Biogeosciences Discuss., 7, C285–C295, 2010
www.biogeosciences-discuss.net/7/C285/2010/

Biogeosciences Discussions

Interactive comment on "Technical Note: Determining the size-normalised weight of planktic foraminifera" *by* C. J. Beer et al.

B. Metcalfe (Referee)

brett.metcalfe@falw.vu.nl

Received and published: 23 March 2010

The importance of determining the impact of size on the weight of planktic foraminifera, and therefore the relevance of weight as a proxy for test thickness is crucial in our understanding of the carbonate ion effect. The paper aims to objectively determine if Size Normalized Weight (SNW) of planktic foraminifera based on measuring size (Measurement based weight – MBW) has an greater accuracy than through using a sieve mid-point size (Sieve based weight – SBW). The paper through the use of an automated image capture device uses a novel approach of measuring shape area (referred to as Test area within the paper) and the mean of the diameter's bisecting the centre of planktic foraminifera to determine test size. There findings indicate that the use of shape area provides a more accurate size measurement upon which to

normalize test weight.

This paper requires Major Revisions in its current form in order to address some key issues: firstly does the experimental methodology adequately meet the main aim of the paper in a non-biased approach – the main issue is that size has not been properly constrained against weight. The use of mean test diameter and mean test area for a single aliquot (where measured test diameter is already the mean of the lengths bisecting the specimens centre) against the mean weight fails to improve upon the reliability and accuracy. Were individuals weights used to compare the different approaches of a) sieve mid-point size estimation (SBW – Barker and Elderfield, 2002; de Moel et al., 2009; Moy et al., 2009) and b) measured size estimation (MBW-this paper) then a more accurate evaluation of the methodologies of SNW could be achieved, as it currently stands the MBW methodology enjoys more degrees of freedom in this experiment.

Secondly, constraining the size of a 3D object through the use of a planar 2D image (length; width) should have had sufficient quantification especially with the loss of the third dimension (depth) in regards to its affect on weight.

Thirdly, in order to critique previous publications (Barker and Elderfield, 2002; Moy et al., 2009; de Moel et al., 2009) it would need a far greater robustness in methodology (e.g. individual weight, individual test min. and max. size). In its current form its potential to become a paper cited just for the statement that the sieve "error estimate is of the same order of magnitude as the change observed in published downcore records of SNW" would lead to erroneous citations in future research without sufficient quantification. Not that it should be rejected because it goes against the grain, the aim of the paper is worthy of praise, however the methodology used to test the hypothesis limits the results and conclusions that can be drawn. Furthermore, as a technical paper it should highlight new methods and allow other researchers to replicate the procedure for their own work, in essence an experimental 'cookbook'. The method section is too brief to enable sufficient replication of the procedures (for example see the interactive

comment posted by Bijma), and is limited by the absence of (light or SEM) image. An image detailing where the software measured the average diameter would have been interesting (was it consistently the same length for the same species?). Researchers from a wide range of backgrounds, from those with simple micrometer lens to those with dedicated computer software packages, would benefit from an agreed dimension to measure for better comparison between work (see: Malmgren and Kennett, 1976 fig. 3).

The paper has no error bars in any of its graphs – yet the use of a mean, and various equipment (from the automated particle analysis to the scales, i.e. Schmidt, et al., 2004) should as a prerequisite require some indication of either the range (min. and max.) or error associated with analysis. How much does the method of mean diameter underestimate 'true' diameter? For that matter the use of the term diameter in this context should be replaced with average length, as diameter in both spherical and non-spherical forms is the largest distance between two points. How much does observed test area correspond to 'true' area were the authors to measure the diameter of each chamber? An expansion of the reasoning for major revisions is outlined below. In summary these include:

1. Individual test weights, rather than mean aliquot weight should have been compared with individual test size measurements, in order to obtain an "error free" data set linking individual shell size to individual shell mass.

2. Planktic Foraminifera are 3D objects and test size should be constrained in 3D (depth; width; length) rather than in one. Dimensions are independent of each other and elongation in one out of the three dimensions can erroneously increase the 'test size'.

3. The authors use test silhouette (2D) as a measure of test area (shape area) based on the traditional micropalaeonotological viewpoint to determine the test area. Has the asymmetrical nature of planktic foraminifera, the overlap and chamber ratio been

factored in?

4. The relationship between size and mass during ontogeny, in relation to increase in mass during Gametogenesis.

5. The use of the 200 – 250 μm limits comparison with previous publications and adapting this to a theoretical error model cannot be done with the data shown.

6. Information on the location of samples used to demonstrate and test the hypothesis is inadequately constrained (i.e. carbonate ion, temperature, season, year, differences in weight between depths and years, etc.).

7. Expansion of the Methodology section is required and inconsistencies in Graphs need clarification.

1. Individual test weights and test size measurements

The measurement of the mean diameter for an individual foraminifera averaged per aliquot against an average weight (per aliquot where n >10), limits the measurement of individual variation (the aim). In order to evaluate the two methodologies individual weight and size measurements (length; width; average chamber size) should have been taken into account. Not only to improve the accuracy but also to account for the large observed ranges (Fig.1 and 2). Measured diameter/area of tests can be affected by the final chamber – its location, size and form (see Malmgren and Kennett, 1976 who preferred width), with shape area being drastically effected depending on chamber overlap, angle of rotation of additional chambers (Olsson, 1973a); size of the last chamber in relation to penultimate chambers (referred to as Kummerforms and Normalforms, e.g. Olsson, 1972, 1973a, 1973b; Malmgren and Kennett, 1976). The ratio of chambers also affects surface area, whilst some chambers that are added may have greater visibility in one orientation than another. The variables connected to shape (measured as test area) do not just include size but also chamber overlap and rotation. The early work of Berger (1969), reprinted in Lipps (1979) and updated by Signes et

al. (1993), brought forward some key issues in respect to growth and size. Especially the importance of growth in regards to species that display allometric and isometric growth during ontogeny (Signes, et al., 1993). The current methodology (SBW) seen in previous publications (Barker and Elderfield, 2002; de Moel et al., 2009; Moy et al., 2009) is limited as correctly stated by the ability to constrain size, a better approach would have been to measure the thickness, weight, size and volume of individual test.

2. Planktic Foraminifera as asymmetrical 3D objects

Foraminifera have in the simplest 3D approximation three variable axises: width; height and; depth. As this paper aims to isolate changes in weight as a function of test thickness, from those of size variation then perhaps time should have been consumed in gaining volume measurements i.e. depth. A simple perusing of the literature shows that G. bulloides has subtle variations (diminutive final chambers, wide apertures, etc) and the centre based on the 'micropalaeontological' viewpoint can miss the 'largest' portion of test. See Spero and Parker (1985 Fig. 2a) who indicate that the maximum length of juvenile O. universa (trochospiral stage) should be measured over the largest chambers (the height). Furthermore, the mean test diameter of the specimens (300 μm) in comparison with sieve size (200-250 μm) should have been sufficiently quantified for publication; little has been done to satisfy this. The concept that there is a difference in 'sieve size' and 'measured size' is not new, e.g. Kroon and Darling (1995 – fig. 10), Lohmann (1995 – fig. 3). First order measurements, such as maximum test height and maximum test width should have been provided to corroborate this statement – not a statistic such as mean aliquot diameter. As only one dimension has been published (referred to as mean diameter), the authors cannot claim that sieving fails to remove the effect of size. In order for them to make this statement publication of data from at least two dimensions (perpendicular to each other, i.e. width and length, length and depth etc.), as the axis measured may be elongated in comparison with the other two dimensions. Species that have a 'bulbous' nature are hard to orient exactly in line with the camera (they may sit at a slight angle), therefore the use of mean diameter rather

than maximum diameter for a specimen is perplexing when many publications state that the max. diameter is least affected by random orientation (Renaud and Schmidt, 2003; Schmidt et al., 2004).

Further, the micropalaeontological orientation of specimens is not necessarily how they may pass through a sieve - rather passing through at the point with the smallest surface area. The realism of this therefore becomes important in understanding how a diameter can be larger than a sieve size. In regards to supplementary figure 2 (theoretical sieve based mean test area against observed measured mean test area) if the authors would have calculated that foraminifera are non-spherical and assumed that the minimum size is 200.5 μm and the maximum is 249.5 μm for 2 out of 3 axis then the grey shaded area, representing theoretical mean test area would actually sit higher, between 40,200.25 to 62,250.25 μm2 (the mid-point being 50,625 μm2) sufficiently covering most of the observations. Furthermore, the second part of supp. Fig. 2 (theoretical sieve based mean test diameter against observed measured mean test diameter) can be explained by using Pythagoras theorem, the largest length of an object passing through the centre is the sum of the squares of its two sides. For example, a sieve represents a square and a line that passing diagonally (through the centre) will be larger than its sides, for a sieve size of 250 μm the hypotenuse (diagonal length) is 353.55 μm. If you visualise this as a diamond pattern each side will be 250 μm but the dimensions running through the centre (both length and width) will be 353.55 μm.

There are also implications for using this method with planktic foraminifera species that have a complex three dimensional morphology, for instance G. sacculifer (with sac) or G. truncatulinoides where the four morphotypes are distinguished by changes in the third dimension, either flatter or more elongate?

3. Does test silhouette adequately determine the test area?

Additionally how much does the size of the aperture affect MBWarea, SEM images within the literature demonstrate the point that some G.bulloides have enlarged aper-

tures, vastly inflating the shape area (referred by the authors as 'test area' based upon the silhouette) whilst having no or little impact on weight (Malmgren and Kennett, 1976). The presence of kummerforms and normalforms may affect the measurement of test area and weight. How much of the true area of a 3D shape is gained from a planar silhouette? The 'measurement' of test area seems to have no correction for the curvature of Foraminifera test. The methodology used within this paper is a measurement more of the shape of planar silhouette than of true test area. de Vargas et al (2001), and Renaud and Schmidt (2003), suggest a method of making an approximation of the three-dimensional morphology through taking two views of specimens (edge view and spiral view). As a biproduct of that method size can then be calculated either through the zeroth harmonic of a Fourier transform analysis, as it is proportional to the size of each specimen or the maximum length.

4. Gametogenetic Calcite

The use of spineless spinose specimens, especially for plankton tows, would understandable lead to the presence of gametogenic calcite. As pointed out by Lohmann (1995) planktic foraminifera grow larger by adding chambers thus during ontogeny a correlation between shell size and mass should be prevalent. Secondary calcification, as for example during gametogenesis, involves the loss of spines, and a thickening of the walls without or very little increase in size, thus decoupling weight and size (see fig.2c Lohmann, 1995). Therefore samples that are measured with the presence of gametogenetic calcite are much more likely to have a diminished correlation between size and weight than those specimens with no presence of gametogenetic calcite.

5. The use of the 200 – 250 μm limits comparison with previous publications

The use of 200 – 250 μm size fraction whilst understandable in the need for a statistically viable result, is perplexing when previous workers have used >250 μm <425 μm (Barker and Elderfield, 2002; Moy et al., 2009; de Moel et al., 2009). Perhaps a range of sizes should have been measured especially in order to accurately predict the theoret-

ical errors (fig. 4). Lines 7 – 29 (page 911) the statement that the observed error in the measured fine size fraction is similar in magnitude to published observations regarding the carbonate ion effect and the only way to decrease this error is through the use of larger specimens (as previous publications do) which have less sieve-based error is a self-correcting argument . Furthermore, as no analysis is shown to have been conducted on larger specimens the comment that there is concern regarding data within previous publications is not justified. Had the authors measured larger sieve fractions then it would be suitable to infer that from their data, and include figure 4.

6. Information on the location of samples used to demonstrate and test the hypothesis

Whilst the most basic of information on the location of the samples was given how long they had been maintained in Formalin (4%) and hexamethyltetramine is not provided, Ganssen (1981) through experiments on Solnhofen Limestone has shown a drop of pH by 1 unit after just a 100hrs (whilst in that experiment the samples were unbuffered –unpublished data has shown that, even with buffering, prolonged time spent in solution has a detrimental affect on sample pH). Little experimental evidence exists for the assumption that samples would have remained unaffected. Was the pH of the samples tested before any experimental analysis conducted? Was there any difference between the four cruises (M12; M21; M26 and M36)? Were there any differences in the (x13) depth intervals utilised? Should the data be 'lumped' together which gives the impression that data points are comparable or more appropriately should not an indication of both depth and age of samples be shown (a simple alternation in symbols would have sufficed)? Whilst this is not the aim of the paper, it should be a prerequisite.

7. Methods and Graphs

As stated previously the methodology needs expansion in order to answer questions such as, once foraminifera were isolated from the sample with a pipette how were they dried? Were they wet/dry sieved? Why were samples not oven dried before weighing – a continuous measurement of the weight over a period of a minute should allow for

a 'back' calculation of original weight to be made (Peeters pers comm.), in regards to samples uptaking moisture. Additionally, why were the samples not treated for organic matter, even with a small density difference (3% contribution to individual mass) this may impact on mean aliquot test weight. Were individual weights recorded then 3% is acceptable, however as weights weren't individual this is an unknown error regarding the proportionality.

Within the diagrams, there are a number of inconsistencies including no trendline for figure 2, the appearance of a data point in fig. 2 of G.glutinata (in comparison with figure 1, at approximately 34,000μm2) thus increasing the range of mean test area. Further are these trendlines tied to the origin? In supplementary fig.2 as a suggestion shouldn't the minimum and maximum be plotted rather than the median and mean test size?

Are the r2 values in supplementary figure 3 for individual species or for all the species measured grouped together? As there are no trendlines and only four r2 values it is presumed the latter (one per graph), which leads to the question is it justifiable to determine the correlation between test area and diameter in this way? If the basis of the statement that there is a lack of correlation between measurement based weight and measured size, thus representing the methodology's removal of the influence of size on weight is based on this figure then clarification is needed. Grouping four different species, with presumably different ecologies, lifespan, vital effects (etc.) makes a lack of correlation more than understandable (i.e. Schmidt et al., 2003: "different adult sizes are mainly influenced by environmentally controlled growth rates since the lifespan is determined and reproduction triggered by the synodic lunar cycles")

Finally, is it reasonable to use mean test area (one of the measured variables) as a measurement of the accuracy of the different approaches, including determining the accuracy of mean test area normalized weight (in fig. 2 and supplementary fig. 3)?

In conclusion, my original intention was to reject this paper; however I believe Major

Revisions are all that are necessary. Were the authors to deal with a lot of the aforementioned comments it would become acceptable. The lack of first order data poses a key problem as the narrative of the paper skips to presenting normalized data, and the essential aim of this paper was what do you do once you have weight data. Do you factor in test size variability via a) measurement or b) a sieve mid-point size? Without the basic data it's difficult for the reader to evaluate the choices. Both size and weight measurements need to be refined in order to enhance the results and conclusions, expansion of both locality data (year, season, carbonate ion, and separation of data points plotted on the graphs into their retrospective samples) and methodology would allow for greatly improve its readability. The objectives of this paper are important for the modern understanding of size-weight relationships in Planktic Foraminifera and I personally look forward to an improved manuscript.

Additional References

Berger, W.H., 1969. Planktonic Foraminifera: Basic morphology and ecologic implications, Journal of Paleontology, 43 (6), 1369 - 1383

de Vargas, C., Renaud, S., Hilbrecht, H., and Pawlowski, J., 2001. Pleistocene adaptive radiation in Globorotalia truncatulinoides: genetic, morphologic, and environmental evidence, Paleobiology, 27 (1), 104 - 125

Ganssen G., 1981. Isotopic analysis of Foraminifera shells: Interference from chemical treatment, Palaeogeography, Palaeoclimatology, Palaeoecology, 33, 271 – 276

Kroon D., and Darling K., 1995. Size and upwelling control of the stable isotope composition of Neogloboquadrina dutertrei (d'Orbigny), Globigerinoides ruber (d'Orbigny) and Globigerina bulloides (d'Orbigny): Examples from the Panama Basin and Arabian Sea, Journal of Foraminiferal Research, 25 (1), 39 – 52

Lipps, J.H., 1979. The Ecology and Paleocology of Planktic Foraminifera, SEPM Short Course, 6, 62 - 104

Lohmann, G.P., 1995. A model for variation in the chemistry of Planktonic foraminifera due to secondary calcification and selective dissolution, Paleoceanography, 10 (3), 445 – 457

Malmgren, B., and Kennett, J.P., 1976. Biometric analysis of phenotypic variation in recent Globigerina bulloides (d'Orbigny) in the Southern Indian Ocean, Marine Micropaleontology, 1, 3 – 25

Olsson, R.K., -1973a. Growth studies on Globorotalia exilis blow and Globorotalia pertenius beard in the hole 154A section, Leg 15, Deep Sea Drilling Project, -1973b. Paleontological Notes - What is a Kummerform Planktonic Foraminifera, Journal of Paleontology, 47 (2), 327 – 329

Renaud S., and Schmidt, D.N., 2003. Habitat tracking as a response of the planktic foraminifer Globorotalia truncatulinoides to environmental fluctuations during the last 140 kyr, Marine Micropaleontology, 49, 97 – 122

Schmidt D.N., Renaud S., and Bollmann J., 2003. Response of planktic foraminiferal size to late Quaternary climate change, Paleoceanography, 18 (2), 1039

Schmidt, D.N., Thierstein, H.R., and Bollmann, J., 2004. The evolutionary history of size variation of planktic foraminiferal assemblages in the Cenozoic, Palaeogeography, Palaeoclimatology, Palaeoecology, 212, 159 - 180

Signes, M., Bijma, J., Hemleben, C., and Ott, R., 1993. A model for planktic foraminiferal shell growth, Paleobiology, 19 (1), 71 – 91

Spero H.J., and Parker S.L., 1985. Photosynthesis in the symbiotic Planktonic foraminifer Orbulina universa, and its potential contribution to oceanic primary production, Journal of Foraminiferal Research, 15 (4), 273 - 281

Interactive comment on Biogeosciences Discuss., 7, 905, 2010.