General Comments

Frieling et al., present records of carbon isotope fractionation from the resting cysts of dinoflagellates to investigate their utility in reconstructing ancient atmospheric CO2. This record of core-top material advances earlier work based on laboratory cultures (and based on sound theoretical basis) and so brings the community closer to confidence that this proxy may work in environmental settings. They show there are differences in carbon isotope fractionation between different species, emphasising the importance of single-species records, and show greater 13-C depletion in their core-top samples compared to cultured, motile organisms. The paper is interesting and makes an important contribution, but some of the analysis is unsatisfactory due to uncertainty about the age of the individual cysts in the "core-top" samples (detailed below). Therefore without a thorough treatment of that uncertainty (which is currently lacking) it's difficult to know whether this proxy has utility. There are certainly hints that it does, but unfortunately this paper does not yet demonstrate that compellingly.

Author response:

We thank the reviewer for recognizing the potential importance of our work and the constructive criticism. In the response below and in our revised manuscript we will further elaborate on (1) how the carbon isotope data from individual cysts has been treated and (2) further elaborate on the uncertainty in the age dating of the core-tops.

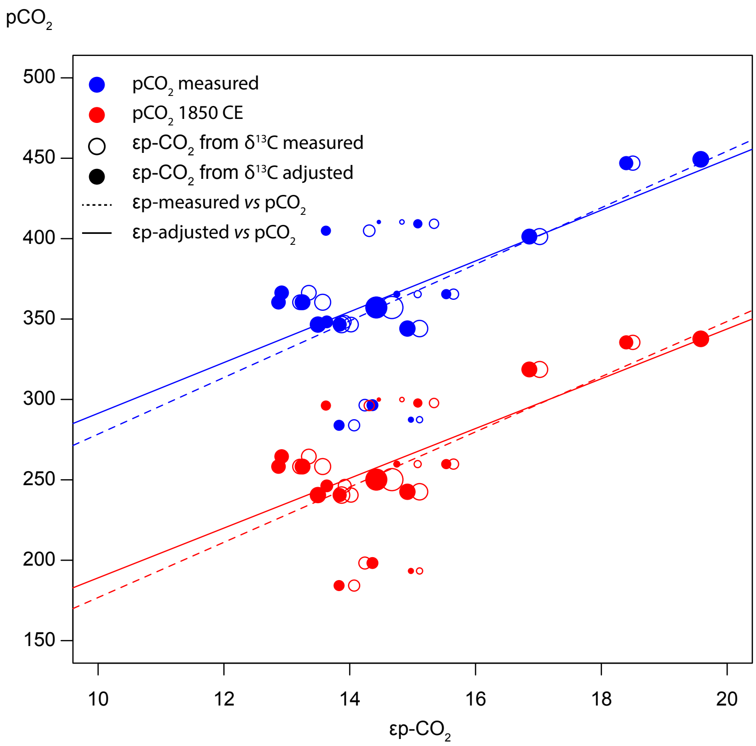
Specific Comments

The problem with using core-top samples is the substantial increase in atmospheric CO2 since the industrial revolution. As the authors note, it is highly uncertain whether the cysts are from the last week, the last year, decades or even centuries ago. The uncertainty around the contemporaneous CO2 is potentially very large. The "rough correction" to 1850 isn't really a correction at all, but an assumption which is not well supported, at best highly uncertain, and not really dealt with satisfactorily in the later analysis. The best approach (although expensive) would be to 14-C date some of these samples to see when this material actually dates to. The cheaper, and for this present study, more plausible approach would be to propagate through what is a really quite large uncertainty and see whether the conclusions still hold. Lines 147-8 state that "With the exception of pCO2, we hence assume all parameters (SST, SSS, nutrients) to be constant over the period the core top samples represent." A fundamental problem here is that the authors have little information (or at least present little data) about how long a period of time the core top samples do in fact represent. I'm not sure that the approach taken to this, systematically removing the most 13-C depleted samples is appropriate. Whilst it is certainly plausible that these individuals represent modern samples, the evidence is fairly circumstantial, and they could represent another confounding variable. What is the impact on the analysis if these samples are not removed?

Author response:

The reviewer comments on the potential of age-mixing of individual cysts in core-tops. This is very much a valid concern as we acknowledge in our original manuscript (lines 196-201). The main challenge is that core-tops (the top-most 2 cm of sediment) contain a range of ages of sedimentary components. While it might be possible to ^{14}C-date carbonate or organic matter also these materials will derive from different times in the past and a single measurement will not show the range of ages of individual components such as for our dinocysts (i.e. the age-distribution of our individual dinocysts). Ideally one would date single-cell using ^{14}C analyses but that is technically not feasible.

As the reviewer correctly points out, it is important to show the potential impact of our data-treatment and we include these analyses in a supplementary file to the revised manuscript. In general, the impact of removing analyses based on exceptionally low amounts of C and exclusion of outliers has no appreciable impact on the regression parameters. The largest uncertainty comes from the recent addition of anthropogenic carbon, the Suess effect. Comparing our calibration including and excluding the Suess effect is to our opinion therefore the best approach to estimate the maximum uncertainty in the regressions.



We will add the figure above to our supplementary information. The figure shows the difference between measured and adjusted pCO_2 and $\delta^{13}C$ (εp-CO_2). Open symbols indicate measured $\delta^{13}C$, closed symbols represent data after eliminating small signals (<0.2 Vs) and outliers. Blue dots represent measured CO_2 values and red dots indicate the CO_2 around 1850 CE. For each dataset a simple linear regression, weighted to the number of measurements, is given. Dashed lines utilize measured $\delta^{13}C$, solid lines are from final $\delta^{13}C$ data. The red solid line is used in Figure 5A.

The figure above illustrates the maximum error that may result from the Suess-effect by constructing a calibration with uncorrected data for each of the parameters (measured $\delta^{13}C$ and CO_2) and we compare this to the original calibration (solid red line). The main difference is the intercept of the calibration which is offset by the difference in atmospheric pCO_2 between 1850 and 2000 CE (*ca.* 100 µatm) resulting from adjustment of measured CO_2 levels to pre-industrial times. The slopes of all combinations are statistically indistinguishable. Further, we must stress the difference between the intercepts of the calibrations is a worst-case scenario and unlikely an accurate reflection of true uncertainty caused by the Suess-effect.

We also fully agree with the reviewer it is critical to show that the data-treatment is free of potential preconceived biases regarding which values to include or exclude. We will therefore further clarify in our revised manuscript how we first corrected for instrument drift at low C. This correction is exclusively based on repeat measurements of the IAEA-PE ($\delta^{13}C$ certified -32.15 ± 0.05‰) standard which show a convergence towards ~ -27‰ at the lowest intensities (original manuscript lines 172-174; see also Van Roij et al., 2015). After performing the drift correction, we subsequently identify outliers within the species-specific populations measured on a single sample. We find a greater proportion of negative outliers (original manuscript line 215) compared to positive outliers (hence a skewed data set; original manuscript Table 1). Since we correct using the PE standard and exclude both negative **and** positive outliers there is no assumption that the most negative are the most recent cysts (which one might assume based on the Suess effect but would be an *a priori* interpretation).

The reason why we argue our data-treatment has removed most of the temporal bias that may exist between samples is as follows:
If we assume that (a) statistical distributions of $\delta^{13}C$ cyst populations are dominated by both pCO_2 and $\delta^{13}C$-CO_2 trends and that (b) these populations include a portion of both pre-industrial and recent times – the trends in both pCO_2 (higher pCO_2 > higher εp > lower δ^{13}Ccyst) and $\delta^{13}C$-CO_2 (lower $\delta^{13}C$-CO_2 > lower δ^{13}Ccyst) would impose a (negative) skew on the δ^{13}Ccyst populations. Prior to correction for skewing at low intensities and outlier removal some, but not all, samples show a negative skew, while after outlier-removal the vast majority is statistically indistinguishable from a normal distribution (Table 1 in the original manuscript). In addition to the supplementary file showing the minor impact of excluding outliers (see above), we will clarify this rationale in our revised manuscript.

Ie L 213-4 "We assume these assemblages are representative of ocean conditions prior to the massive increase in anthropogenic carbon emissions." This is very sizable assumption which whilst plausible is not currently supported by very much evidence.

Author response:
Please also see response above – we now supply a more elaborate reasoning why we assume the vast majority of the analysed cyst populations is considered to represent pre-industrial times. We will clarify in our revised manuscript that this is also based on the assumption that core-top material includes an age range, likely centuries up to millennia. For core-top samples a major impact of the Suess effect may be detectable only in the relatively few cysts that were formed in the past 50 years.

It appears from Figure 5a that no uncertainty at all has been applied to the assumed CO2 value – is this correct? This is not a fair assumption given the uncertain age of each cyst. In fact a plain reading of Figure 5a suggests that, rather than supporting a function between CO2 and ep apart from at <240 uatm CO2, ep is effectively constant, and only slightly higher above 310. Why therefore has 240 uatm been emphasised?

Author response:
In our revised manuscript we will clarify that after removal of outliers and, arguably, calculating the maximum effects of fossil-fuel derived CO_2, we do not expect any other major biases in the CO_2 gradient (see also Figure in previous response). This implies, regardless of the uncertainty on the actual CO_2, the slope of the relation would not significantly change. However, the reviewer is correct in pointing out that this may influence our interpretation at what value of CO_2 εp might become (in)sensitive to CO_2 changes. We will add further nuance to these statements in our revised manuscript and explain potential pitfalls in our approach. We also make it clearer that the 240 µatm – level is potentially a low estimate of (in)sensitivity and that, for practical reasons, the quadratic calibration should not be used below this level.

In addition, we will clarify that our initial Figure 5D and 5E we propagated (via Monte-Carlo analyses) a 5% error on the measured CO_2 and nutrients, as well as the standard error on the mean dinocyst $\delta^{13}C$. We erroneously omitted these uncertainties in Figure 5A-C and these will be added to the revised Figure. We will clarify our approach in the revised text.

Whilst the data presented here are interesting and important, the analysis at present is not sufficient to support the conclusions drawn robustly.

Author response:
We hope with the clarifications above, the calculated (maximum) uncertainty and the proposed adaptations of our revised manuscript we have alleviated the reviewers' concerns.

Technical corrections
22 use of "significantly" if this is meant in the statistical sense, please add p and n values, else reword.
Author response:
Changed to 'appreciable' as these data are not directly compared here.

24 ibid "significant"
Author response:
We will retain this statement but given the number of comparisons (20) we cannot provide p or n values of each comparison in the abstract of the manuscript. For significance we refer the reader to Figure 3 and keep the original generalised statement in the abstract.

40-42 This is a slightly eccentric choice of papers to cite here. At a minimum add an "e.g." but better to make it clear why these papers or a more comprehensive survey of the pCO2 proxy literature.
Author response:
We will include "e.g." before the cited references.

43 "However, many of the organic compounds used for CO2 reconstructions are not related to a single species, genus or even group." A fairly sweeping statement here not supported by any references. Which records and compounds are you referring to?

Author response:
We clarify in our revised manuscript this refers to proxy substrates in general (bulk organic matter and biomarkers such as phytane and alkenones) mentioned in the previous paragraph and we will refer to literature examples to illustrate this.

54 "extremely long-ranging" in space or time? Please be specific and it time list age range.
Author response:
We will add a statement to clarify this refers to *Operculodinium centrocarpum* and *Spiniferites* species as used here (see lines 65-67 for their age-range).

70 Should "cyst species" by cyst-forming species?
Author response:
To clarify we will rephrase this to 'εp derived from motile cells from controlled growth experiments can be translated to that of cysts formed in the natural environment.'

90 "Using standard palynological techniques" Please provide a citation.

Author response:
We will refer to Brinkhuis et al. (2003) who described the standard cold HCl/HF acid-digestion procedures employed to obtain palynological residues.

94 “ultraclean water” what is this? Ie quote a specific measure such as resistivity if reverse-osmosis teqnique has been used
95 “milliQ” is a brand name not a type of water. Please revise.
Author response:
We will change both these to demineralised water.

L104-5 is the 0.3-0.4 permil number precision or accuracy? How has accuracy been determined.
Author response:
We thank the reviewer for pointing out this unclear reference. We will remove ‘accuracy’ here as this statement was meant to refer to precision.

L343-4 “Badger, 2003, 2021;” These are two difference Badgers. Check BG style but likely need to include initials (lots, because they share first first name initial too).
Author response:
We thank the reviewer for highlighting this – we will revise to “Badger, M.R, 2003, Badger, M.P.S., 2021”.

L 385. I’m not sure this is sufficient to meet the journal data policy. Pangaea doi should be available at publication.
Author response:
We will include the DOI as soon as that is available.