Analysis of fossil planktonic foraminifera: the sieve mesh effect.

V. DI DONATO[1], J.A. MARTÍN-FERNÁNDEZ[2], J. DAUNIS-i-ESTADELLA[2] and P. ESPOSITO[1] ,
[1] Dipartimento di Scienze della Terra - Università degli Studi di Napoli “Federico II”, Italy, valedido@unina.it
[2]Dept. Informàtica i Matemàtica Aplicada - UdG, Spain

1. Introduction

The choice of the sediment size fraction in the analysis of fossil planktonic foraminifera is of great importance in determining the composition of assemblages. In past studies several size fractions have been utilised. Imbrie and Kipp (1971) stated that “smaller fractions give rise to too many uncertainties in the identification of small specimens and require too long to process. Large mesh size yield undesiderable loss of small species, and small specimens of larger species”. The adoption of coarser sieve meshes has the effect to reduce the percentages of small sized (usually living in cold waters) species (Bé and Hutson, 1977). In a comparison between >63 μm and >150 μm size-fraction of planktonic foraminifera assemblages from NW Atlantic Ocean, Smart (2002), stated that because particular smaller species are either under-represented or even absent from the larger (>150 μm) size-fraction, the smaller (>63 μm) size-fraction must be included in studies of planktonic foraminifera. Di Donato et al. (2008) highlighted in the >150 μm size fraction of a core from the Tyrrhenian Sea, a great loss of the small-sized species *Turborotalita quinqueloba* in glacial samples, where this species is very abundant. This caused an increase in warm water species such as *Globigerinoides ruber*, apparently reaching typical Holocene values in some full glacial levels.
From the above mentioned statements, it is clear that treatment changes can strongly influence the results of palaeoclimatic reconstructions based on planktonic foraminiferal assemblages.

As stated by Aitchison (1986, 1992) scale invariance and subcompositional coherence are fundamental properties of the compositional data analysis (CODA). The main goal of this paper is to verify if the variable relationships pointing out from CODA of foraminiferal assemblages are or not influenced by the size fraction and, more in general, to test the robustness of CODA respect to treatment changes related to different preparation techniques for the analysis of planktonic foraminifera.

2. Material and methods

Core GNS84-C106 has been recovered in the Tyrrhenian Sea. Its stratigraphical record includes sediments ranging from the Last glacial period to the present (Buccheri et al., 2002). The planktonic foraminiferal assemblages has been firstly analysed on the >106 μm size fraction (Buccheri et al., 2002). Later the whole core was re-analysed on the >150 μm size fraction in order to reconstruct sea surface temperatures through modern analogues (Di Donato et al. 2008), a technique requiring the comparison of fossil assemblages with modern ones. Thus the comparison between the compositions obtained with these two treatments can be made on a dataset consisting of pairwise observations corresponding to 58 different stratigraphical levels.

Following Daunis-i-Estadella et al. (2008), the zero replacement which is required by the CODA has been carried out by adopting a Bayesian approach to the zero replacement, based on a posterior estimation of the parameter of the multinomial distribution with Jeffreys and Uniform priori, combined with a multiplicative substitution.

In order to evaluate the significance of the differences among the obtained compositions, it has been analysed if the average perturbation difference vector is the neutral vector of not, from two different point of view: contrast and bootstrap. In the contrast based approach: it has been tested by means of a Hotelling's one sample T2-test the null hypothesis that the average perturbation difference vector is equal to neutral vector. This is equivalent to contrast if the ilr-transformation vector is equal to vector of zeros.

In the bootstrap approach a one thousand resampling from the parwise perturbation data set has been made, after which the average perturbation difference vector of each of the 1000 data sets have been calculated. For each part the 5% and the 95% percentiles have been calculated so as to compare it with the neutral perturbation e=[1/D,…,1/D]..

As an exploratory analysis CODA biplots (Aitchison and Greenacre, 2002) have been carried out to evaluate the differences in the covariance structure of data obtained by the analysis of the two size fractions. Moreover a Mantel test for similarity has been applied to the covariance matrices obtained from ilr-trasformed data. The test works by permutating the rows and columns of the two matrices randomly. It returns a P-value corresponding to the quantile of the observed statistic in the permutation distribution. The null hypothesis states that the correlation matrices are not related.

Statistical methods required have been applied from packages ade4, ICSNP and MASS (R Development Core Team, 2010).

3. Results

As regards the evaluation of the differences in the compositions obtained with the >106 µm and the >150 µm size fractions, the performed Hotelling's one sample T2-test (T.2 = 135.0615, df1 = 11, df2 = 47) yielded a P-value < 2.2e-16, allowing to reject the null hypothesis of no difference of the ilr-transformation vector from a vector of zeros. In other terms this result indicates that it exist a statistically significant perturbation vector that moves the >106 µm size fraction assemblages to the >150 µm ones.

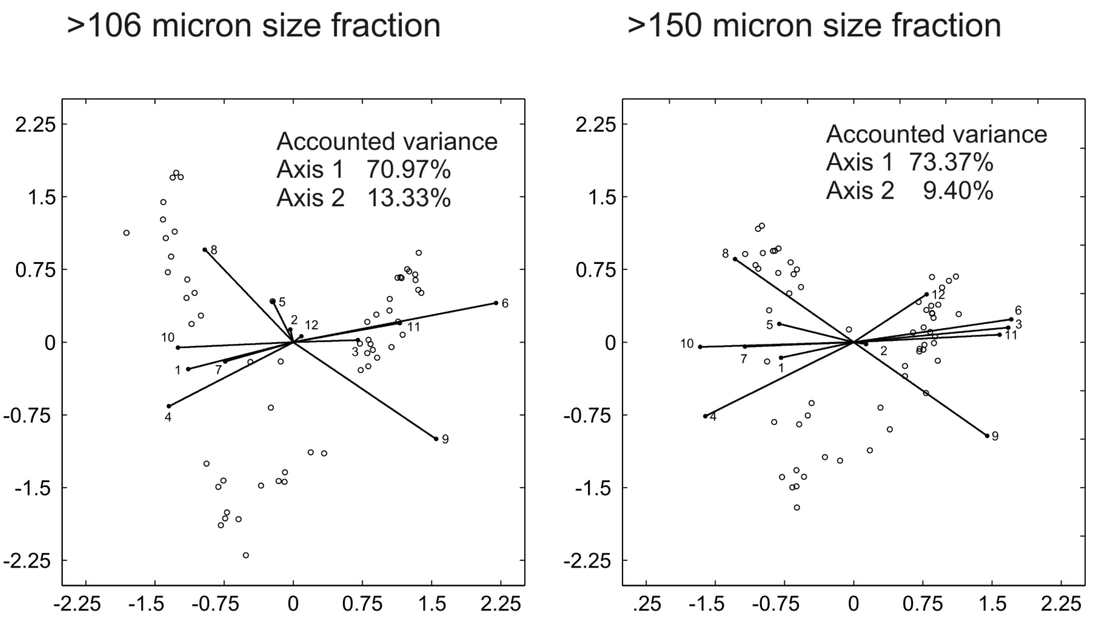

Figure 1. Covariance CODA biplots computed from >106 and >150 size fraction planktonic foraminiferal assemblages of core GNS84-C106. 1) *Globigerinoides ruber* 2) *Globigerina bulloides* 3) *Globigerinita glutinata* 4) *Globorotalia inflata* 5) *Globigerinoides sacculifer* 6) *Globorotalia scitula* 7) Warm species group (including *Globigerinita glutinata* and *Globoturborotalita* spp.) 8) *Globorotalia truncatulinoides* 9) *Neogloboquadrina* spp. 10) *Orbulina universa* 11) *Turborotalita quinqueloba* 12) undetermined.

The covariance CODA biplots computed from >106 µm and >150 µm size fractions are shown in fig. 1. The main difference is represented by the different amount of total variability accounted by some of the planktonic foraminifera species, as shown by the changes in the length of vectors. In particular it can be noted that *Globigerinoides ruber,* the most abundant warm water species throughout the core, as well as *Globorotalia scitula,* has a slightly reduced variability in the >150 µm size fraction. An opposite behaviour is shown by *Globigerinita glutinata* and the undeterminables specimens, with an increased variability in the >150 µm size fraction. It can be noted, however, that angular relationships among column vectors obtained from >150 µm and >106 µm size fraction are substantially similar. In particular, whatever the treatment, the first axis shows an opposite relationship between warm and cold species, allowing the first component to be interpreted in terms of sea surface temperatures changes. It can be also noted in both cases an evident opposition between *Globorotalia truncatulinoides* and *Neogloboquadrina* spp., the length of the corresponding link indicating a large standard deviation of their logratio. It is noteworthy the location of the apex of *Globigerina bulloides* column point near the origin of the biplots independently from the treatment, meaning that this species, one of the most abundant throughout the core, is rather invariant respect to the compositional changes occurring from glacial to Holocene.

So as to compare the ilr-covariance matrices a MBox-test was performed. The resulting P-value was greather than 0.9 (xi-square=222.1508; df=66) suggesting that it can be assumed that no difference exist between both matrices. Nevertheless when the normality assumption was checked both produced P-values (multivariate Shapiro-test) were below 0.012. Changing the strategy, we concentrated to evaluate the degree of association between both matrices. The performed Mantel test for evaluating the similarity of the correlation matrices obtained from the two different size fractions (fig. 2) gave a P-value of approximatively 0.01, allowing to reject the null hypothesis that the correlation matrices are not related. This means that the compositional differences between >106 and >150 µm size fractions highlighted by the Hotelling's one sample T2-test does not imply strong changes in the covariance structure.

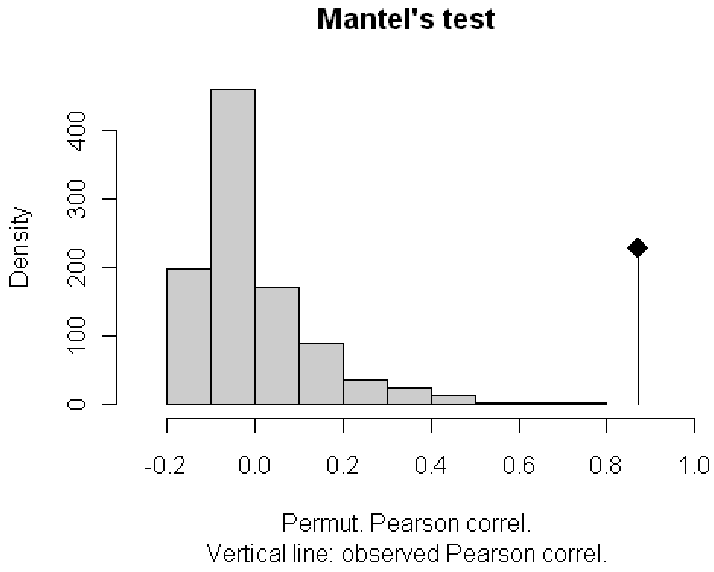

Figure 2. Mantel test of the similarity of the correlation matrices obtained from >106 and >150 µm size fractions planktonic foraminiferal assemblages of core GNS84-C106

The associated variability in our analysis was checked via the Bootstrap methods (b=1000). As regards to the average perturbation-difference vector the 5% and 95% percentiles of each part were calculated. Figure 3 shows that only three parts (x_8, x_9 and $x_{10,}$ corresponding to *Globorotalia truncatulinoides, Neogloboquadrina* spp. and *Orbulina universa*) are not affect by a perturbation. The rest of parts receive a perturbation between the two cores.

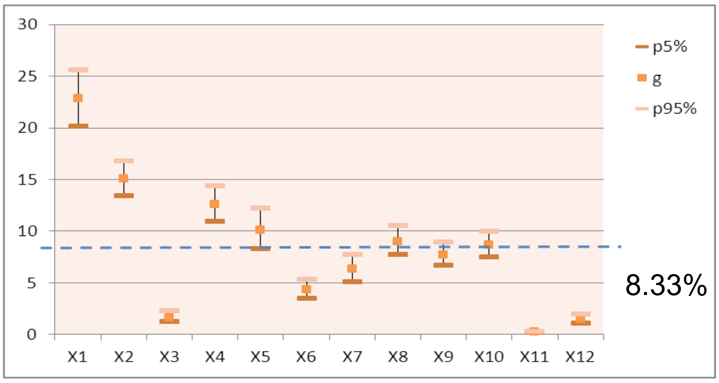

Figure 3. Perturbation-difference vector in the Bootstrap analysis. Horizontal line stands for the neutral-difference 1/D value. Percentile interval is plotted as vertical-bars. x1) *Globigerinoides ruber* x2) *Globigerina bulloides* x3) *Globigerinita glutinata* x4) *Globorotalia inflata* x5) *Globigerinoides sacculifer* x6) *Globorotalia scitula* x7) Warm species group (including *Globigerinita glutinata* and *Globoturborotalita* spp.) x8) *Globorotalia truncatulinoides* x9) *Neogloboquadrina* spp. x10) *Orbulina universa* x11) *Turborotalita quinqueloba* x12) undetermined.

Concerning the comparison of the correlation matrices, Pearson correlation coefficient between both ilr-covariances matrices across the Bootstrap analysis was computed. Figure 4 shows that this coefficient mostly varies in a range [0.75, 0.95] and gives support to the association between both covariance structures.

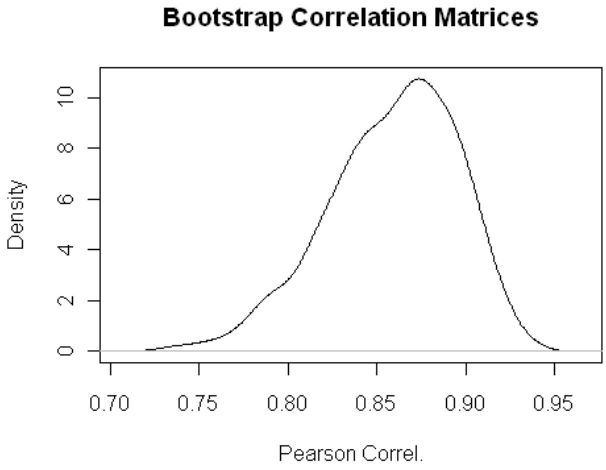

Figure 4. Kernel density estimation of Pearson correlation coefficient in the Bootstrap analysis.

The obtained results evidence on the whole the robustness of covariance structure of CODA respect to sieve mesh changes in the analysis of planktonic foraminiferal assemblages. This behaviour is not unexpected, if one consider that the shifting from a small size fraction to a larger one implies a decrease in the abundance of small sized species acting as they would be multiplied by a reducing factor. The scale invariance and subcompositional coherence properties of CODA techniques result in a robustness of covariance structure which cannot be achieved in raw percentage data analysis.

References

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data.* Monographs on Statistics and Applied Probability. Chapman & Hall Ltd., London (UK). (Reprinted in 2003 with additional material by The Blackburn Press). 416 p.

Aitchison, J. (1992). *On Criteria for Measures of Compositional Difference.* Mathematical Geology, vol. 24, No. 4, pp. 365-379.

Aitchison, J., Greenacre, M. (2002). *Biplots of compositional data*. Applied Statistics 51 (4): 375–392.

Bé, A.W. H., Hutson, W. H. (1977). Ecology of planktonic foraminifera and biogeographic patterns of life and fossil assemblages in the Indian Ocean, Micropaleontology, 23 (4), 369-414.

Buccheri, G, Capretto, G, Di Donato, V, Esposito, P, Ferruzza, G, Pescatore, T, Russo Ermolli, E, Senatore, M.R., Sprovieri, M, Bertoldo, M, Carella, D, Madonia, G. (2002). *A high resolution record of the last deglaciation in the southern Tyrrhenian Sea: environmental and climatic evolution*. Marine Geology 186: 447-470.

Daunis-i-Estadella, J., Martín-Fernández, J.A., Palarea-Albaladejo, J. (2008). Bayesian tools for count zeros in compositional data sets. *In: Daunis-i-Estadella, J. and Martín-Fernández, J.A. (Eds.), Proceedings of CODAWORK'08, The 3rd Compositional Data Analysis Workshop, May 27-30, University of Girona, Girona (Spain)*, CD-ROM (ISBN: 978-84-8458-272-4).

Di Donato, V., Esposito, P., Russo-Ermolli, E, Scarano, A., Cheddadi, R. (2008). *Coupled atmospheric and marine palaeoclimatic reconstruction for the last 35 kyr in the Sele Plain-Gulf of Salerno area (southern Italy).* Quaternary International 190: 146–157.

Imbrie, J. and Kipp, N. (1971). *A new micropaleontological method for quantitative paleoclimatology: application to a late pleistocene caribbean core*. In: The late Cenozoic glacial ages, pages 71-181. K. K. Turekian, Yale Univ. Press, New Haven.

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.