sentence added to this section. |
| 512. We have now cited Olson (2003) in this section and clarified how our aperture data fit with his model. |
| 515. We have added discussion of our 0.65 exponent in terms Olson (2003) and Schultz et al. (2008) as well as a recent study by Mayrhofer et al. Schultz et al. and added citations |
| 520. We have moved this section to section 2 – Geological setting to provide evidence to support the Group 3 fractures being similar to Clair structures |
| 522. We have changed this to make it clear that it is the fractures that contain the fault rocks. We have deleted the 'hydrothermal' form this section as its somewhat irrelevant to point that we are making here – that the fractures and their fills show similar features to Clair. |
| 523. We refer to Dichiarante et al 2016 where these structures are described |
| 534. We refer to Dichairante et al (2016) where evidence for this was discussed. |
| 548. We agree – but have not changed our text as this is covered by 'dramatically reduce' flow |
| 50? We have explained more clearly at the start of this section what we are attempting to do. |
| 551. corrected |
| 553. We have clarified the Coney et al study is based on wells and aeromagnetic surveys. Coney described these as faults sets so we have changed the terminology to this.Yes they are spaced at these intervals |
| 560-561. Yes – if we can use spacing/intensity to predict length then we can use it also to predict aperture given the aperture/length scaling relationship we presented and discussed in an earlier section. We have added discussion of this assumption in |
| 562. We have clarified what we mean here |
| 562. We now discuss this assumption explicitly. In this application, we use the length attribute at regional scale and consider what the 'aperture' would be given the scaling relationship. It is using the assumption that we know from onshore evidence that faults are contributing to the overall subs-surface fluid flow. |
| 567. corrected |
| 568. We have clarified that the structure was found in the core. |
| 563-8. We have now made a new paragraph that explicitly discusses the assumptions and limitations of our approach |
| 574. We have limited the use of fracture corridors to the regional scale – which is what has been suggested by previous authors. We now refer to the mesoscale examples as clusters of interconnected fractures as the reviewer suggested. We have removed the sentence about 'minimum' estimate as this suggestion is not well constrained. |
| 584. We have added into this discussion the insights from the spatial correlation analysis. |
| 582. We have limited the use of fracture corridors to the regional scale where we do mean 'abnormally closely spaced subparallel fractures'. We have define our usage of fracture corridors. |
| 584-587. We have revised this section to make it clear what we mean and restricted our comments to fracture connectivity in 2D. |
| 589-595. We have now report the spatial correlation results and 2D spacing estimate in the results section - see sections 5.1 and 7.1.1 |
| 602. We have reorganised the discussion so that the flow is more logical. |
| 606. Yes |
| 602-612. We have rewritten this section to improve clarity. |
| 608. No these data re not available |
| 607-608. Rewritten to make it clearer |
| 614-625. We hope the rewritten section is now clearer. |
| 616. We have revised this paragraph and included the Philip et al reference – we are grateful to the reviewer for bring it to our attention. |

| |
|---|
| 614-615. We have modified the sentence slightly to change the emphasis away from being economically 'useful' to providing a useful insight into subsurface fracture properties. |
| 619. We have clarified our assumptions and limitations with respect to this in the discussion |
| 628. Added the Primaleon paper which is now published |
| 634-638. Correct. We have removed this text. |
| 640-641. We have toned down the assertion and now describe the results from Orcadian basin then mention the application to Clair |
| 652. Corrected to 'fracture' |
| 673. Drilling strategy is mentioned in the discussion – its more clearly flagged now.<br>Despite the uncertainty the application is still useful. |
| 872. No as Walsh and Watterson pointed out the artefacts arise when large areas of no exposure are included in a fracture pattern. This is not the case in this example |
| 876-878. We have deleted b and c from Fig 1. |
| 876. These are example distributions. We have clarified this in the caption |
| 898. We added a reference to the locations in Figure 3. All observations are from 'target' sandstones. |
| 902. We have removed 'present-day' and replaced with an imposed stress. |
| 909. Replaced with 'kinematic aperture' |
| Figures. As noted above we are not able to compare our results with Hooker et al 2014 as their data are not available. We have made qualitative comparison in the discussion |
| Fig. 2. We have given the alternative classification in the in the figure caption but retain this classification as we find it useful. |
| Fig. 3 We have fixed the labelling and improved the description of the figure |
| Fog. 6. See comment above with regard to Hooker et al. |
| Fig. 8. We have done this in the text. |
| Fig. 9. We have corrected this to 2 decimal places |