Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2020-269-AC2, 2020

Interactive comment on "Technical note: Single-shell δ^{11}B analysis of *Cibicidoides wuellerstorfi* using femtosecond laser ablation MC-ICPMS and secondary ion mass spectrometry" *by* Markus Raitzsch et al.

Markus Raitzsch et al.

mraitzsch@marum.de

Received and published: 9 September 2020

AC: Lubos Polerecky's very helpful comments and suggestions, particularly from his analytical view, on our manuscript are very much appreciated and will all be addressed in the revised manuscript (listed below).

Raitzsch et al. provide a detailed comparison between SIMS and LA-MC-ICPMS measurements of delta-11B in individual shells of benthic foraminifera. They show that intra-shell and inter-shell variability is significantly lower for the LA-based technique

Printer-friendly version

Discussion paper

compared with SIMS, which they attribute to the larger volume sampled by the LA-based technique. Importantly, they show that both techniques yield very similar "average" values to those obtained by the traditional bulk measurements based on dissolved specimens. They conclude that the traditional bulk-based analysis is still the preferred approach for paleo applications, but demonstrate clearly the advantages and limits of the microanalytical techniques. The manuscript is well written and clearly organized. Also the figures are clear and of excellent quality. I only have a few minor comments and questions. I recommend the manuscript for publication after these minor issues have been resolved by the authors.

Technical comments/questions/suggestions: l.69: Please formulate more clearly the *aim* of the study. 'What' do you want to achieve, and especially 'why'?

AC: That's right, the aim of the study is not clearly enough explained, but just vaguely outlined between l. 38-47. We will emphasize this in the revised version.

l.154-155: Please clarify how this variability was calculated. Since 2*sigma is used, it may be confused with 2*sigma of the individual measurement's precision. And since permil units are used, it may be confused with the coefficient of variation (which is in percent). To avoid confusion, best would be to clarify in one sentence that 2*sigma here actually corresponds to 2*SD of n individual measurements (if I understand it correctly). Or is it 2*SE (standard error)?

AC: Yes, the reported 2*sigma variability is the 2-fold standard deviation, derived from the individual measurements. This is not to be confused with the measurement uncertainty (=precision), which is dependent on the ablation time and is also given as 2*sigma (=2*SD). So we use sigma as a statistical expression just to clarify that the SD is reported, and not the SE. We will explain it in more detail in the revised manuscript.

l.157: unclear why such inistinguishability should affect variation in measured data. Please explain, or provide an alternative explanation.

AC: Right, this part is quite confusing and probably also not reflecting the truth. It is true that we observed the largest variability in the knob area that might be attributed to a signal mixture from multiple juvenile chambers, but it may also be due to the higher number of measurements compared to the other chambers. In addition, there were also similarly large variabilities found in chambers f-8 and f-9. So we have to revise (or delete) the according statement. Thanks for hinting at this inconsistency.

l.168-169: Please clarify how this was derived/deduced. Intuitively it is expected that variability in measurements is lower if larger volume is sampled. But it is unclear how you arrived to those values (e.g. âĹij0.3 permil).

AC: Thanks, we missed to describe how we calculated this. We simply applied the following function to estimate the variability reduction: u(V2) = u(V1)/sqr(V2/V1), where u(V1) is the variability for a quoted volume, and V1 and V2 represent the two different volumes that are compared. We will add this information.

l.176/fig.4: Please clarify representation of the data in polar plots in Fig. 4. I understand that the "phi" coordinate corresponds to the chamber, but it took me a while to figure out that the r-coordinate (scale -7 to 3) corresponds to Delta-delta11B. Also I am wondering whether it would be more beneficial/transparent to show each Delta-delta-11B datapoint rather than average Delta-delta-11B deviations derived from measurements of multiple specimens. Averages may be misleading, as we know.

AC: I agree, the so-called coxcomb chart seems to be difficult to read at the beginning, but once it's understood, it is a nice way of presenting such data. In the updated figure, it will be clearer that red represents positive and blue negative Delta-delta values. I have tried a couple graph types, also plotting all individual datapoints, as Lubos suggested, but this resulted in a quite confusing graph due to the large number of measurements. For our aim to eventually observe trends, plotting averaged deviations seems to be the catchiest way.

Did you test whether the decreasing trend between f and f-5 is significant, or you can only state "the deviation tends to decrease"?

AC: Yes, we applied the Wilcoxon-Mann-Whitney approach to test the significance of d11B differences between chambers (l. 178 ff). Also, Dennis Mayk made this a subject of discussion, and I realized that I have walked right into a trap when testing the null hypotheses. Because of the few datapoints for each chamber, I applied the statistical test on Monte-Carlo simulated d11B values (n=10) around quoted uncertainties, yielding p values smaller than 0.05, meaning that the differences in d11B between chambers f-1 and f-5 are statistically significant at a 95 % SL. This artificially increased population size, however, led to a biased uncertainty estimation, which was also subject to papers in mathematical journals (e.g. Lin et al. (2013), Too Big to Fail: Large Samples and the p-Value Problem, http://dx.doi.org/10.1287/isre.2013.0480). If just the original data are taken into account, both the WMW and Welch t-test yield p values ∼0.07, and hence the d11B differences between the chambers are not statistically different at a 95 % SL, based on the small datasets of this study. I have to apologize for this incautious and naive application of statistical tests on our data. This will be corrected in the revised manuscript.

l.200: I am wondering why the authors report median instead of, for example, the mean? If it leads to a different mean in comparison to the bulk-based analysis, it should be discussed why such a difference exists. In any case, I think it needs to be clarified why median was used. Similar on l.254.

AC: The differences between medians and means are small, e.g. the SIMS median is 16.08 ‰ and the mean is 16.19 ‰ while the LA median is 15.91 ‰ and the mean is 16.17 ‰. However, the median is less sensitive against outliers than the mean and also represents the average value of a non-Gaussian distribution. The median is also equivalent to the average shown by the violin/box plots in Fig. 5. We will insert a small sentence why we chose the median instead of the mean.

l.211: yes, I agree, but it would be useful to expand this argument towards the *source* of this variability (e.g., shell-to-shell differences in the intra-shell heterogeneity?).

AC: We think this goes slightly beyond the scope of this study, and is difficult to answer, based on our database. Maybe the range of isotopic compositions of the trigonal BO3 hosted in the calcite lattice is very large (not yet examined on the molecular scale), and hence better resolved the smaller the scale of the analytical technique is.

Figure 1: It is rather confusing to see signals for 10B and 11B centered on the same mass (10.25). Is it really so? And why? I am not familiar with the Daly detector principle.

AC: On a multicollector ICP-MS, the Quad lenses are tuned in a way that the peaks of different isotopes (here 10B and 11B) coincide, i.e. the incoming ions hit the respective detectors simultaneously, where the one for high mass (11B) is on either and the low mass (10B) on the other side of the Center cup. Once the peaks coincide and the peak center is set, the information on the position of the peak center is read by the Center cup. Of course, 11B is measured on 11.009 amu and 10B on 10.013 amu, but the position of the peak center is in relation to the Center cup, and might slightly change on a daily basis, depending on the tuning parameters. We will slightly modify the figure caption to make it clearer.

Figure 3: Please verify the expression for 2*sigma in the graph (in red). First, the factor 1000 does not make sense if cps is in counts per second (perhaps it does if it is in kilo-counts per second). Second, if I substitute 300,000 and 300,000/4.9 for 11B and 10B, I get a factor 8.8, not 8.2. In my opinion, the formula should read as 2*sqrt(1/counts(11B)+1/counts(10B)), where counts(11B)=cps(11B)*1s*n and similarly for 10B. This is a formula for the Poisson error of 11B/10B based on counting statistics. In this formula the factor is then 0.00887 at B11=300,000. Please verify cps vs. kcps.

AC: The factor 1000 is because the boron isotopic composition is given in permil (see eq. 1 in the main text), so it's expressing the relative difference from the standard value. The measurement uncertainty (i.e. the internal precision) must hence be multiplied with 1000 to have the number in permil as well. We agree that the formula we provided is quite cumbersome, but it is mathematically correct. Based on the simpler formula provided by Lubos, we slightly modified it to more easily enter the number of cycles (n), and multiply if with the factor 1000 to obtain the result in permil expression. The final formula is now 2*sqrt(1/cps(11B)+1/cps(10B)) / sqrt(n)*1000, which will replace the current one in the revised figure. Concerning the obtained factor at a countrate of 300,000 cps for 11B, the ratio between 11B and 10B is in nature in the order of 4, and not 4.9 (11B=80.1 %, 10B=19.9 %). Therefore, if 11B is measured at 300,000 cps, 10B is recorded at approximately 75,000 cps, thus giving a factor of 8.8 using the formula above. We will add the expected countrate on 10B in the revised figure.

Editorial suggestions: l.24: unclear why the word "presumably" is used in the abstract. It would help if the sentence is reworded to clarify what is certain and what is not (i.e., what is estimated).

AC: Right, "presumably" is a too careful term. We will replace it with "estimated to be".

l.39: would be useful to cite few examples of such studies.

AC: These are the same references as in lines 32-33, but we will list them here as well.

l.104: perhaps it should read "45 cm3 *and* ablated"?

AC: Thanks, will be corrected.

l.126: remove "and" before "that"

AC: OK.

In caption to Fig. 4, it should read "inset", not "inlet". Similar on l.195.

AC: Right, slip of mine.

l.184: remove "large-scale"

AC: Thanks for that, was a leftover from a former sentence.

l.279: "French" - uppercase F.

AC: Will be corrected.

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2020-269, 2020.