Contents lists available at ScienceDirect

Deep-Sea Research II

journal homepage: www.elsevier.com/locate/dsr2

Meiofaunal community structure of the deep-sea Gulf of Mexico: Variability due to the sorting methods

Elva G. Escobar-Briones [a],[*], Citlali Díaz [b], Pierre Legendre [c]

[a] Universidad Nacional Autónoma de México, Instituto de Ciencias del Mar y Limnología, Unidad Académica Ecología Marina,
A.P. 70-305 Ciudad Universitaria, 04510 México, D.F., Mexico
[b] Universidad Nacional Autónoma de México, Posgrado en Ciencias del Mar y Limnología, Ciudad Universitaria, 04510 México, D.F., Mexico
[c] Département de sciences biologiques, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec, Canada H3C 3J7

ARTICLE INFO

ABSTRACT

Studies designed to describe the structure of soft bottom meiobenthic communities have used diverse sorting methods, thus making the comparison of results difficult. This study compared sorting specimens using the manual and centrifuge density gradient sorting methods, for two deep-sea depth zones [continental slope (1630–1860 m) and abyssal plain (3720–3830 m)] in the southwestern Gulf of Mexico. The manual sorting method produced more taxonomic groups and higher abundance and biomass values, compared to the differential density Ludox-centrifuge sorting method. The results obtained from the two sorting methods suggest that distribution patterns of deep-sea meiofaunal samples can be an artefacts of the sorting.

1. Introduction

Meiofaunal studies have used different sorting techniques making it difficult to compare results. Extraction of deep-sea meiofaunal specimens from fine grain-size sediments (clay, silt) is routinely accomplished by bubbling, sieving and examining the residue, followed by centrifugation in density gradients and floatation (Higgins and Thiel, 1988; Burgess, 2001). Centrifuge density gradient sorting assumes that the density of the specimens is close to $1.15\,g\,ml^{-1}$, while the density of silt and clay is larger. The density of the solution used to create the gradient is larger than the density of most specimens but slightly less than sediment density, allowing the specimens to float and remain in the supernatant after centrifugation. The Ludox-centrifuge used to sort meiofauna has a density range of 1.20–$1.40\,g\,ml^{-1}$.

Other protocols using isopycnic methods have produced large differences in the efficiency of extraction in different sediment types. These differences are the result of the physical interaction between sediment, pore space and the specimens (de Jonge and Bouwman, 1977; Nichols 1979; Barnett 1980). Some authors have suggested a reduction in time when using the centrifuge extraction method on meiofaunal samples from muddy sediment or sediment containing a lot of detritus (de Jonge, 1979; Alongi, 1986; Burgess, 2001). The extraction efficiencies recorded are larger than 80% in shallow-water samples; however, other authors have suggested large losses of specimens during all steps of the centrifuge method (Schwinghamer, 1981; Bloem et al., 1986). An additional disadvantage of the centrifuge method is the small size of the processed samples ($< 10\,ml$) due to the maximum capacity of the centrifuge tube size ($50\,ml$). Care must be taken to maintain the stability of the Ludox-AM® density; a column where organisms can be retained in the supernatant must use a density gradient with a ratio of 1:3 of Ludox (L. Hyde, pers. comm.).

The issue of sorting techniques originated because meiofauna is similar in size to the average grain size if it inhabits sand (Higgins and Thiel, 1988). However, manually sorting mud (silt and clay) or sediment rich in organic matter requires extensive time at the microscope (Heip et al., 1974; de Jonge, 1979). This inefficiency has also been considered, in addition to precision and accuracy. This study compares the efficiency of two sorting methods for meiofaunal samples from a subtropical deep-sea basin to establish quality criteria contributing to extractions of the fauna from abyssal sediment at low latitudes.

1.1. Study area

The study area located in the southwestern sector of the Gulf of Mexico includes the continental slope and the SIGSBEE abyssal plain (Fig. 1), the greatest depth in the Gulf of Mexico (3900 m). Sediment in the northern portion of the abyssal plain is of continental origin contributed by the Mississippi River discharge; the carbonate content is less than 25% (Bouma, 1972). Sediment in

* Corresponding author. Tel.: +52 55 5622 5841; fax: +52 55 5616 0748.
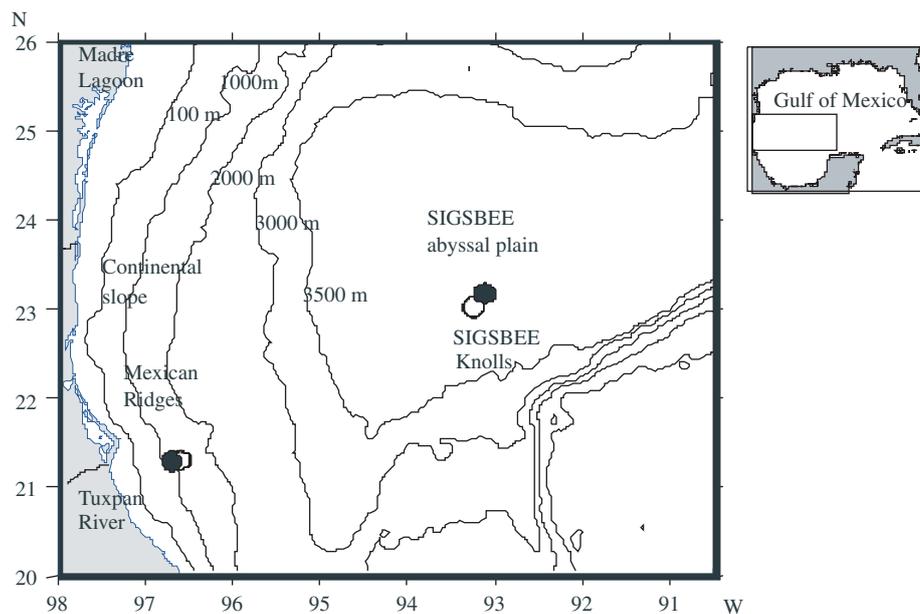E-mail address: escobri@mar.icmyl.unam.mx (E.G. Escobar-Briones).

Fig. 1. Study area and locations of the sampling stations. White markers: SIGSBEE4 cruise; solid black markers: SIGSBEE5 cruise.

the central and southern portions of the abyssal plain is hemipelagic in origin; it is mostly composed of pelagic foraminiferan shells.

2. Materials and methods

2.1. Field work

Meiofauna samples were collected from the sediment from the continental slope (1630–1860 m) and the abyssal plain (3405–3830 m) using a 0.25 m^2 US-NEL box corer (Fig. 1), during the June 2001 SIGSBEE4 and July 2002 SIGSBEE5 cruises, following the protocol described by Hessler and Jumars (1974).

At each station during each cruise, six 6.6-cm^2 meiofaunal replicates (subcores) were taken using 50-ml, 2.89-cm diameter hand corers with 5 cm penetration. Additional subcores were taken to analyse associated sediment variables: grain size, chloroplastic pigment concentration, and elemental organic carbon and nitrogen. The subcores for pigments and elemental carbon and nitrogen were frozen to −20 °C onboard the vessel and analysed later in the laboratory. Onboard the ship, the meiofaunal subcores were separated from the finest sediment by washing them through a sieve with 42-μm mesh size after exposure to a freshwater shock. Subcores were fixed with 95% ethanol and Rose Bengal vital stain and stored at −20 °C in 75-ml jars.

2.2. Sample processing

Each meiofaunal subcore was hand-sorted first using a stereoscopic microscope with 5× magnification. The subsamples of sediment were placed in a Petri dish and the specimens were extracted with fine-tip forceps; 14 major taxonomic groups were identified. The time required to sort each subcore was used as an estimate of efficiency, in addition to the total number of taxa, abundance, and biomass of the sorted specimens. The additional time required for sample preparation for the centrifuge sorting was taken into account in the comparison of methods.

The sorted meiofaunal specimens were put back into the sediment and homogenised gently. A centrifuge density gradient was created using Ludox-AM® to sort the specimens anew. The nematodes and harpacticoid copepods were extracted with the centrifuge strategy placed in a loop slide (4 × 3 loops per slide) with a glycerine drop for digital photography with a stereoscopic Zeiss microscope with a 10/8 objective. These images were used to calculate biomasses based on biovolumes, following the protocols described in Feller and Warwick (1979) and Warwick and Price (1979). Biovolumes were determined following the protocol described in Baguley et al. (2004, 2006); our measurements are reported in microns. Biomasses were estimated following the protocols described in Warwick and Gee (1984) using taxon-specific conversion factors.

2.3. Data analyses

The Sørensen similarity index (Eq. (1)) was computed to estimate the similarity between subcores, within each depth zone and between depth zones:

$$CCs = 2(N_{common})/(N1 + N2) \tag{1}$$

where N_{common} = number of taxonomic groups common to two subcores, $N1$ = number of groups found in subcore 1, and $N2$ = number of groups found in subcore 2 (Sørensen, 1948; Legendre and Legendre, 1998).

For the taxa presence–absence, abundance, and biomass data tables, multivariate analyses of variance for two crossed factors (sorting methods and depth zones) were carried out by redundancy analysis (RDA). The abundance and biomass data for the 15 taxonomic groups were square-root transformed prior to RDA analysis in Canoco version 4.5 (ter Braak and Smilauer, 2002). All tests of significance were permutation tests involving 9999 permutations of the residuals under the reduced model in the program Canoco.

Extraction efficiency of the centrifuge sorting method (CSM) was estimated with the following equation:

$$E_{Ef} = Ind_{CSM}/Ind_{Total} \tag{2}$$

where E_{Ef} is the extraction efficiency index, Ind_{CSM} is the number of organisms found in the Ludox supernatant, and Ind_{Total} is the

sum of organisms remaining in the sediment plus those found in the Ludox supernatant. Since manual sorting provided the initial counts, the initial count values were considered to represent the total abundances of organisms and were the baseline to compare the efficiency of the two sorting methods.

The time required to extract the specimens, the taxonomic richness (total number of taxa), and the total abundance were also tabulated.

3. Results

The environmental characterization for the locations on the continental slope and in the abyssal plain in the two SIGSBEE cruises is presented in Table 1.

Table 1
Sampling stations and environmental factors in the study area

	Continental slope		Abyssal plain	
	SIGSBEE 4	SIGSBEE 5	SIGSBEE 4	SIGSBEE 5
Latitude (N)	21°19.09	21°21.1846	23°12.30	23°04.0635
Longitude (W)	96°25.19	96°21.4864	92°50.69	93°00.4049
Depth (m)	1630	1860	3830	3720
T (°C)	5	5	4	4
Salinity (psu)	35.05	35.02	35.43	35.40
% Clay	92.08	93.22	94.73	98.29
DO (mg L^{-1})	5.44	4.4	5.94	4.2
% o.m.	0.22	0.24	0.22	0.27
% C_{org}	0.90	0.96	0.90	1.09
% N	0.08	0.09	0.08	0.14
C/N	11.25	10.67	11.25	7.79
Chl-a ($\mu g\,cm^{-2}$)	0.77	0.51	1.24	0.60

Abbreviations: psu: practical salinity units; DO: dissolved oxygen concentration in water; %o.m.: organic matter content in sediment; %C_{org}: organic carbon in sediment; %N: nitrogen in sediment; Chl-a: chloroplastic pigments in sediment.
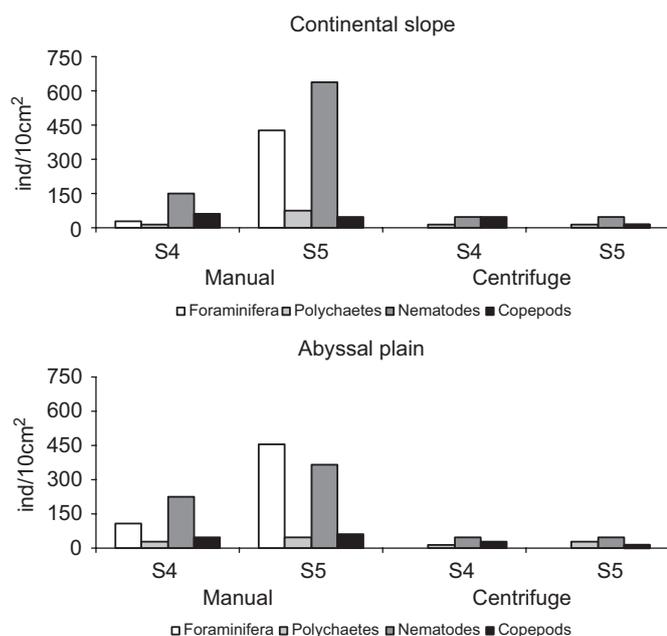
3.1. Taxonomic composition

A total of 8 phyla were recognized when using the manual sorting method and included a total of 14 invertebrate taxonomic groups. The centrifuge sorting strategy missed several phyla in the samples from both cruises and under-represented the abundance of the dominant taxa (Fig. 2).

Nematodes were the dominant taxon in both depth zones, irrespective to the sorting method used, with exception of samples collected on the continental slope during cruise SIGSBEE4 and sorted with the CSM (Fig. 2). Harpacticoid copepods and foraminiferida were next in abundance. Differences in the mean abundance values, recorded for each depth zone, cruise, and sorting method illustrate the inefficiency of the CSM (Fig. 2).

The Sørensen similarity index (CCs) between subcores from the continental slope in the SIGSBEE4 cruise, computed for the manually sorted data, was quite variable; values ranged from 0.33 to 0.86 ($n = 15$, Fig. 3). The similarity was higher and less variable in values computed from the SIGSBEE5 cruise and among subcores from the abyssal plain; the mean value was 0.75 ± 0.12 ($n = 15$, Fig. 3).

3.2. Taxonomic richness

Of the 14 taxonomic groups recognized, 11 were recorded in subcores from the continental slope and 11 in subcores from the abyssal plain. The average number of taxonomic groups recorded with the CSM was 2.16 ± 0.93 ($n = 12$) for the continental slope and 2.41 ± 0.79 ($n = 12$) in the abyssal plain. Using the manual sorting method, a mean value of 4.41 ± 1.08 taxa was recorded in

Fig. 2. Mean abundances (ind. $10\,cm^{-2}$) of the dominant meiofaunal taxa from the SIGSBEE4 and SIGSBEE5 cruises using the manual and the centrifuge sorting methods in the two depth zones. Upper panel: continental slope. Lower panel: abyssal plain. Abbreviations: For: foraminiferida; Pol: polychaeta; Nem: nematoda; Cop: harpacticoid copepods.
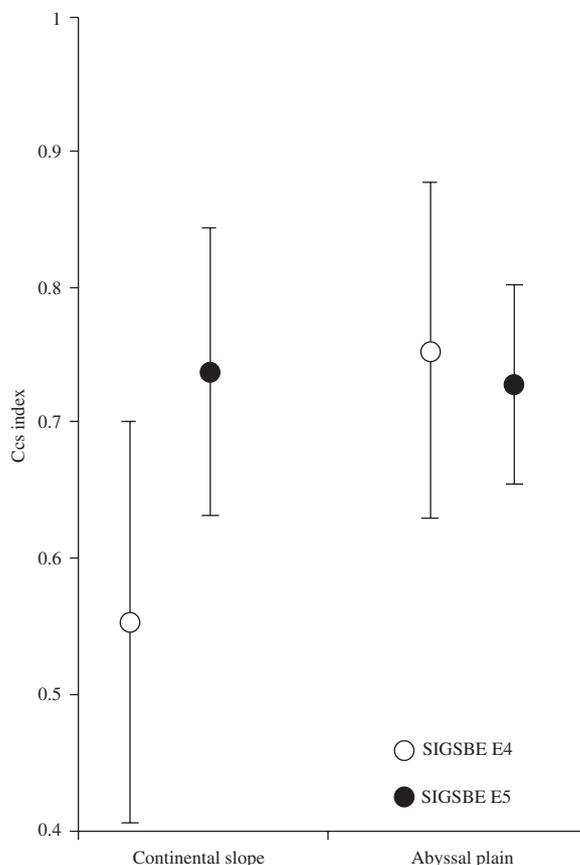
Fig. 3. Means and standard deviations of the similarities among subcores, computed from the manually sorted data, estimated using the Sørensen similarity index (CCs), for the two depth zones in cruises SIGSBEE4 and SIGSBEE5.
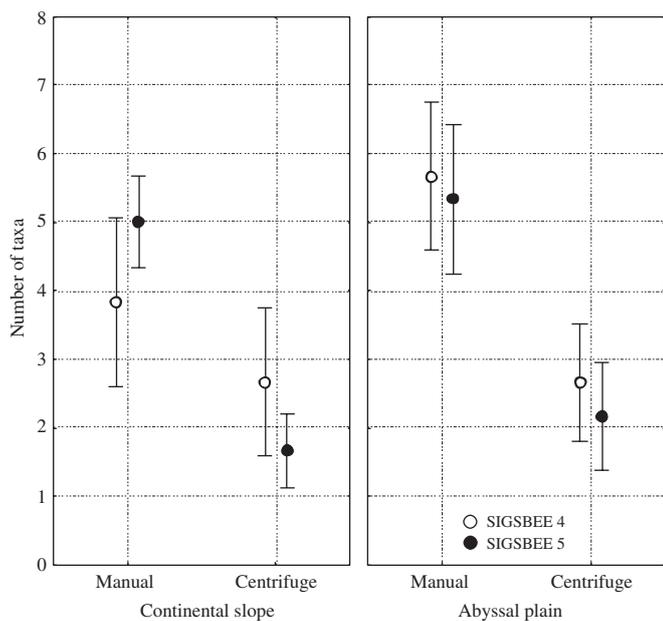
Fig. 4. Taxonomic richness (number of taxa) of the meiofauna in the two cruises and the two depth zones using the two sorting methods, manual and centrifuge. The vertical bars represent 95% confidence intervals about the means (symbols).

the two cruises SIGSBEE4 and SIGSBEE5 in the continental slope samples ($n = 12$); the mean value was larger in the abyssal plain with 5.50 ± 1.00 taxa ($n = 12$). Differences were significant between sorting methods for the two depth zones during both cruises (Fig. 4). In all cases, manual sorting produced more taxa than centrifuge sorting.

3.3. Abundance

Higher abundance values were recorded with manual sorting (continental slope mean abundance values 16 ± 11 ind. $6.6 \, \mathrm{cm}^{-2}$ and 78 ± 10 ind. $6.6 \, \mathrm{cm}^{-2}$ for cruises SIGSBEE4 and SIGSBEE5, respectively; abyssal plain mean abundance values 28.5 ± 9 ind. $6.6 \, \mathrm{cm}^{-2}$ and 59 ± 8 ind. $6.6 \, \mathrm{cm}^{-2}$ for cruises SIGSBEE4 and SIGSBEE5, respectively) than with centrifuge sorting (1–19 specimens per subcore) which in most cases showed significant differences (Newman–Keuls $p = 0.03$ for the continental slope; $p = 0.0001$ for the abyssal plain; Fig. 5).

The abundance values displayed a similar pattern in the two cruises. When subcores were sorted using the centrifuge method, the mean values were slightly higher in the continental slope and lower in the abyssal plain. Higher values were recorded in the SIGSBEE4 cruise in comparison with the SIGSBEE5 cruise (Fig. 5). This was not the case with the samples sorted manually. Values from one cruise to the other were significantly different and the patterns with depth varied, decreasing in the SIGSBEE5 cruise and increasing in the SIGSBEE4 cruise (Fig. 5). The differences were significant with the two sorting methods in the two depth zones (Newman–Keuls $p = 0.0002$ in the continental slope; and $p = 0.0001$ in the abyssal plain).

3.4. Biomass

The biomass values recorded in samples from the two depth zones using the CSM were lower than biomass values recorded from samples sorted by hand. The differences recorded for biomass values between sorting methods were significant in samples from the abyssal zone (Newman–Keuls $p < 0.05$). The difference between sorting methods in samples from the
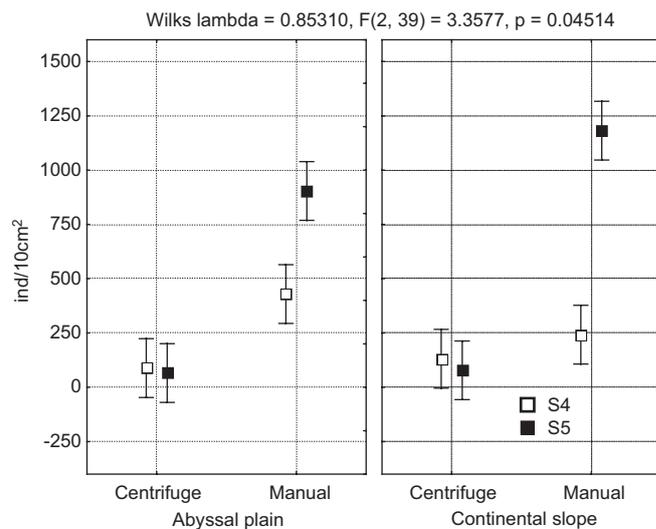

Fig. 5. Total abundance values (ind. $10 \, \mathrm{cm}^{-2}$) of meiofaunal samples recorded in the SIGSBEE4 and SIGSBEE5 cruises using the two sorting methods, manual and centrifuge, in the two depth zones. The vertical bars represent 95% confidence intervals about the means (symbols).

Fig. 6. Total biomass values of meiofaunal samples recoded in the two SIGSBEE cruises using the two sorting methods, manual and centrifuge, in the two depth zones. The vertical bars represent 95% confidence intervals about the means (symbols).
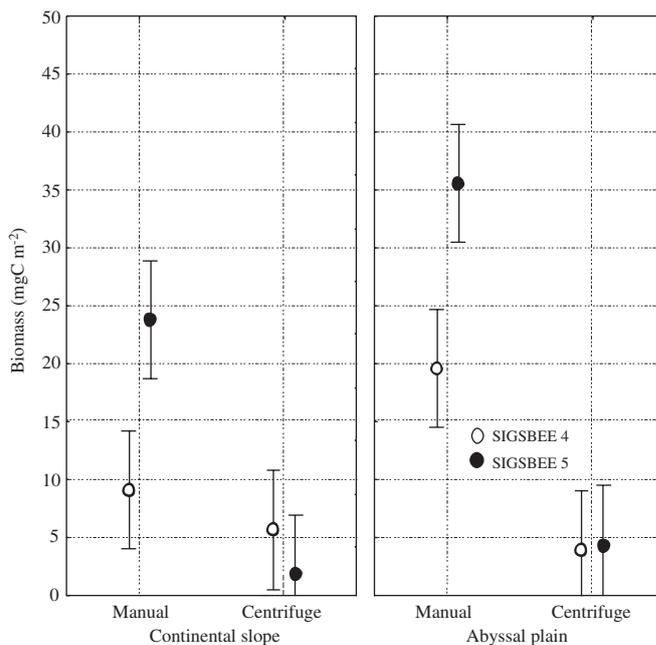
continental slope was significant only in the SIGSBEE5 cruise (Newman–Keuls $p < 0.05$).

The biomass values recorded from the continental slope were low (Fig. 6; 3.80–15.6 mgC m^{-2}); the average value was 9.1 ± 5.08 mgC m^{-2} ($n = 6$) for the SIGSBEE4 cruise. In contrast, the values obtained for samples from the abyssal plain were higher (Fig. 6; 13.4–33.3 mgC m^{-2}), with an average value of 19.5 ± 7.2 mgC m^{-2} ($n = 6$). The biomass values from subcores processed using the CSM and collected in the abyssal plain were lower (0.56–12.1 mgC m^{-2}) than those obtained by manual sorting; the average value was 4.1 ± 4.2 mgC m^{-2} ($n = 6$) in the SIGSBEE4 cruise. The biomass values from samples obtained in the

(A) Presence-absence data
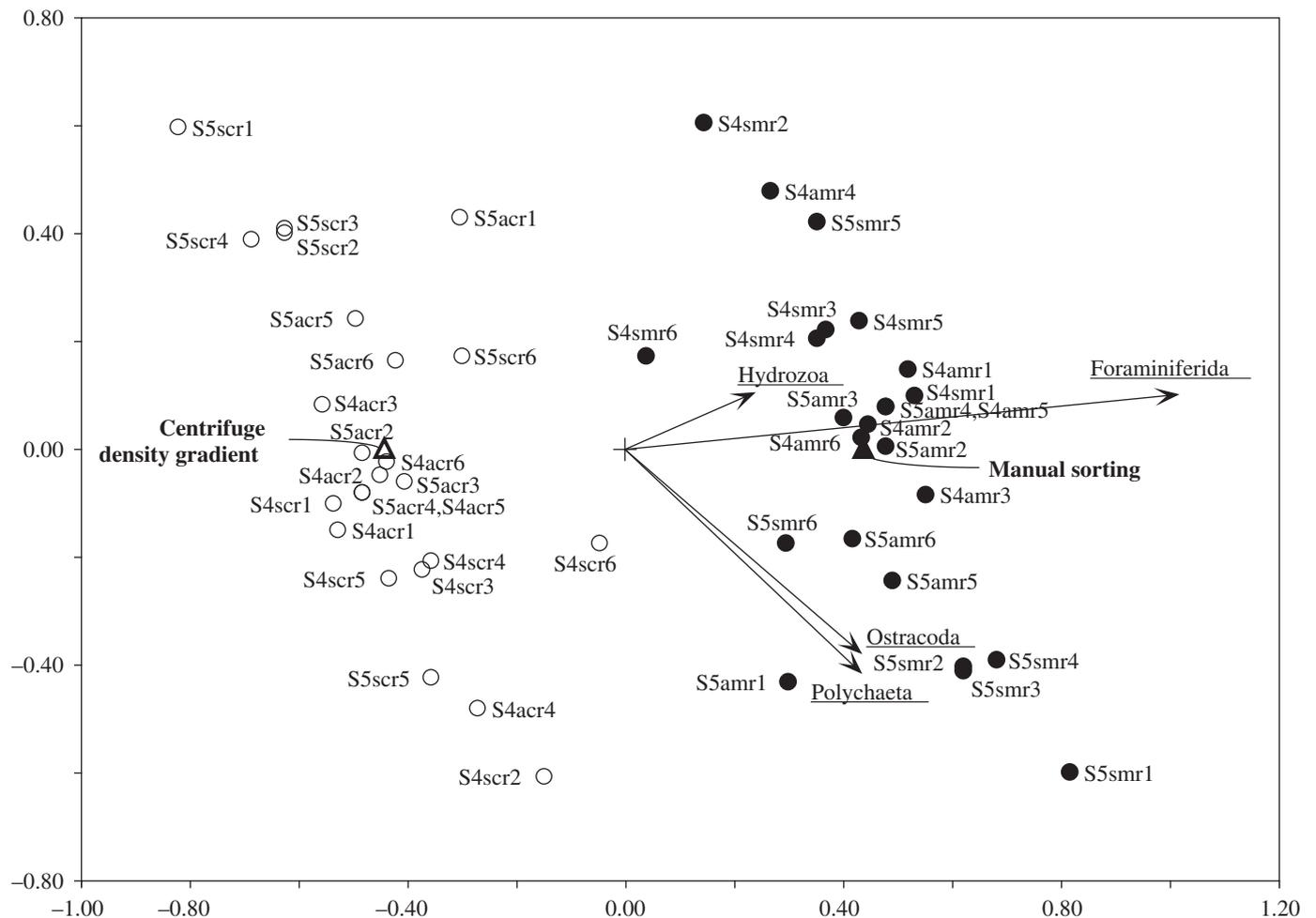
(B) Abundance data (square-rooted)
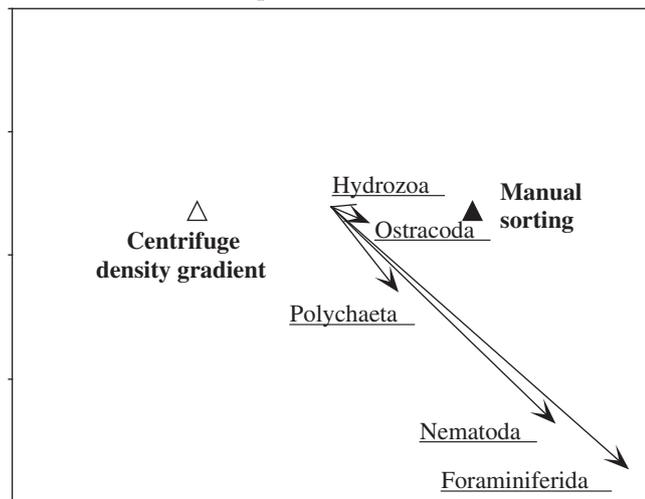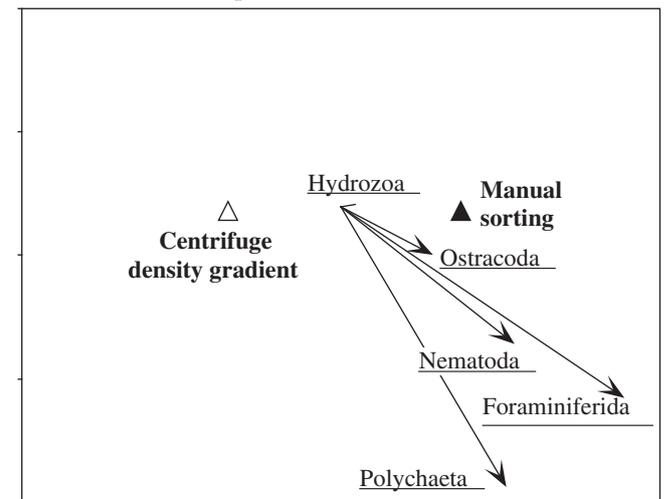

(C) Biomass data (square-rooted)

Fig. 7. RDA biplots illustrating the results of the multivariate Anovas for differences between sorting methods. The triangles in the biplots display the centroids of the subcores for each of the two sorting methods. The arrows represent the five dominant taxa; they were fitted at more than 10% by the single canonical axis (abscissa). Species arrow lengths were divided by 3 in (A) and by 2 in (B) and (C) for clarity of the graphs. Nematoda is not well fitted at all by the canonical axis in the presence–absence analysis (fit = 0%) because nematodes were found in every subcore. Subcores are identified by circles in (A); the filled circles on the right represent manual sorting results; the empty circles on the left are centrifuge density gradient sorting results. Subcores are identified by a code with four elements: cruise (S4 = SIGSBEE4, S5 = SIGSBEE5); zone (s = continental slope, a = abyssal plain); sorting method (m = manual, c = centrifuge density gradient); replicate (r1 to r6).

SIGSBEE5 cruise and processed by the manual sorting method displayed the same pattern as values from the SIGSBEE4 cruise (Fig. 6). The lowest values were recorded in samples from the continental slope (18–91 mgC m^{-2}), with an average value of 25.5 ± 6.0 mgC m^{-2}. These values were 3 times higher than those recorded in the SIGSBEE4 cruise. In contrast, the values recorded in samples from the abyssal plain were higher (21–48 mgC m^{-2}) with an average value of 35.5 ± 11.4 mgC m^{-2}, twice as high as values recorded in the SIGSBEE4 cruise.

3.5. Multivariate canonical analysis of variance

The multivariate anova results showed that the interaction between the main factors, as well as the main factor depth zone, was not significant. The main factor sorting method was significant for all three data sets, however. Plots drawn to illustrate the nature of the differences between sorting methods (Figs. 7A–C) show that 5 taxonomic groups were present more often (Fig. 7A), were more abundant (Fig. 7B), and had higher biomasses (Fig. 7C) in the manual than in the centrifuge density gradient sorting results, in the two SIGSBEE cruises. The Nematoda occurred in almost all samples and hence were not a distinctive group to establish differences in the presence–absence plot (Fig. 7A).

The efficiency values using the Ludox-centrifuge density sorting method in samples from the continental slope were $27 \pm 22\%$ ($n = 12$) and $20 \pm 13\%$ ($n = 12$) in the abyssal plain.

4. Discussion and conclusions

It is interesting to find an increasing trend in animal density with increasing depth in data obtained by manual sorting, when plotting the published density results from studies carried out in tropical oceans. These results are consistent with the results observed in the SIGSBEE4 cruise (Fig. 5), and opposite to the diminishing trends with depth in the Northern Gulf (Baguley et al., 2006), which agree with the results obtained with the centrifuge sorting method (Fig. 5) and suggest that this pattern could be an artefact of the sorting strategy. A similar suggestion was made by Thiel (1983) with regard to sieve mesh sizes, and by Shirayama and Fukushima (1995) for the sampling methods. The results reported by Alongi and Pichon (1988) suggest that abundance diminished with increasing depth due to an increased limitation in food supply.

The two sorting methods produced different results in deep-sea meiofaunal samples. The extraction efficiencies with the centrifugation method have been recognized to be lower in foraminiferida at Ludox solution densities below 1.15 g cm (Burgess, 2001). Other authors have suggested keeping Ludox solution density higher than 1.31 g cm and centrifuging the samples several times, increasing the abundance values significantly (Gambi and Danovaro, 2006), although the effects of this differ (Somerfield and Warwick, 1996; Baguley et al., 2006; Fleeger et al., 2006). This study used the method developed by Baguley et al., (2004) for biomass estimation, to be consistent with Northern Gulf of Mexico studies.

Different sorting techniques affect biomass values as well, especially for nematodes, harpacticoid copepods, and foraminiferida which are important to model and simulate energy flows in ecosystems. Most studies describe a diminishing trend of change with depth for both abundance and biomass (Tietjen, 1992). The biomass values in the literature increase with increasing depth when the manual sorting method was used. In contrast the centrifuge sorting data show an opposite trend. These opposite trends also were found in our samples.

The efficiency of extraction has been reported to be quite variable in the literature; the trends change according to the taxa involved (Schwinghamer, 1981; de Jonge and Bouwman, 1977; Burgess, 2001; Cromar and Williams, 1991). The results obtained in the present study allowed us to recover significantly more foraminiferida, hydrozoans, polychaetes, nematodes, and harpacticoid copepods with manual sorting.

The inclusion of foraminiferida in the counts increased the total abundance values by 29–99% for meiofaunal abyssal samples. These results are consistent with those authors who reported an under-representation of foraminiferida, molluscs and ostracodes in results obtained from the centrifuge sorting method (de Jonge and Bouwman, 1977; Olafsson and Elmgren, 1997; Kröncke et al., 2000).

The taxonomic groups that have rounded shapes and carbonate shells (i.e. foraminiferida, ostracoda) or are heavier and larger (e.g. polychaete larvae, hydrozoa) were found during manual sorting, but were missed with the centrifuge sorting method and remained in the sediment. They account for the lower abundance and biomass values recorded in the samples sorted with the centrifuge method.

References

Alongi, D.M., 1986. Quantitative estimates of benthic protozoa in tropical marine systems using Silica Gel: a comparison of methods. East Coastal Marine Science 23, 443–450.

Alongi, D.M., Pichon, M., 1988. Bathyal sediments of the Western Coral Sea: distribution and abundance in relation to microbial standing stocks and environmental factors. Deep-Sea Research 35, 491–503.

Baguley, J.G., Hyde, L.J., Montagna, P.A., 2004. A semi-automated digital micro-photographic approach to measure meiofaunal biomass. Limnology and Oceanography Methods 2, 181–190.

Baguley, J.G., Montagna, P.A., Hyde, L.J., Kalke, R.D., Rowe, G.T., 2006. Metazoan meiofauna abundance in relation to environmental variables in the Northern Gulf of Mexico deep sea. Deep-Sea Research 53, 1344–1362.

Barnett, B.E., 1980. A physico-chemical method for the extraction of marine and estuarine benthos from clays and resistant muds. Journal of the Marine Biological Association of the United Kingdom 60, 225–256.

Bloem, J., Bär-Gilissen, B., Cappenberg, T.E., 1986. Fixation, counting, and manipulation of heterotrophic nanoflagellates. Applied and Environmental Microbiology 52, 1266–1272.

Bouma, H.A., 1972. Distribution of sediments and sedimentary structures in the Gulf of México. In: Rezak, R., Henry, V.J. (Eds.), Contribution on the Geological and Geophysical Oceanography of the Gulf of Mexico. Texas A&M University, Oceanography Studies, vol. 3. Gulf Pub, Houston, Texas, pp. 35–65.

Burgess, R., 2001. An improved protocol for separating meiofauna from sediments using colloidal silica sols. Marine Ecology Progress Series 214, 161–165.

Cromar, G.L., Williams, D.D., 1991. Centrifugal flotation as an aid to separating invertebrates from detritus in benthic samples. Hydrobiología 209, 67–70.

de Jonge, V.N., 1979. Quantitative separation of benthic diatoms from sediments using density gradient centrifugation in the colloidal silica Ludox-TM. Marine Biology 51, 267–278.

de Jonge, V.N., Bouwman, L.A., 1977. A simple density separation technique for quantitative isolation of meiobenthos using the colloidal silica Ludox-TM. Marine Biology 42, 143–148.

Feller, R.J., Warwick, R.M., 1979. Energetics. In: Higgins, R.P., Thiel, H. (Eds.), Introduction to the Study of Meiofauna. Smitsonian, Washington, pp. 181–196.

Fleeger, J.W., Carman, K.R., Weisenhorn, P.B., Sofranko, H., Marshall, T., Thistle, D., Barry, J.P., 2006. Simulated sequestration of anthropogenic carbon dioxide at a deep-sea site: effects on nematode abundance and biovolume. Deep-Sea Research 53, 1135–1147.

Gambi, C., Danovaro, R., 2006. A multiple-scale analysis of metazoan meiofaunal distribution in the deep Mediterranean Sea. Deep-Sea Research 53, 1117–1134.

Heip, C., Smol, N., Hautekiet, W., 1974. A rapid method of extracting meiobenthic nematodes and copepods from mud and detritus. Marine Biology 28, 79–81.

Hessler, R.R., Jumars, P.A., 1974. Abyssal community analysis from replicate box cores in the central North Pacific. Deep-Sea Research 21, 185–209.

Higgins, P.R., Thiel, H., 1988. Introduction to the Study of Meiofauna. Smithsonian Institution Press.

Kröncke, I., Vanreusel, A., Vincx, M., Wollenburg, J., Mackensen, A., Liebezeit, G., Behrends, B., 2000. Different benthic size-compartments and their relationships to sediment chemistry in the deep Eurasian Artic Ocean. Marine Ecology Progress Series 199, 31–41.

Legendre, P., Legendre, L., 1998. Numerical Ecology, second English ed. Elsevier, Amsterdam.

Nichols, J.A., 1979. A simple flotation technique for separating meiobenthic nematodes from fine grained sediments. Transactions of the American Microscopical Society 98, 127–130.

Olafsson, E., Elmgren, R., 1997. Seasonal dynamics of sublittoral meiobenthos in relation to phytoplankton sedimentation in the Baltic Sea. Estuarine, Coastal and Shelf Science 45, 149–164.

Schwinghamer, P., 1981. Extraction of living meiofauna from marine sediments by centrifugation in a silica-sol-sorbitol mixture. Canadian Journal of Fisheries and Aquatic Sciences 38, 476–478.

Shirayama, Y., Fukushima, T., 1995. Comparisons of deep-sea sediments and overlying water collected using multiple corer and box corer. Journal of Oceanography 51, 75–82.

Somerfield, P.J., Warwick, R.M., 1996. Meiofauna in marine pollution monitoring programmes. A Laboratory Manual, Directorate of Fisheries Research (MAFF), Lowestoft, UK, 71pp.

Sørensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analysis of the vegetation on Danish commons. Biologiske Skrifter 5, 1–34.

ter Braak, C.J.F., Smilauer, P., 2002. Canoco Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5). Microcomputer Power, Ithaca, New York.

Thiel, H., 1983. Meiobenthos and nanobenthos of the deep-sea. In: Rowe, G.T. (Ed.), Deep-Sea Biology. Wiley, NY, pp. 167–230.

Tietjen, J.H., 1992. Abundance and biomass of metazoan meiobenthos in the deep-sea. In: Rowe, G.T., Pariente, V. (Eds.), Deep-sea Food Chains and the Global Carbon Cycle. Kluwer Academic Publishers, Dordrecht, pp. 45–62.

Warwick, R.M., Community structure of estuarine meiobenthos. Marine Ecology Progress Series 18, 97–111.

Warwick, R.M., Price, R., 1979. Ecological and metabolic studies on free-living nematodes from an estuarine mud flat. Estuarine and Coastal Marine Science 9, 257–271.

describing

Two-way canonical analysis of variance for paired observations

Pierre Legendre, mars 2004

The analysis of variance problem at hand is difficult because it involves two crossed factors, the observations are paired for one of the factors, and the response data are multivariate (15 taxa). Canonical redundancy analysis (RDA; Rao 1964) can be used to compute a multifactorial and multivariate analysis of variance (Legendre & Anderson 1999; ter Braak & Smilauer 2002, Section 3.7.6). RDA is the direct extension of multiple regression to the modelling of multivariate response data (Legendre & Legendre 1998).

This section shows how the factors and pairing variables must be coded for this type of problem. We will use an example data set presented by Zar (1999, p. 162) to illustrate the paired-sample *t*-test and show that RDA can be used to obtain the same test, provided that the factors are coded in an appropriate way. We will then modify Zar's example to include a second factor crossed with the first one, thus creating a univariate analogue of the multivariate data studied in the main paper.

Comparing the means of two groups of paired observations

The original example of Zar compares the means of the lengths of the hind and fore legs of 10 deer using a paired-sample *t*-test (Table A1.1). The paired-sample statistic for that example is $t = 3.41379$ with 9 degrees of freedom and an associated two-tailed parametric probability P =

0.0077 (the numerical results reported by Zar (1999), are imprecise and differ slightly from these values). The mean length of the hind legs is 144.7 cm and that of the fore legs is 141.4 cm. The mean difference is thus 3.3 cm. The test indicates that, if the sample data are representative of the population, the hind legs are significantly longer than the fore legs in that population of deer.

Using coded pairing variables, we will recompute the t-test by (1) multiple regression and (2) canonical redundancy analysis (RDA). The dummy variables representing the main factor "Hind/Fore leg" as well as the pairing of the legs for each deer are shown in Table A1.1. "Hind/Fore leg" is coded as $+1$ and -1 for a reason that will be explained in the next section. For the simple example of the present section, it could have been coded as 0 and 1, with the same result.

The tests of significance of the factor "Hind/Fore leg" is obtained by regressing "Leg length" on "Hind/Fore leg", using as covariables the nine dummy variables shown in Table A1.1 which pair the hind and fore legs pertaining to the same animals. In regression analysis, one only has to include all variables in the analysis and look for the test of the parameter associated with "Hind/Fore leg" in the multiple regression results. The regression results are found in Table A1.2. The t-statistic for the test of significance of the partial regression coefficient of the term "Hind/Fore leg", as well as the P-value, are identical to those of the t-test for paired observations reported in the previous paragraph. Thus, inclusion of the P1-P9 dummy variables as covariables in the multiple regression effectively produced a test for the parameter of the "Hind/Fore leg" variable equivalent to a t-test for paired observations.

The exact same results can be obtained by canonical redundancy analysis (RDA), using the program Canoco (ter Braak & Smilauer 2002). "Leg length" can be analysed using "Hind/Fore leg" as the explanatory variable, and variables P1 to P9 as covariables. The statistic

for the test of the explanation of "Leg length" provided by "Hind/Fore leg" is $F = 11.654$, which is the square of $t = 3.41379$. Canoco tests the F-statistic by permutation; the P-value estimated in that way remains valid when the normality assumptions of the parametric test are not met; it is also entirely appropriate for small samples, as in the present example (Legendre & Legendre 1998, Section 1.2; Anderson & Legendre 1999). For the deer leg problem, the P-value provided by Canoco was 0.0114 after 999 random permutations of the residuals under the reduced model. This is very close to the parametric probability reported in a previous paragraph.

The canonical analysis solution also provided a statistic, called *trace* in the Canoco program, which estimates the proportion of the variance of "Leg length" explained by "Hind/Fore leg": *trace* = 17.855%. This is equivalent to a partial R^2 statistic. Since the covariables P1-P9 are orthogonal to the explanatory variable "Hind/Fore leg" in this particular example, *trace* is equal to the coefficient of determination (R^2) of the simple linear regression of "Leg length" on "Hind/Fore leg": $R^2 = 0.17855$ or 17.855%. The F-statistic, which is the ratio of two *independent* portions of the dependent variable's variation, is computed as follows from the R^2 values (Legendre & Legendre 1998; ter Braak & Smilauer 2002; see Fig. A1.1, leaving out for the moment the contributions of "Sex" and "Interaction"):

$$F = \frac{R^2_{\text{explanatory variables}} \big/ m}{(1 - R^2_{\text{explanatory variables}} - R^2_{\text{covariables}}) \big/ (n - 1 - m - q)} \tag{1}$$

$$F = \frac{0.17855 \big/ 1}{(1 - 0.17855 - 0.68356) \big/ (20 - 1 - 1 - 9)} = 11.654$$

with the number of data rows $n = 20$, the number of explanatory variables $m = 1$ and the number of covariables $q = 9$ (Table A1.1).

Fig. A1.1 shows why, in the case of paired observations, a test involving the pairing variables as covariables is necessarily equally or more powerful than an unpaired *t*-test or an anova for unpaired data. In a test for unpaired data, the denominator of the *F*-statistic would not include the variation accounted for by the pairing variables (68.356% in this example), nor the degrees of freedom associated with these variables (9). The *F*-statistic would be:

$$F = \frac{0.17855 / 1}{(1 - 0.17855) / (20 - 1 - 1)} = 3.912$$

This value, which is the square of the statistic one would obtain in a *t*-test for independent data (*t* = 1.97802), leads to a higher probability of the data under H_0 (parametric P = 0.0634, two-tailed test, 18 degrees of freedom) than the corresponding *t*-test for paired observations (P = 0.0077, two-tailed test, 9 degrees of freedom). The *t*-test for independent observations would not reject H_0 at the 5% significance level in this example, whereas the *t*-test for paired observations does. The latter is more powerful because it is more capable of identifying that a small difference between the means of the groups of observations is significantly larger than what is expected from random variation.

Two-way analysis of variance for paired observations

The next challenge was that there are not one, but two factors in the benthos data analysed in the main paper: the two geographic zones and the two counting methods. We could have analysed the two zones separately, but there was real interest in determining if there were significant differences in the infauna between the two zones. So, the design of the analysis was: two crossed fixed factors: the geographic zones called "Zones", and the counting methods called "Methods". The observations were paired for the second factor since the *same cores* were analysed and counted using the two methods. For illustration of the statistical method of analysis,

the Zar example was turned into a two-factor problem by pretending that animals 1 to 5 were males and animals 6 to 10 were females (Table A1.1). A possible interaction between "Hind/Fore leg" and "Sex" can be studied through an "Interaction" variable obtained by computing the product of the dummy variables representing the main factors.

In Table A1.1, the variables "Hind/Fore leg" and "Sex", which are crossed, are coded using orthogonal dummy variables. The codes for each one sum to 0 and their scalar product (or dot product) is 0; these variables are said to be *orthogonal*. For balanced study designs, the row-by-row products of variables coded in this way produce an interaction variable which is orthogonal to the two crossed variables. Thanks to this property, the fractions of variation explained by the main factors and the interaction term do not overlap in Fig. A1.1. If we had used ordinary dummy variables coded by 0's and 1's, the interaction variable would not have been orthogonal to the main factors and their circles would have overlapped in Fig. A1.1. Various methods can be used to code an experimental factor into orthogonal dummy variables; an easy method is described in Appendix C of Legendre & Anderson (1999).

The relationships among the explanatory variables were complex in this example. Whereas the variables coding for "Hind/Fore leg", "Sex" and "Interaction" are orthogonal to one another (their correlation is 0), the pairing dummy variables P1-P9 explain "Sex" completely in a multiple regression ($R^2 = 1$). Table A1.1 shows that the dummy variables P1 to P5 separate the males from the females, so that variables P1 to P5 are sufficient to explain "Sex" entirely. A Venn diagram of the coefficients of determination (*trace* statistics, or R^2) of the explanatory variables on the "Leg length" response variable, drawn in the spirit of the variation partitioning diagrams pioneered by Borcard et al. (1992) and Legendre (1993), is presented in Fig. A1.1. In other studies where the pairing affects the two crossed factors, if the factors are orthogonal to

each other (as it is the case if the design is balanced), they are also orthogonal to the set of pairing dummy variables. In such a case, the surfaces explained by the two main factors, the interaction, and the pairing variables do not overlap in the Venn diagram.

Because, in the present design, "Sex" was entirely explained by (and was thus collinear to) the pairing variables, we computed the regression residuals of P1-P9 on "Sex" before proceeding with the analysis. The rank of the covariance matrix of the 9 residual variables was 8; projection on "Sex" had resulted in the loss of one dimension. The 9 residual variables were subjected to principal component analysis (PCA). The object scores along the 8 axes (Axis 1 to Axis 8) that had non-zero positive eigenvalues were linearly independent of "Sex"; they became the new explanatory variables in the analysis of variance that follows.

The tests of significance of the main effects and interaction are tests of the partial regression coefficients of these terms in a multiple regression of "Leg length" against the two main effects and interaction variables, when the principal components are also included as covariables in the regression analysis.

Test of the "Interaction" — From the regression analysis results (Table A1.3), the statistic associated with "Interaction" is $t = -1.70941$, P = 0.1257 (parametric probability). The same result was obtained from a redundancy analysis (RDA) of the response variable "Leg length" by the explanatory dummy variable coding for "Interaction", with "Hind/Fore leg", "Sex", and the 8 principal components representing pairing residuals (Axis 1-Axis 8) as covariables ($F = 2.922$, P = 0.1222 after 9999 permutations of residuals of the reduced model in Canoco). With $m = 1$ and $q = 10$, we find, using in Eq. 1 the R^2 values from Fig. A1.1:

$$F = \frac{0.03689 / 1}{(1 - 0.17855 - 0.15757 - 0.52599 - 0.03689) / (20 - 1 - 1 - 10)} = \frac{0.03689}{0.10100 / 8} = 2.922$$

"Hind/Fore leg", "Sex", and Axis 1-Axis 8 are used as covariables in this analysis. $F = t^2$ since, in this example, the number of degrees of freedom of the numerator of the F-statistic is $m = 1$. If there were more than 2 classes in one of the main factors, "Interaction" would be coded by more than one variable and the anova could only be done by partial multiple regression for a single response variable, or partial canonical analysis for single or multiple response variables. In the present example, a non-significant interaction indicates that the effects of the classes of each factor do not significantly differ between the classes of the other factor. We can thus proceed with the analysis of the main factors in the two-way context.

Because "Sex" is entirely explained by, and collinear to, the pairing variables P1-P9, an equivalent method is to exclude "Sex" from the analysis and analyse "Leg length" by the explanatory dummy variable "Interaction", with covariables "Hind/Fore leg" and the original 9 dummy variables P1-P9 representing pairing. If one uses a program, such as Canoco, capable of handling collinearity among the covariables, "Sex" can remain among the covariables in the analysis. All these forms produce the same results for the test of "Interaction".

Test of the "Hind/Fore leg"—We expect the test of the factor "Hind/Fore leg" to either have the same power or be more powerful in a two-way anova than the t-test of that factor alone reported in the previous section. The reason is that the variability explained by "Interaction" was not included in the denominator of the F-statistic in the previous one-factor analysis of variance (Fig. A1.1). If "Interaction" contributes to lowering the value of the F-statistic more than a random factor would, it may reduce the P-value of the test.

In the regression analysis (Table A1.3), the t-statistic associated with this factor is 3.76070, P = 0.0055 (parametric probability). This result shows slightly more power than the t-test reported in the previous section (P = 0.0077). An equivalent result was obtained from a

redundancy analysis (RDA) of the response variable "Leg length" by the explanatory dummy variable coding for "Hind/Fore leg", with "Sex", "Interaction", and the eight principal components representing pairing (Axis 1-Axis 8) as covariables (F = 14.143, P = 0.0085 after 9999 permutations of residuals of the reduced model in Canoco). $F = t^2$ since, in this example, the number of degrees of freedom of the numerator of the F-statistic is $m = 1$.

Because "Sex" is entirely explained by, and collinear to, the pairing variables P1 to P9, an equivalent method is to exclude "Sex" from the analysis and analyse the response variable "Leg length" by the explanatory dummy variable coding for "Hind/Fore leg", with covariables "Interaction" and the original 9 dummy variables representing pairing.

Test of the "Sex"—In the regression analysis table, the t-statistic associated with this factor is −3.53276, P = 0.0077 (parametric probability). The same result was obtained from a redundancy analysis (RDA) of the response variable "Leg length" by the explanatory dummy variable coding for "Sex", with "Hind/Fore leg", "Interaction" and the 8 principal components representing pairing (Axis 1-Axis 8) as covariables (F = 12.481, P = 0.0099 after 9999 permutations of residuals of the reduced model in Canoco). $F = t^2$ since, in this example, the number of degrees of freedom of the numerator of the F-statistic is $m = 1$.

Discussion

The results reported in the previous sections show that, in the case of a single response variable, multiple regression or canonical redundancy analysis (RDA) can be used to obtain tests equivalent to a paired-sample t-test (for a single factor) or a paired-sample two-way analysis of variance (for two crossed factors). For multivariate response data tables, RDA is the instrument of choice to conduct such tests since it can handle multivariate response data tables and allows

incorporation in the analysis of dummy variables coding for the main factors, the interaction, as well as pairing variables tailored to the problem under study. For the multivariate case, RDA offers the additional advantage of allowing the representation of the manova results in the form of biplots displaying the responses of the dependent variables to the manova factors. RDA is used in the main paper to analyse multivariate data of the same form as the univariate example presented in this Appendix.

Tests of significance of the fixed factor in a mixed model (i.e., an anova with a fixed and a random factor), and of any factor in a model II anova (i.e., an anova for two random factors), require the use of the interaction mean square in the denominator, instead of the residual mean square. How to obtain a correct test is described in Section 3.7.6 of the Canoco manual (ter Braak & Smilauer 2002). For paired observations, the pairing variables must be included in the analysis as covariables, as shown in the present example.

References

Anderson, M. J. and P. Legendre. 1999. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. *J. Stat. Comput. Simul.* 62: 271-303.

Borcard, D., P. Legendre and P. Drapeau. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73: 1045-1055.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology* 74: 1659-1673.

Legendre, P. and M. J. Anderson. 1999. Distance-based redundancy analysis: testing multispecies responses in multifactorial ecological experiments. *Ecol. Monogr.* 69: 1-24.

Legendre, P. and L. Legendre. 1998. *Numerical ecology, 2nd English edition.* Elsevier Science BV, Amsterdam.

Rao, C. R. 1964. The use and interpretation of principal component analysis in applied research. *Sankhyaá A* 26: 329-358.

ter Braak, C. J. F. and P. Smilauer. 2002. *Canoco reference manual and CanoDraw for Windows user's guide: software for canonical community ordination (version 4.5).* Microcomputer Power, Ithaca, New York.

Zar, J.H. (1999) *Biostatistical analysis, 4th edition.* Prentice Hall, Upper Saddle River, New Jersey.

Table A1.1. Example data set: lengths of the hind and fore legs of ten deers (Zar 1999). Variables P1-P9 identify the fore and hind legs pertaining to the same animals. The example was modified by creating a second variable "Sex" which is crossed with "Hind/Fore leg". The dummy variable coding for "Interaction" is the product of the main factors' ("Hind/Fore leg" and "Sex") dummy variables.

Deer No	Length (cm)	Pairing dummy variables P1 P2 P3 P4 P5 P6 P7 P8 P9									Hind/Fore leg	Sex	Interaction
Hind legs													
1	142	1	0	0	0	0	0	0	0	0	1	1	1
2	140	0	1	0	0	0	0	0	0	0	1	1	1
3	144	0	0	1	0	0	0	0	0	0	1	1	1
4	144	0	0	0	1	0	0	0	0	0	1	1	1
5	142	0	0	0	0	1	0	0	0	0	1	1	1
6	146	0	0	0	0	0	1	0	0	0	1	–1	–1
7	149	0	0	0	0	0	0	1	0	0	1	–1	–1
8	150	0	0	0	0	0	0	0	1	0	1	–1	–1
9	142	0	0	0	0	0	0	0	0	1	1	–1	–1
10	148	0	0	0	0	0	0	0	0	0	1	–1	–1
Fore legs													
1	138	1	0	0	0	0	0	0	0	0	–1	1	–1
2	136	0	1	0	0	0	0	0	0	0	–1	1	–1
3	147	0	0	1	0	0	0	0	0	0	–1	1	–1
4	139	0	0	0	1	0	0	0	0	0	–1	1	–1
5	143	0	0	0	0	1	0	0	0	0	–1	1	–1
6	141	0	0	0	0	0	1	0	0	0	–1	–1	1
7	143	0	0	0	0	0	0	1	0	0	–1	–1	1
8	145	0	0	0	0	0	0	0	1	0	–1	–1	1
9	136	0	0	0	0	0	0	0	0	1	–1	–1	1
10	146	0	0	0	0	0	0	0	0	0	–1	–1	1

Table A1.2. Regression analysis of the variable "Leg length" by the dummy variables coding for "Hind/Fore leg" and the 9 dummy variables (P1-P9) pairing the hind and front legs of the 10 animals. The t-statistic and P-value for "Rear/Front leg" are in bold.

Variable	Regression coefficient	Standard regression coefficient	t-statistic	P-value
Intercept	147.00000	0.00000	96.17686	< 0.0001
Hind/Fore leg	1.65000	0.42256	**3.41379**	**0.0077**
P1	–7.00000	–0.53780	–3.23844	0.0102
P2	–9.00000	–0.69146	–4.16371	0.0024
P3	–1.50000	–0.11524	–0.69395	0.5052
P4	–5.50000	–0.42256	–2.54449	0.0315
P5	–4.50000	–0.34573	–2.08186	0.0671
P6	–3.50000	–0.26890	–1.61922	0.1399
P7	–1.00000	–0.07683	–0.46263	0.6546
P8	0.50000	0.03841	0.23132	0.8222
P9	–8.00000	–0.61463	–3.70108	0.0049

Table A1.3. Regression analysis of the variable "Leg length" by the dummy variables coding for "Hind/Fore leg", "Sex", "Interaction", and the principal components (Axis 1-Axis 8) derived from the residuals of the regression of dummy variables P1-P9 on "Sex". The t-statistic and P-value for "Rear/Front leg", "Sex", and "Interaction" are in bold.

| Variable	Regression coefficient	Standard regression coefficient	t-statistic	P-value
Intercept	143.05000	0.00000	326.04120	< 0.0001
Hind/Fore leg	1.65000	0.42256	**3.76070**	**0.0055**
Sex	–1.54999	–0.39694	**–3.53276**	**0.0077**
Interaction	–0.75000	–0.19207	**–1.70941**	**0.1257**
Axis 1	–2.30585	–0.18674	–1.66196	0.1351
Axis 2	–1.84359	–0.14930	–1.32877	0.2206
Axis 3	–0.26967	–0.02184	–0.19436	0.8507
Axis 4	2.57978	0.20892	1.85937	0.1000
Axis 5	–4.37359	–0.35419	–3.15228	0.0136
Axis 6	6.16895	0.49959	4.44626	0.0021
Axis 7	–0.60997	–0.04940	–0.43964	0.6718
Axis 8	6.00003	0.21730	1.93398	0.0892
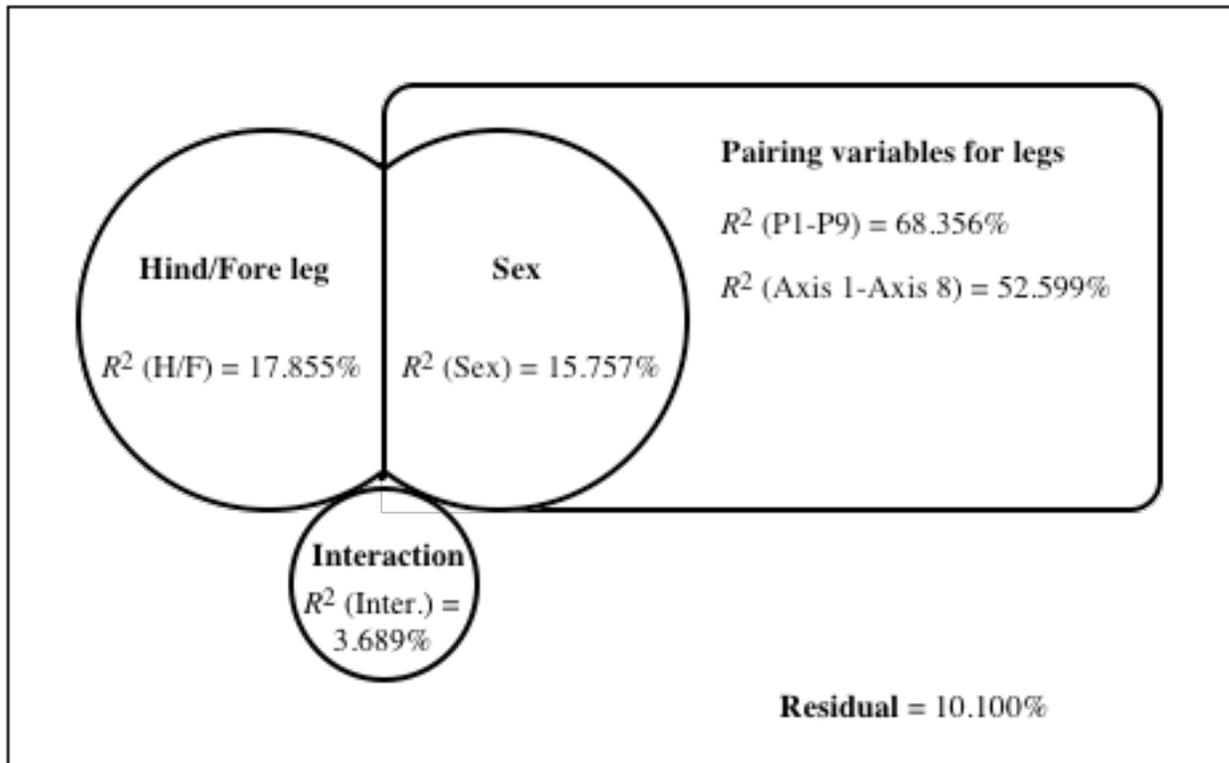
Fig. A1.1 – Venn diagram showing how the variation of the response variable (outer rectangle) is partitioned among the explanatory variables "Hind/Fore leg", "Sex", "Interaction, and the pairing variables for the legs. The pairing variables (68.356%) accounts entirely for the variation explained by "Sex" (15.757). The variation uniquely explained by the pairing variables and not by "Sex" is 68.356 – 15.757 = 52.599%. 10.100% of the variation of the response variable is not explained by the explanatory variables used in the model.