Authors’ response:

Dear Dr. Jubb,

Thank you for taking the time to study our manuscript and for writing such a helpful and comprehensive review. We have taken your comments very seriously and have, as a result, made significant changes to the manuscript. Our specific responses are detailed below each major and minor comment individually (blue text), but we thought it would be useful to summarise our changes as follows:

- We have focussed the paper on the changes to individual Raman spectral parameters rather than geothermometric equations, only including the geothermometry as a minor discussion point later on in the text. We believe that this makes the manuscript far clearer and more focussed, not to mention easier for the reader to surmise the key points of the study.
- Transects across faults and shear zones are presented in more detail.
- Error ranges have been added to both text and figures.
- Methodology and definitions have been significantly tightened.

Once again, we would like to thank you for your time and for your help in improving our manuscript. We believe that with your comments we have been able to make substantial improvements in terms clarity, and we are particularly grateful for your comment summarising the key questions answered by our paper as this had the effect of highlighting those points that we had made clear and those we had not.

Yours sincerely,

L. Kedar, C. E. Bond, and D. Muirhead.

Major Revisions

1. Geothermometers were not applied correctly. Specific problems with each detailed below.

Thank you for highlighting your concerns with our use of geothermometers in this manuscript. Upon consideration, we determined that the original manuscript lacked clarity and tried to ‘do too much’ – we believe that it is more useful to the reader and the scientific community to focus the revised manuscript on the response of individual Raman parameters only, rather than including various geothermometers too. The geothermometry side of the original manuscript could be addressed in a future follow-on paper, which would of course address each of your concerns below.

For each comment below, we have described the changes made, and where appropriate, have explained our reasoning. However, for the most part, the points relating to geothermometric techniques no longer apply to the revised manuscript as we have removed the geothermometry section in order to focus the paper on the individual spectral parameters and their response to strain. This means that the Lahfid and the Kouketsu thermometers are no longer included in this paper; however, the Schito-

Corrado equation is discussed as an implication of our findings relating to individual spectral parameters.

a. Lahfid thermometer: T = (RA1-0.3758)/0.0008, here RA1=(D1+D4)/(D1+D2+D3+D4+G) using peak areas from Lorentzian peaks fit to Raman spectra collected with 514 nm laser. To apply this thermometer correctly you need to:

(i) Collect the Raman data with a 514 nm laser because dispersion of the D peak will change the calculated area. See the following references for discussions on dispersion effects on Raman spectra of geologic organic matter:

Lünsdorf, 2016, Raman spectroscopy of dispersed vitrinite – Methodical aspects and correlation with reflectance, Inter. J. Coal Geol., 153, 75-86.

Sauerer et al., 2017, Fast and accurate shale maturity determination by Raman spectroscopy measurement with minimal sample preparation, Inter. J. Coal Geol., 173, 150-157

Jubb et al., 2018, High microscale variability in Raman thermal maturity estimates from shale organic matter, Inter. J. Coal Geol., 199, 1-9

(ii) Fit the collected Raman data using a sum of Lorentzian profiles as Lahfid did, not the pseudo-Voigt profile that was used here.

Thank you for pointing out this oversight. In future we will bear this in mind; however, we now no longer use the Lahfid equation.

(iii) Fit the collected Raman spectra with five peaks, not two.

Thank you for highlighting these points. A 514nm was used; this has now been specified (Line 240). The original reason for using 2 peaks as opposed to 5 was that some work (e.g. Schito et al., 2017; Schito and Corrado, 2018; Henry et al., 2019) suggests that the two methods produce comparable trends, and so although not absolutely accurate, we reasoned that it would be possible that some authors may try to equate the two methods, and therefore it was reasonable to investigate the results that might be produced using the two-peak approach. We appreciate that this was not made clear in the original manuscript, and any future work will address this problem. The revised manuscript does not include the Lahfid geothermometer any longer as we focus on individual parametric responses.

b. Kouketsu thermometer: T = -2.15(D1-FWHM) + 478, collected with 532nm laser, fit using pseudo-Voigt profiles.

The number of peaks fit to the Raman spectra for the Kouketsu thermometer was dependent on several qualitative and quantitative parameters. See flowchart (Figure 3) from Kouketsu 2014 here:

For the sample temperatures in this study (left end of the flow chart), either 4 or 5 peaks needed to be fit to the Raman data in order to correctly apply this geothermometer.

Once again, thank you for these comments. As with the Lahfid thermometer, the Kouketsu thermometer is not included in the revised manuscript owing to our general refocussing of the paper. Again, the idea was that general use of such geothermometers might result in cross-overs of methods such as this. However, future work will ensure that correct methodology is followed and that any deviations from this are accompanied by valid reasoning.

c. Schito and Corrado estimation of VRo% for use in the Barker and Pawlewiscz geothermometer: T = (ln(VRo%) + 1.68) / 0.0124.

Here the authors have correctly used relationships from Schito and Corrado to determine an equivalent vitrinite reflectance (VRo%) from the Raman spectra. This parameter is then used to determine a temperature using the 1986 thermometer proposed by Barker and Pawelewicz. However, the data that this thermometer is based on are highly uncertain (see Figure 1 from Barker and Pawelewicz 1986 below), making it semi-quantitative at best. Certainly, correlating vitrinite reflectance to an absolute temperature is a large challenge for the field and one that warrants careful study in its own right. Regardless, for the work reported here, an estimation of uncertainty in the determined temperatures using the Barker and Pawelewicz equation **must** be included if this geothermometer is to be used.

The revised manuscript addresses these points by highlighting the uncertainties described above (Line 488). The Schito and Corrado equation is now primarily used as a discussion point rather than a focus to the paper.

2. Analytical uncertainties need to be added to Figures 5, 6, & 7 (where appropriate) and better discussed throughout the text. Specifically, what do uncertainties represent (e.g., propagated precision from fits, standard deviations, reported uncertainties from original geothermometers?), what confidence interval do these uncertainties represent, and what the uncertainties mean for the interpretations drawn from the data.

We have now included uncertainties in the text and corresponding error bars in the figures, and have also expressed what the uncertainties were derived from (e.g. Line 323). These are mostly based on standard deviations for the individual parameters, but for the discussion relating to the Schito and Corrado equation and the resulting conversion through Barker and Pawlewiscz, uncertainties are derived from the original equations.

3. More detail is needed on the organic matter comprising the samples and on the Raman analyses. Specifically:

a. What types of organic matter made up the extracted kerogen? Different organic matter types will have different Raman response due to differences in their molecular structures.

Organic matter types were not analysed, but we have highlighted the fact that changes in lithology may influence the resultant spectra. However, we are more interested in trends in the data and anomalies observed within the transects (which are now presented in more detail: Figure 6).

b. How does the kerogen isolation procedure change the Raman response? If this wasn't' tested, some text describing why this procedure isn't anticipated to impact the Raman spectra needs to be included.

This procedure has been performed on many examples where the carbon content may be low and shows no evidence of spectral shift:

J.S. Schmidt, R. Hinrichs, C.V. Araujo, 2017: Maturity estimation of phytoclasts in strew mounts by micro-Raman spectroscopy. International Journal of Coal Geology, Volume 173.

Muirhead, D. K., Parnell, J., Spinks, S. and Bowden, S. A., 2017: Characterization of organic matter in the Torridonian using Raman spectroscopy, Geol. Soc. London, Spec. Publ., 448(1), 71–80)

c. What mesh size were the samples powdered to prior to demineralization?

Samples were not powdered to a specific mesh size (though most were smooth to the touch) but were instead crushed for a length of time that was consistent between samples to minimise the chance of different durations potentially shearing the samples to varying extents, although the effect of this was thought to be minimal. See refs ready above

d. What grade of HCl and at what temperature was used in the demineralization procedure?

Room temperature and 10% HCl were used; these detailed are now included in the manuscript (Lines 236 and 239)

e. Were there entrained clay particles in the kerogen isolates that remained after demineralization? If so, could that impact your results?

During the demineralisation process, clay that was contained within the bulk sample tended to float to the surface of the solution, particularly during neutralisation/rinsing. Once dried, this clay could be scraped off the surface of the residue.

f. What evidence do you have that graphitization has not occurred in any of your samples?

We predict that due to burial depths of 6 to 8km it is extremely unlikely that graphitisation has occurred due to burial conditions. There were no intrusions to elevate temperatures and samples were taken away from the influence of hydrothermal fluids (with the exception of fault surfaces which could of course be conduits for fluids). Finally, none of the observed spectra were graphitic in nature.

g. How did you assess the presence/absence of thermal alteration from the excitation laser during Raman collection?

All acquisitions were the same length and duration, with laser power kept low. Samples were visually checked for signs of burning after each spectrum was recorded.

h. What microscope objective was used? What was the numerical aperture? What was the laser polarization?

The microscope objective was 50x magnification. Numerical aperture was 0.75 and laser was polarized from source.

i. What function was used to smooth the data? How did smoothing the data impact the Raman parameters from the fits?

Smoothing the data aided the visual identification of spectral peaks. Where smoothed spectra still appeared to have significant noise, this was taken as an indication that errors due to the software fitting curves might be increased.

4. Figure 4 shows ‘typical’ changes in Raman parameters vs. temperature and strain. However, no discussion (or very little) is given to how these ‘typical’ trends were determined, especially for temperature. A better representation of this kind of analysis can be seen in Henry et al., 2019, Earth-Sci. Reviews. I recommend removing this figure (at least with regards to temperature) and directing readers to the trends shown Henry et al.

Figure 4 is intended to be a visual summary of general trends in Raman spectral parameters and is not to scale. Trends are based on data summarised in Henry et al. (2019) but are simplified further and are not tied to specific numerical values. The figure is simply to visually represent the written description in Section 4.3.

5. The discussion of the Raman band separation (RBS) parameter is unclear throughout and is incorect in several places. Specifically:

a. Section 5.2: “…RBS appears to show no consistent pattern throughout the stratigraphy…”. To me, the RBS parameter shows a consistent, if weak, trend decreasing with depth.

We originally described this lack of trend because the variation between adjacent samples was often of the same or similar magnitude to the overall change through the stratigraphic section, so any apparent weak trend could feasibly be a result of this variation. However, we accept that it is still worth mentioning that there does appear to be a weak trend, as you point out. Therefore, we have included a note on this in Lines 348-349: “Raman band separation (RBS) varies through the stratigraphy (Fig. 5b), with what appears to be a prevalence of values >265 in the upper stratigraphy and <265 in the lower sequence.”

b. Section 5.2: +/- 4 cm-1 for the RBS parameter is referred to as “high error”. What does this error represent (see comment #3) and furthermore, this degree of uncertainty is <+/-2% the determined value. Why do the author’s think this level of error is high?

The error is considered to be high because the variation between adjacent samples is of similar magnitude to the overall change observed through the whole sequence (see above comment). We have clarified this in Line 349: “The average error associated with each sample is +/- 4, suggesting that the change through the stratigraphic sequence is not significant.”

c. Section 7.2: The statement “It is possible that the samples in this study are not of high enough maturity for a trend in RBS to be seen….” is not correct. Relatively low temperature organic matter show trends in RBS. In fact, both of the citations provided to support this statement show RBS data with trends for low temperature source rocks. I do not understand where this statement comes from.

This statement originally stemmed from the high degree of variation in low-temperature RBS results presented in the studies cited; however, we accept that this statement is perhaps misleading. We have therefore adapted the manuscript to fit this: Lines 432-434 now read “Raman band separation (RBS) is reported to increase with increasing temperature (Fig. 4b; Zhou et al., 2014; Bonoldi et al., 2016; Sauerer et al., 2017), and so should increase with depth towards the basal thrust in our study.”

d. Section 7.2: The statement “This suggests that frictional heating does not play a significant role in changing the Raman spectral parameters on thrust planes.” is not supported by the data. RBS is calculated from D- and G-peak frequency. FWHM and peak height are other Raman parameters not included in the RBS parameter.

We suspect that our wording here was unclear and we thank you for highlighting this. In our discussion (Lines 434-440) we now include the following points: “If frictional heating on fault planes were the primary control on changes in RBS, and we assume an approximate instantaneous slip magnitude of ~1 m, then it would be expected that temperatures could rise by >400°C (Savage et al., 2014). This should be enough to produce a shift in RBS which is greater than the general variation we see in our samples. However, Nakamura et al. (2019) report that in addition to temperature, RBS is sensitive to lithology and the effects of fluids, which may explain the variable results we see in this study.”

6. Line 372: “….we only observe a decrease in FWHM[d], whilst FWHM[g] changes very little…”. This suggests that something weird is going on with either the samples or the analysis as this observation is fairly unexpected. More discussion is needed to explore this observation as FWHM[g] is usually considered one of the “best” indicators of maturity in Raman geothermometry.

We are aware of this discrepancy but thank you for pointing it out! It is beyond the scope of this study to investigate this as we are merely aiming to present the data observed in relation to how the parameters change across locally strained rocks; however, it is worth further investigation, as you correctly suggest.

7. Section 7.8: The statements “From our observations…..equation is less effected by strained environments than…” and “The Kouketsu equation is also more suited to strained environments.” illustrate a real disconnect in the author’s perspective on Raman geothermometers and what is *actually* being measured when you probe organic matter with a Raman instrument. The Raman response from organic matter is dictated by the molecular structure of the organic matter ensemble in the probe volume of the Raman microscope. Raman geothermometers work because the thermal alteration of organic matter structure as it reacts toward a graphite endmember is deemed irreversible. The interesting question this study is trying to ask is “Does strain, independent of temperature, change organic matter structure such that these effects need to be accounted for when determining a temperature from Raman spectra of geologic organic matter?”. That is, all Raman geothermometers should be affected by strained environments if strain is changing the molecular structure of the probed organic matter. Stating that one geothermometer is more appropriate than another for strained environs misses this point. I strongly suggest reworking the discussion and conclusion with perspective on this point.

We are particularly grateful for this comment as it highlights the key points which we have made clear and those which we have not. Your (correct) interpretation of the key question that this paper is trying to answer was useful for helping us to refine the focus of the manuscript. We also appreciate that all geothermometers will be affected by strain if our results are correct; however, the point we tried to make was that since certain spectral parameters are more affected by strain than others, a geothermometric equation which utilises those parameters which are *less* affected by strain (e.g. FWHM[d]) would logically be *more* applicable in strained environments, although of course this only further highlights the simplicity assumed by many existed Raman geothermometers!

We now only include the Schito and Corrado equation in our manuscript and only do so as a discussion point. We mention the sensitivity of I[d]/I[g] and look at this in the context of our results, if the equation were to be applied (Lines 505-514): “The most significant term in the equation is I[d]/I[g], and our data shows that I[d]/I[g] is strongly affected by strain-related spectral changes. It therefore follows that the equation should be sensitive to strain, but the fact that not all strained samples produce calculated temperature shifts of the same direction or magnitude suggests that the process is more complex than simply strain or temperature having an effect. Regardless of cause, however, an error in calculated temperature of ±10°C in a stratigraphic sequence with an overall temperature range of only 25°C highlights the importance of context when estimating temperatures using this method. For example, if using this temperature data to reconstruct a burial history, then a strained sample might be ‘out’ by over a kilometre, or it might give the correct value. It is therefore important that more work is done to calibrate Raman geothermometers in rocks which have undergone strain in natural environments.”

8. Finally, some discussion should be given to the work examining the Raman response of kerogen to high pressures in laboratory settings. Certainly, the rate of strain between natural samples (as examined here) and laboratory strained samples will be different, but I believe that prior laboratory experiments can provide much insight into the processes under study here. Start with:

Huang et al., 2010, In situ Raman spectroscopy on kerogen at high temperatures and high pressures, Phys. Chem. Minerals, 37, 593-600.

Thank you for this suggestion. We have added as part of Section 6.2 discussing RBS that “pressure also affects peak positions” in Line 441, and have found the references within the Huang paper particularly useful.

Minor Revisions

1. Line 30: define “reliable”. This word is subjective without context.

The word “reliable” has been removed, with the sentence now reading “…to develop temperature equations that are based on Raman spectral parameters and are applicable across a range of settings…” (Line 27)

2. Line 40: The use of the term “organic carbon nanostructure” is misleading. Raman spectroscopy of complex geologic organic matter typically only reports on the aromatic character of the organics due to resonance effects (i.e., the electronic bandgap of the aromatic moieties matches the energy of visible wavelengths

commonly used as Raman excitation sources which pumps the response from these functional groups). Suggest rewording as “organic matter aromaticity”.

Thank you for this suggestion. However, we have decided to continue with the term “nanostructure” because the aim of this paper is not to analyse the precise structure, bonds, or resonance of the molecules involved but rather to address changes in Raman spectral parameters due to strain. We do not try to discern what these changes precisely imply in terms of the molecular structure of the carbon involved as this is beyond the scope of this paper (but would be worth future study). The term “nanostructure” is only mentioned for the purpose of putting the spectral changes into some context; we do not wish to overcomplicate the concepts involved. We refer the interested reader to the appropriate literature (Line 2).

3. Line 44: Define “FWHM” at first usage.

Definition now included (Line 46). Thank you for pointing this out.

4. Line 154: Define “BRGM” at first usage.

Clarified as the French geological survey (Line 201).

5. Line 171: The statement “Raman spectroscopy measures the wavelengths of backscattered….” is not entirely true. First, Raman scatter occurs in all directions to a degree, not just in the backward direction. Second, Raman is used to measure much more than just different forms of organic carbon. Rephrase.

Thank you for pointing out this oversimplification. We have rephrased the sentence to read “Raman spectroscopy measures the wavelengths of radiation produced by inelastic (Raman) scattering during the de-excitation of electrons in different molecular bonds, in this case focussing on those involved in different forms of organic carbon.” (Lines 222-224)

6. Line 172: More than just “peak temperature and strain conditions” are important for determining the molecular structure of geologic organic carbon. I would argue that biologic origin, depositional conditions, erosion, exposure to oxidants, and microbial activity are just as important as temperature and strain (and perhaps more so!).

This is a valid point and certainly highlights the depth to which organic carbon has been (and still needs to be) studied. However, due to the risk of overcomplicating what actually our methodology section, we have only included an additional note on biological origins here. The sentence now reads: “…depending on many factors during both deposition and burial: these include, but are not limited to, initial kerogen type, peak temperature, and the strain conditions…” (Line 225)

7. Line 173: The statement “Initially, the carbon will exist in the form of fossilized organic matter.” is not correct. Initially all of this organic carbon was from living carbon-based life.

We have altered the statement so that it now reads: “In the initial stages of burial, the carbon will have a nanostructure resembling that of kerogen…” (Line 230)

8. Line 180: What does “excess inorganic carbon” mean? Excess of what?

The word “excess” is indeed unnecessary and has been deleted.

9. Line 185: Change “lots” to “co-adds”.

Done, thank you.

10. Line 188: In figure 3 the caption states that a “Gaussian-Lorentzian hybrid” was used to fit the spectra. This is commonly termed a Voigt or pseudo-Voigt profile. Regardless, which profile shape was used to fit the data.

Thank you for pointing this out; we have changed the terminology to fit.

11. Line 204: The statement “The intensity of a single peak is a direct product of signal strength,…” is tautological. Raman intensity is proportional to the fluence of the input excitation source, the number density of oscillators in the probe volume, and the Raman cross-section (itself a function of the molecular polarizability).

Although of course correct, we consider this explanation to be too complex for the point we are trying to make. When dealing with sediments such as those in this study, absolute intensity of the spectral peaks can vary by an order of magnitude whilst intensity *ratios* remain constant, which is why intensity values of individual peaks are not used. We have, however, rephrased the sentence to make it clearer (Lines 269-274): “The intensity of a single peak is a direct product of signal strength, i.e., how many Raman-scattered photons come into contact with the detector. This can be affected by several factors including the amount of carbon present within the laser spot, or the strength of the laser. It is therefore more common to use the ratio between the D- and G-peaks (I[d]/I[g]), which will be characteristic of the nanostructural features regardless of signal strength.”

12. Line 206: The statement “The G-peak is in fact a composite of three spectral bands…” is not correct. For less ordered carbonaceous materials the G-peak is best represented by a single peak, for higher metamorphic grade organic matter there can be another peak or two in there, but for graphene (arguably the most ordered carbonaceous material) there is only one G peak whereas for single-walled carbon nanotubes the G-peak is split into G- and G+ peaks. What I am saying is that Raman spectra of carbonaceous materials is incredibly complicated and so definitive statements such as this are inappropriate.

You are correct to highlight that such a definitive statement is inappropriate; the wording has been changed to reflect this (Lines 279-282): “The G-peak defined here can be considered a composite of up to three spectral bands (D2, G, and D3) depending on metamorphic grade, but at low maturities such as those in this study they are difficult to distinguish and can be collectively referred to as a single peak (Beyssac et al., 2002; Muirhead et al., 2021).”

13. Lines 216-217: What do you mean by “pure graphite”?

Changed to read “complete graphitisation” (Line 289)

14. Line 219: The statement “…shows an increase in RBS with increasing temperature at higher maturities…” is tautological. Increased temperature = higher maturity.

What you refer to as a tautology is intended to highlight that at lower maturities this trend is not always observed.

15. Line 245: The wrong paper by Barker and Pawlewiscz is cited. You are looking for Barker and Pawlewiscz, 1986, The correlation of vitrinite reflectance with maximum temperature in humic organic matter, Lecture notes in Earth Science, Vol. 5, Paleogeothermics, Edited by G. Buntebarth and L. Stegena, Springer-Verlag, Berlin.

Thank you for pointing out this mistake! This has since been corrected (Line 488).

16. Line 283: Quantify "weak trend".

Actual values have now been included in the text to quantify this (Line 358).

17. Lines 309-310: Remove "(although reversed)".

Changed wording for clarity, but we believe it is beneficial to leave it in (Line 442)

18. Line 373: The statement "…is kerogen-like carbon rather than coal,…" is misleading. Kerogen is operationally defined as insoluble sedimentary organic matter. Hence, coal is kerogen. Usually, coal kerogen is termed Type I, or gas-prone kerogen. Change statement accordingly.

We have changed "kerogen-like" to "amorphous" (Line 449.

19. Line 404: "…(as noted previously by Muirhead et al., in review)…". I don't love citing unpublished work. Also, this is definitely not "noted previously" as it is unpublished. Finally, this citation is not included in the bibliography. I suggest removing this citation.

This paper has now been published and hence the citation has been adjusted to fit.

20. Line 427: The statement "Since the Lahfid et al….." is unclear. Rephrase.

We have removed this section from the paper: see earlier comments.

21. Line 536: Space needed between "Michael" and "Raman".

Thank you for highlighting this.