Orbital CO$_2$ reconstruction using boron isotopes during the late Pleistocene, an assessment of accuracy – de la Vega et al

review by William Gray (william.gray@lsce.ipsl.fr)

Boron isotopes in planktic foraminifera are a widely used proxy for reconstructing ancient CO2. As other studies have done previously, de la Vega et al use ice core CO2 as a test of accuracy; to this end, they present new G-IG boron isotope data, including data from a new sediment core site in a location close to equilibrium with the atmosphere today. This is only the second record of its kind (i.e. a record from a site close to equilibrium and over multiple glacial cycles) and is a very welcome addition as it helps overcome the assumption that a single site (i.e. ODP 999) has remained in equilibrium with the atmosphere. They go on to assess the accuracy with which CO2 can be reconstructed, and possible causes of the (albeit relatively minor) discrepancies (dissolution, second carb system parameter, calcite-borate calibration). Overall, I found the manuscript to be clear, well structured, and well-reasoned.

I have the following suggestions:

I think a greater exploration/discussion of the thermal influence on the carbonate/borate system (via the dissociation constants) in paleo CO2 reconstruction is warranted – basically, how sensitive is CO2 to accurate SST reconstruction? More sensitivity tests could be implemented - including keeping temperature constant throughout. Overall, I think further exploration/discussion of the thermal effects are warranted as, my guess is, this could be an important source of bias/uncertainty and it's helpful to understand what we need to improve. The authors mention they also use our iterative pH correction, and briefly mention this in the text, but I think it warrants further discussion; there is very good evidence from culture studies that pH influences Mg/Ca, and this is an influence we know is going to covary with atmospheric CO2.

The authors mainly rely on the calibration of Anand et al 2003 to derive Mg/Ca SSTs to use in their CO2 reconstruction. Although widely applied I don't think the calibration of Anand et al accurately describes the relationship between Mg/Ca and temperature (see figures 5 and 6 in Gray et al 2018). This is very apparent if you compare the measured CTD temperatures at the sediment trap site used by Anand, and the temperature calculated using Anand's own Mg/Ca data and their calibration line (see figure below). Using the calibration of Anand et the Mg/Ca are almost always too warm, and the seasonal cycle is about half of what it should be as winter temperatures are 4 degrees too warm (there is no way to explain this by 'sampling issues' as the Mg/Ca SSTs are warmer than any individual CTD measurement ever taken at the site in winter, and this is BATS so it's about the most heavily sampled place in the ocean). Basically, if you use Anand et al on their own data, you get the wrong answer. If the authors want to use Anand et al, I think there needs to be more justification (and 'it gives a better fit to the CO2 data/the core top SSTs' isn't a great reason).
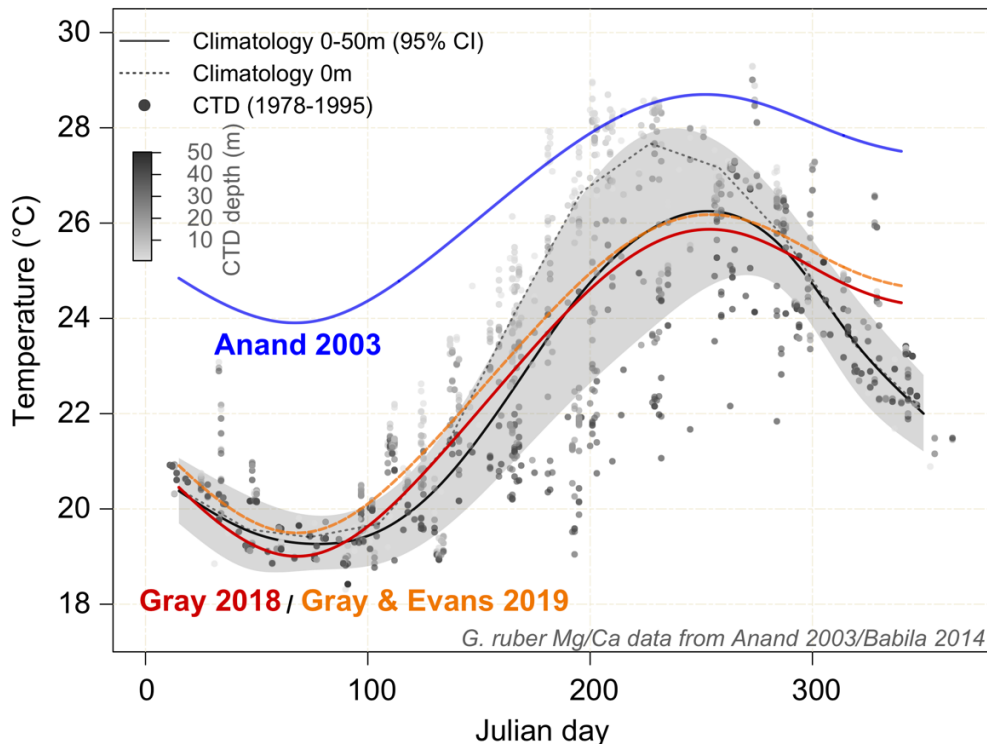
Figure above shows a comparison of Mg/Ca and CTD temperatures at the Sargasso Sea sediment trap site (the site used by Anand to derive their Mg/Ca calibration). I'm showing a LOESS fit to the Mg/Ca data (colored lines), rather than the individual data points to make it legible. The grey dots show all the individual CTD measurements ever taken at this site. The black line/grey shaded area is the expected temperature (and 95% CI) at the habitat depth of G. ruber.

For the leaching experiment is T. sacculifer with or without sacc? This is really important as the lines of argument regarding the differential dissolution of the gametogenic calcite versus the rest of test only hold if it has a sacc (note there are similar arguments for a d18O dissolution effect in this species). The authors mention they didn't assess weight loss (a shame), but are there really no TE data for these samples? Was an aliquot not analysed as part of the boron isotope analysis to check for cleaning etc? If TE data are available it would be really informative to show how Mg changes with the leaching, as having some metric to be able compare to real world samples would make these leaching experiments much more useful.

For the uncertainty and the discussion of changing DpCO2 its worth noting that Earth System Models under glacial forcings typically simulate very small changes in DpCO2 – this can be seen in the figures below which show the pH difference minus the mean ocean pH difference in the IPSL model between PI and LGM forcings (taken from Gray and Evans 2019) – basically most of the ocean just reflects the change in atmospheric CO2 (95% range within ±0.05 units, equivalent to about ±40 uatm DpCO2). There is almost no residual pH change at the two sites used in the present study. I think this is really encouraging for paleo CO2 reconstruction from boron isotopes and I think we could exploit ESMs to understand/quantify the DpCO2 aspect of the calculation. Doing this kind of exercise with a larger ensemble would be really good way to test this central assumption in surface ocean carbonate system based pCO2 reconstructions in the future.

For the present study, I'd be happy to provide some version of the figures below (perhaps in DpCO2 space, rather than pH) to include in the manuscript (say below Figure 1) if the authors thought it would be helpful and wished to redraw them and include them.
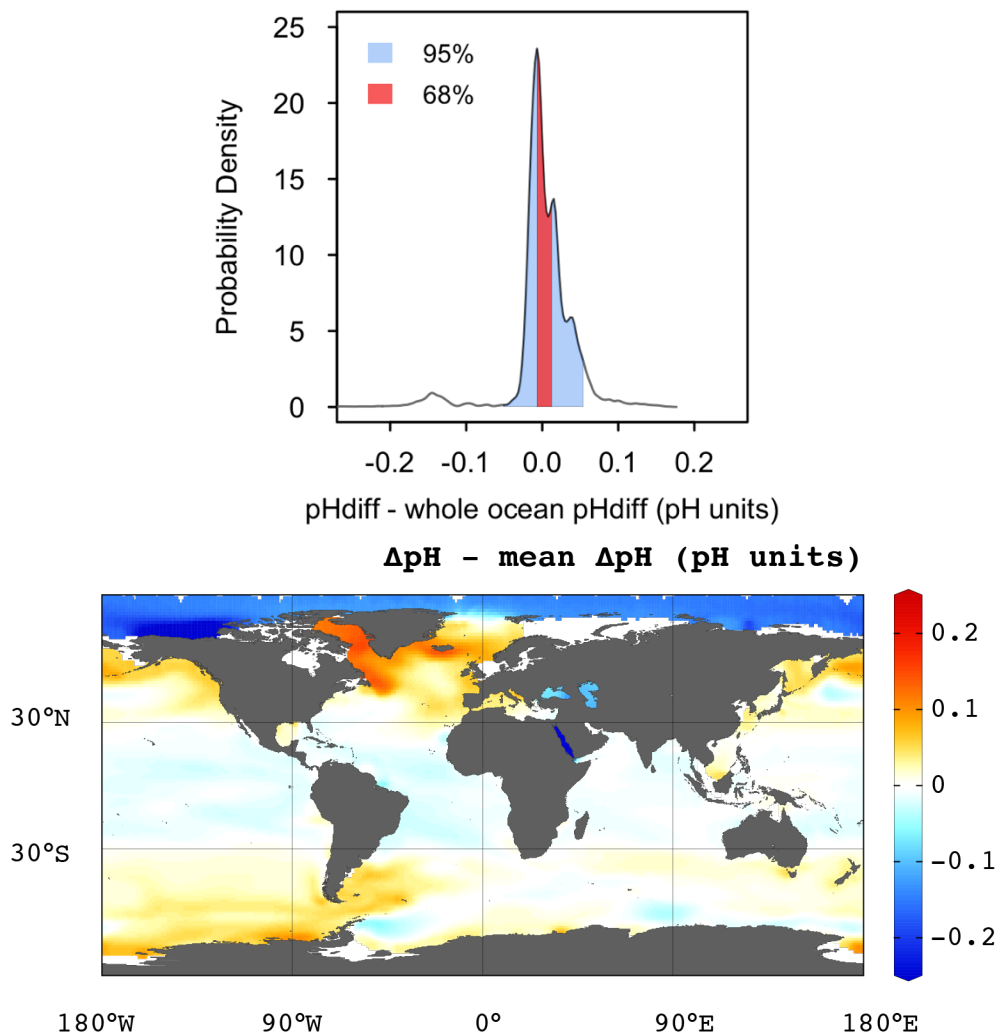

Figure above (from Gray and Evans 2019) pH difference minus the mean ocean pH difference in the IPSL model between PI and LGM forcings. Most of the surface ocean shows very little pH change beyond the impact of changing atmospheric CO2, which is reassuring for trying to reconstruct atmospheric CO2.

For the second carbonate system parameter the authors use ALK, taking a flat 175 umol/kg range about the modern value – how is this range distributed around modern value? Is it weighted more heavily to higher values to account for likely higher glacial ALK a la Martinez-Boti 2015? They suggest the ALK variations needed would be too large to be sole cause of discrepancy, but we I guess cannot rule out more minor/systematic ALK changes over G-IG cycle cannot explain some part of discrepancy.

Detailed comments
line 58 – dissolved inorganic carbon (DIC)
line 138 – I think you could sure up this assumption or try to quantify the likely uncertainty in DpCO2 using ESM model output (see point above)
line 153 – Rebotim et al 2017 (biogeosciences) is a good reference for G. ruber habitat depth

line 198 – is 0.25% the relative concentration of the TPB? It could still have a big effect if it has a funky d11B, so better to report the absolute per mil values of correction
line 201 – with sacc or without sacc?
line 237 – I think using Anand needs some real justification (see discussion/figure above)
line 241 – an aside as you are not accounting for dissolution downcore, but it should really be parameterized as a function of Omega rather than depth
line 246 – I think more details are need of the method you used for the pH corrected approach, and greater discussion of the results are warranted later
line 254 – how is the ALK range distributed around modern value
line 273 – why not use the ESM output to try to quantify the likely DpCO2 uncertainty?
line 390 – sacc or no sacc?
line 431 – im not sure this is the reason Pacific CaCO3 preservation increases during glacials, as there is a lot of evidence for reduced ventilation of the deep Pacific in glacials (e.g. Anderson et al 2019). An ALK increase due to the reduction of CaCO3 burial in the deep Atlantic seems more likely (e.g. Cartapanis et al 2018), but a lot of this discussion seems somewhat superfluous. Its enough to say that the fragmentation should give an indication of relative changes in dissolution and describe the patterns seen.
line 461 – why?
line 467 – in Gray et al 2018 we found no systematic Mg/Ca offset between morphotypes across the Indian and Atlantic
line 483 – I really think ESMs can be useful here
line 524 – w/ or w/o sacc?
line 545 – really no TE data?
line 595 – worth noting optimized calibration is still very similar to other planktic calibrations i.e. it doesn't require something radically different to what we might expect
line 623 – the disequilibrium pump has garnered a lot of attention recently and is worth noting (e.g. Egglestone and Galbraith, 2018)

figure 1 – better to use red-white-blue colour scheme for positive and negative anomalies. why not also add a map of modelled LGM-PI DpCO2 differences below with the core sites indicated?
Figure s14 – why not plot on the other planktic lines to enable comparison with optimized line