Response to Reviewer 1

We thank the reviewer for their positive comments, and hope that these responses are satisfactory

Comment: The main conclusion of the paper is based on the comparison of CM in samples from different locations. While a quite large number of Raman spectra were collected, there is no description of how statistical analysis was carried out in the Methods section, for instance, how were differences determined for datasets with different number of samples? What statistical analysis was used? How was error propagated for datasets with analytical errors and replicates? How was PCA carried out? Were the data normally distributed? These are key questions regarding the robustness of the conclusion.

Actually, looking at Table 2, I would think that the statistical differences only occur for the Highly Graphitised CM in distal ESAS samples relative to the others and for Mildly Graphitised CM in terrestrial samples relative to the others. The others are mostly similar with a big standard deviation. I recommend using box graphs showing the median, mean and percentiles for each sample group rather than using Figure 3 (which is redundant showing less information than Table 2).

Response: We agree that statistical analyses form a key portion of the manuscript, and will make suitable changes to the text to explain these to the reader. Briefly, significance was calculated using the mean, standard deviation and number of samples, using a t test calculation to generate P values. PCA was carried out using the 'prcomp' function within the software package R. Instrumental and sampling error were not investigated for this study. Instrumental procedures were previously tested in detail (Sparkes et al., 2013). Each sample consisted of spectra collected from ~30 individual pieces of organic carbon, and a repeated sampling could have led to a different subset of organic particles being measured. However, the consistency within and demonstrable difference between each group of samples, is in line with previous analyses, indicating that 30 spectra per sample is enough to produce a robust characterisation of each location.

Offshore Raman studies involve the mixing together of terrestrial OC from multiple sources, and so the null hypothesis in any such project should be a homogenous distribution across the shelf. Figures 2 and 3, and Table 2, demonstrate this trend – there is a large degree of similarity between the samples, but careful PCA analysis shows that small differences seen in the distribution of the different spectral classes are significant and systematic.

We will produce the box graphs requested by the reviewer and include them in place of Figure 3 in the revised manuscript.

Comment: Regarding the presentation of data, I prefer to see the distribution of CM drawn in the format of Figure 2 rather than in Figures 4 and 6. While the latter is truly impressive, how reliable are the schemes given the scattered and uneven distribution of sampling locations?

Response: This presentation style, using standard algorithms from ArcMap, has been employed in a number of studies, utilising the ISSS-08 samples and around the world (see examples from the reference list: Vonk et al., 2012; Sparkes et al., 2015; Bischoff et al., 2016; Sparkes et al., 2016). Non-uniform sample distributions are never ideal but we feel that there are enough sample locations to support these interpolations.

We propose including the requested figures as supplementary information, if required by the editor.

Comment: The authors mentioned that black carbon particles smaller than submicron size is not detectable by Raman spectroscopy. However, it is also mentioned in the methods that hours of grinding does not affect CM crystallinity, suggesting no effects on Raman spectroscopy. I am a bit confused here. How big is the pool of black carbon "undetectable" for Raman in the total CM or black carbon budget? Is it possible that, during transport and winnowing, CM may be physically ground to smaller particles to escape the analytical window? How would this affect your data interpretation? In the end, I think it is very important to frankly point out drawbacks of the method as no method is perfect.

Response: Grinding for many hours reduces grain size but does not affect the degree of crystallinity of the graphite particles (see Nakamizo et al., 1978; Sparkes et al., 2013). Therefore, there is a risk of material becoming too small for the analytical window, but until that point the technique should robustly characterise the distribution of CM types.

Using a 50x objective, as this study employed, Raman spectroscopy can measure particles of about 1-2 µm and above. This includes silt and clay size fractions, and even in distal samples consisting of fine mud, individual grains can still be identified clearly. Colloidal material smaller than this is less likely to settle in shelf sediments, and could be winnowed into the deep ocean. The fraction of material ground to this small size is hard to quantify, but we believe that the majority of the CM present is within the appropriate size window for Raman analysis. Whether the material is sufficiently crystalline to produce measurable Raman spectra is another matter. Raman analysis of atmospheric SBC produces spectra with very broad peaks (Catelani et al., 2014), similar to very disordered lignite-grade terrestrial CM.

Atmospheric soot particles are typically in the nm range, these would be too small for this study and would not be identified as individual CM particles. Large SBC particles, eroded and transported by fluvial and coastal processes, alongside sedimentary matter, may be included in this analysis.

We are happy to clarify all of this in the text.

Comment: For the discussion part, I think it makes more sense to introduce PCA analysis first, followed by comparison of group mean values. I also think that some descriptions are repeated and can be shortened to increase the readability.

Response: *We are happy to switch sections 4.1, 4.2 and 4.3 so that PCA analysis is introduced and discussed before Highly Graphitised CM. Suitable alterations will be*

made to ensure continued internal consistency, and reduce redundancy as identified by the reviewer.

Comment: There are some minor mistakes:
Page 13: Line 15: …have been caused by…
Line 29: … is that it is preferentially…
Page 14: Line 7: no offshore trends

Response: *We will change the text accordingly.*