Dear Editor,

Please find my general and specific comments below for my review of the revised manuscript by Kirilova et al., titled “Micro- and nano-porosity of the active Alpine Fault zone, New Zealand”.

General comments

In their revised manuscript the authors have chosen to retain the manuscript largely as is. The few additions that have been made in the methods section now help the reader not make the mistake I made when first reading the initial submission.

I understand why the authors have been reluctant to change the manuscript as it is well written and has a flow that guides the reader but I do not find the answers to my last comments satisfactory in addressing the fundamental limitations of the data. The chief concern I have is that the porosity segmentation is too simplistic to allow the interpretation that follows in the discussion of a porosity and permeability gradient. In the following specific comments I have tried to clarify why I think that the authors may be over-interpreting their data and as such why I think that the authors should make room for more discussion about the limitations of the data.

To be very clear I think that the work is good and the results can be published but first there is a need for more transparency in the methodological workflow used and how it may affect any interpretations. If the authors wish to make the claims they currently make then I think these must be made more cautiously and with enough information given to allow the reader to evaluate the discussion points. As the manuscript currently stands, the results presented do not allow a gradient in porosity and permeability to be interpreted nor the further interpretation of this that variations in dissolution-precipitation must exist, which are both key discussion points of the current manuscript.

Best,

James Gilgannon

Specific comments

I have kept to the sections previously defined and have written my comments to the authors responses under headers of the same names. My original comments are in the lightest grey, the author's responses are in italics and an intermediate grey colour, while my reply to the author's response is in black coloured font.

Lines 123 - 130: Pore shape descriptors

The manuscript would significantly benefit from a figure illustrating the relevant aspects of the use of the covariance matrix. For example, I do not understand the author's characterisation of sphericity. I may have misunderstood the description but the ratio of two eigenvalues, which are both contained within a plane, surely cannot describe the deviation/tendency to a sphere, or have I misunderstood the metric you present? I am more familiar with sphericity being the ratio of the equivalent surface area of a sphere with the same volume as the pore volume over the actual surface area of the pore volume (e.g. Wadell, H. (1932))?

For this reason, I think that the section would benefit enormously from an example figure that corresponds to, and visualises the explanation of the metric. I imagine this would be best done with some specific examples of pore volumes from your data set. If the authors have not come up with the method themselves then I think that a citation for the more curious reader is also necessary.

Wadell, H., 1932. Volume, shape, and roundness of rock particles. The Journal of Geology, 40(5), pp.443-451.

Response: Thank you for this comment. No, we have not come up with the methods ourselves. All the shape analyses we performed are based on functions embedded in Avizo software that yield volumetric and shape characteristics for each segmented material in numerical format (lines 118). We simply plotted the results on bivariate histograms by using Matlab as stated in line 125.

We do not find it necessary to include in the manuscript a description of how the software produces those results as Avizo software is trusted source and every user/reader can refer to their library. However, here we provide a brief explanation of the functions we have used:

The covariance matrix is built on the basis of the moments of inertia and can be written as: …

(matrix provided and implemented by Avizo software)

By using this matrix, the software computes the three eigenvalues by using a Singular Value decomposition. In an elongated ellipsoid the largest eigenvalue will describe the longest axis of the 3d object.

In this context, the deviation of the spherical form (i.e. anisotropy - a value extracted from Avizo software) is measured as 1 minus the ration of the smallest to largest eigenvalue. In a 3D object if the smallest and the longest axis are equal, the medium will have the same value as well, describing a spherical object and having numerical value = 0.

Reply to Response:

I agree that the authors should not repeat the Avizo user manual in their manuscript but I was not suggesting that this was necessary. The authors already do a great job of explaining in writing the ratios they implement and as such if the authors do not wish to provide a figure for explanation then this is a choice I understand.

However, I do not think that the term sphericity should be used as this is already a term expansively used in the literature. As the authors wish to defer to the Avizo user manual then the metric should be changed to the term that it is defined as in Avizo, which the authors refer to above as anisotropy. I say this because the Avizo user manual (at least in the copy for v. 9) has a separate section on custom functions where the user is shown how to implement the well known sphericity calculation I cited in my last review. Calling a metric by a name it is not known as, and which happens to be the name of an already established metric, is confusing.

2. Questions/concern regarding total porosity calculations
My questions/concern is regarding the uncertainty associated with the filtering of pore data used and how this translates into the discussed differences in the magnitudes of the total porosity from different samples. Your TEM results show that very small fractures exist, which you identify as fracture porosity and, by the general argument of the paper, could have resulted from coring. While these fractures shown in fig. 8d are below the XCT resolution, I am brought to wonder how many slightly larger pores exist that are actually induced fractures. For example, the fact that so many small pores identified by XCT are almost completely flat in shape (fig. 6) might reflect that many small fractures, that are not syn-kinematic, are retained in the analysis. Therefore for me a question that presents itself is; does a simple size threshold, as you have used, have an appropriate amount of filtering information to allow a discussion about a difference of 0.14% porosity? Stated more plainly, how do you know if the variation between samples is not just a function of the degree to which each lithology experienced the coring and retrieval? Alternatively, can you rule out that the variation of 0.14% may just be related to the uncertainty of the polynomial fitting used to cap the pore size for integration?

I am uncertain if it is correct to straightway interpret this difference of 0.14% as meaningful. I think that more interrogation of this result needs to appear in the discussion before it is taken forward as independent confirmation of other literature. It might be that the authors wish to use the bore hole and laboratory measurements of permeability that are mentioned in the text to quantitatively check if the difference of 0.14% in total porosity can account for these differences in permeability.

I am aware that this would require some assumptions when calculating but it would provide a base to the interpretation that a difference of 0.14% porosity between samples is meaningful. As currently presented I think that the result only convincingly shows that each calculated porosity is of the same order of magnitude.

Response: We understand the reviewer's concern that big pores and small fractures could get easily misinterpreted/mislabeled in XCT datasets. This exactly is the prime reason why we decided against calculating total porosities in these samples by simply using 'connected components' and instead we fitted the data to a polynomial curve (mentioned in lines 115-123). We believe that implementing a mathematical approach is much more trustworthy than limiting the data based on the interpreter's bias. Furthermore, our total porosity calculations (by using the polynomial fit) roughly coincide with the total porosities yield by calculating the total porosities based on connected components with up to 200 voxels. You can see these numbers on the table below:

The reviewer also expressed concern about the fact that some of the very flat pores may represent fractures. We acknowledge the validity of this statement. However, we believe that our approach of excluding cracks is efficient and possibly the best methodology for analyzing these samples (i.e. fitting the data to a polynomial curve). Furthermore, the shape of these pores is also very likely to result from their distribution along grain boundaries, especially of clay minerals (lines 190). The authors of the manuscript are in favour of this second scenario.

And last but not least, the difference of 0.14% of total porosity in these samples may seem insignificant to the reader. However, all of the samples contain extremely low porosities, and thus only 0.14% more pores actually result in doubling the amount of pores in sample DFDP-1B 69_2.57 in comparison to the rest of the samples. Thus, a discussion here is not only meaningful but also required, and very well related to changes in lithology in between the samples (lines 219-226, now in lines 222-229), and previous permeability measurements of these rocks (lines 233-236, now in lines 236-240).

Reply to Response:
Thank you for your answer and considering my concerns. From your reply I take away that you feel the use of the polynomial fit leads to a more trustworthy total porosity calculation. I am not sure that I agree that there is much difference between using the 'connected components' method vs the polynomial fit of pore data.

The first reason I do not agree is that:

1. They are both just pore ‘size’ filters. The connected components caps the total porosity to the pores below 200 voxels and the integration of the polynomial simply caps the total porosity to pores below those deemed as the largest by the x intercept.

The fact both methods produce similar total porosities (cf. table in your reply to my comments) probably just reflects that the size filtering is similar and there are not so many pores between the 200 voxels cap (~430 μm^3) and the higher polynomial caps (~580,~680, ~860, ~760 μm^3).

The second reason, which I think is more problematic, is:

2. A simple grey scale threshold has enormous uncertainty associated with it and this uncertainty can account for variations on the order of those that are reported (i.e. doubling/halving).

This choice of segmentation method would affect the outcome of any later data filtering, be that using the connected components calculation or the polynomial fit.

For the purposes of illustration I have constructed a fake porosity microstructure (fig. 1) to quantitatively highlight that the choice of threshold can give a range of values, in some cases the variation can be as much as double. This fake porosity microstructure has been constructed to have ~2% porosity (fig. 2a). When one chooses three different thresholds to segment the pores by (fig. 2b and c) one gets three different values in total porosity (fig. 3). One can see in figure 2c that all of the chosen thresholds are reasonable for the fake pores and their resolution, but the difference between those choices is as much as double the total porosity (1.1% vs 2.1%).

In your methods you have not reported the thresholds chosen or how data between samples looks (for example, are different core’s XCT data of similar quality or not?, and how do their microstructures compare in XCT?, how do the greyscale histograms, and hence thresholds, compare between samples?). From the 2D sectioned images you provided in your reply to my comments I can see that the small pores are made of not very many pixels in a slice, which makes them susceptible to the sensitivity I have highlighted in my quantification of the fake microstructure. Therefore it is very hard to believe in the significance of a 0.14% difference in your samples.

In my original review I suggested that if you wanted to pursue the claim that the 0.14% difference is significant you would need to either bring compelling microstructural evidence to support such a claim or qualify if the porosity variation you report can account for the permeability gradient reported by other studies that you cite. As neither of these, nor any other supporting line of argument, has been included I think that you cannot consider this difference beyond the uncertainty of the method used. The consequence of this is that as the results stand it is erroneous to discuss the observed differences as independent verification of a permeability gradient reported by other studies or to infer that there is more or less dissolution-precipitation processes active in one sample over the other.

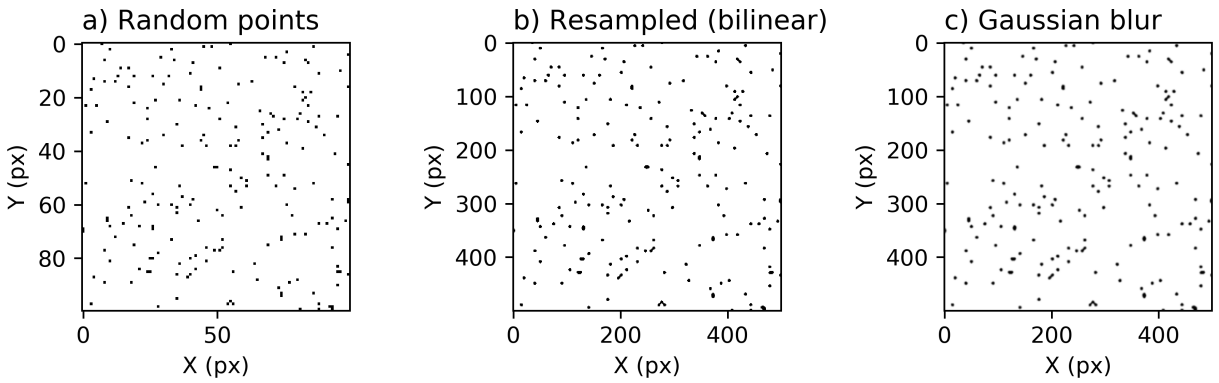

Figure 1: Generating a fake porosity microstructure. (a) ~200 pixels that represent pores were randomly generated in a 100x100 matrix. This represents a total porosity (ϕ) of ~0.02. The simple single pixel pores were then resampled to increase their resolution (b) and then a greyscale gradient along pore edges was introduced by applying a Gaussian blur to the image (c).

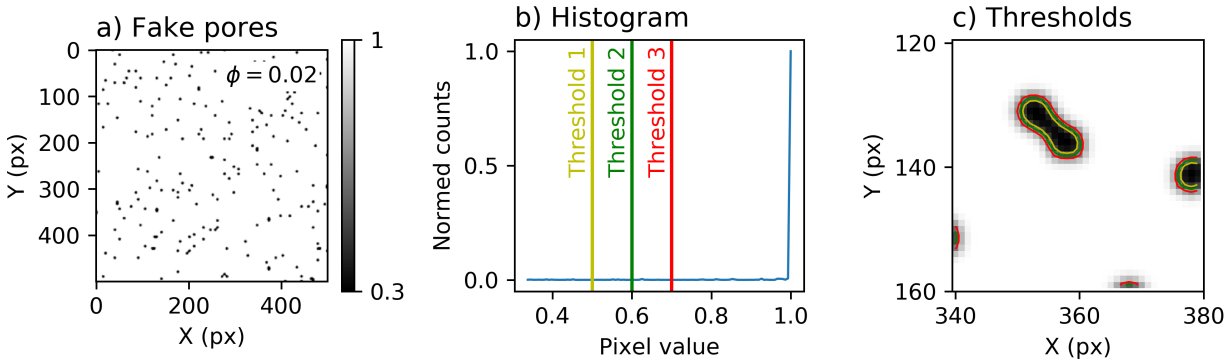

Figure 2: Segmenting the fake porosity microstructure with three different threshold values. (a) is the same microstructure as shown in fig. 1c. The histogram of pixel values in fig. 2a is presented in (b). Alongside this three different thresholds to segment porosity are shown in fig 2b. These thresholds are visualised for example pores from fig. 2a in (c).

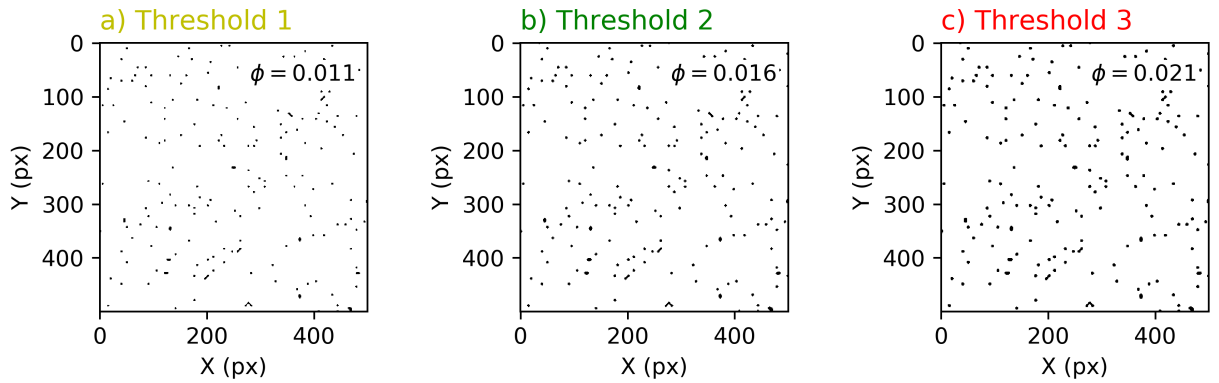

Figure 3: The total porosity calculated with three different threshold values.