



# Scale-dependent distribution of soft-bottom infauna and possible structuring forces in low diversity systems

Patrik Kraufvelin<sup>1,2,\*</sup>, Jens Perus<sup>2</sup>, Erik Bonsdorff<sup>2</sup>

<sup>1</sup>ARONIA, Coastal Zone Research Team, Åbo Akademi University and Novia University of Applied Sciences, Raseborgsvägen 9, 10600 Ekenäs, Finland

<sup>2</sup>Environmental and Marine Biology, Åbo Akademi University, Artillerigatan 1, 20500 Turku/Åbo, Finland

**ABSTRACT:** Despite their central roles in ecology, patterns and scales of variation in the distribution of species and their densities are seldom satisfactorily described for most ecosystems. This is mainly due to spatial structures being scale-dependent across broad ranges of scales with partly different processes operating at different scales. In this study, scale-dependent distributional patterns in soft-sediment macrofaunal communities of naturally low species richness were studied on the Åland Islands, northern Baltic Sea, in 3 model areas (I, II and III) representing different environmental regimes with regard to openness (water renewal rate), depth, and organic and nutrient content as well as an increasing intrasite variability. A systematic grid design comprising 64 biotic and 16 abiotic sampling points, hierarchically arranged at 3 (at distances 10, 100 and 1000 m) and 2 spatial scales (100 and 1000 m), respectively, was used in each area. The intrasite distribution of macrofauna was predictably the most uniform in Area I, the least variable environmentally, followed by Area II, which was intermediate with respect to both abiotic and biotic variability, and then Area III, which was always the most variable area. This was true when analysed by both univariate and multivariate statistical means. When examining the contribution of each spatial scale to the total variance, most variability was observed at the smallest sampled level (10 m) for the majority of the studied variables. Despite a high overlap in species composition, the areas differed significantly in animal community structure. This was driven by both (1) the top-dominant species (*Macoma*, *Mytilus*, *Monoporeia*, *Oligochaeta*, *Marenzelleria*) and (2) the proportion and distribution of rare species. These systems with low diversity showed clear species distribution patterns driven by local environmental conditions, but also showed a substantial biotic component seen as the proportion and role of rare species and the high spatial variability observed at the smallest scale. Our understanding of spatial variability (or 'stability') is thus influenced by both the habitat and the species composition. The results are valuable for the refinement of sampling design, improved interpretation of monitoring programs (including detection of anthropogenic impact) and incorporation of scaling issues into questions of marine biodiversity and conservation.

**KEY WORDS:** Baltic Sea · Zoobenthos · Soft sediment communities · Spatial patterns and processes · Scales · Heterogeneity · Hierarchical sampling · Components of variation · Rare species

Resale or republication not permitted without written consent of the publisher

## INTRODUCTION

Patterns and scales of variation in the distribution of species and their densities are not easily described, despite their central roles in ecology, and may thereby

inhibit our understanding of natural phenomena (Levin 1992, Thrush et al. 1997). This is mainly because spatial structures are usually scale dependent across broad ranges of scales with partly different processes operating at different scales (Wiens 1989, Thrush

\*Email: patrik.kraufvelin@abo.fi

1991). In marine soft-sediment communities, differences in physical environmental factors, such as water depth, water currents and sediment characteristics, are generally believed to determine large-scale patterns of zoobenthic distribution (Gray 1974, Barry & Dayton 1991). At smaller spatial scales, zoobenthic patchiness may be due more to biotic processes such as competition, predation and reproduction, although, for example, local differences in oxygen levels and sediment organic content can also play roles (Thrush et al. 1989, Bonsdorff & Blomqvist 1993, Bonsdorff et al. 2003, Chapman et al. 2010). The mechanisms through which patterns are expressed also differ with scale. At large scales, zoobenthic patterns are mainly established through differences in recruitment and mortality, while at small scales, behavioural responses, competition and active or passive movements of organisms may be more involved in the formation of patterns (Hewitt et al. 1996, Huston 1999, Norkko et al. 2001, Valanko et al. 2010). However, the spatial scales of variability are seldom known before the sampling has been undertaken, and variability at any spatial scales between that of the sampling units (small scale) and the locations sampled (large scale) will not always be revealed by the sampling design. The greater the distance between one level of sampling and next, the more interpolation of scale of variability will occur (Morrisey et al. 1992). Appropriate study scales can, however, be revealed by the use of nested sampling designs (Underwood 1981, 1997, Andrew & Mapstone 1989), where each of a series of successively smaller scales is nested within the scale above. This will provide an estimate about how much each single scale will contribute to the total variation among samples at the greatest scale.

An understanding of processes in complex assemblages is dependent on a good insight into spatial and temporal structural patterns at various scales (Legendre 1993, Underwood & Chapman 1996, 1998) and is crucial for accurate measurements of community structure and attempts to apply results to bigger contexts (Levin 1992, Thrush et al. 1997). During the last decades, an increasing number of studies have been carried out with sampling designs capable of unravelling the confounding patterns of biotic community structure at different spatial and temporal scales (e.g. Borcard et al. 1992, Ysebaert & Herman 2002), although there seem to be conflicting views about the best way in which to accomplish this, i.e. by hierarchical analyses or geostatistical continuous methods (Cole et al. 2001). The reasons for studying spatial and temporal scales and their ecological roles are, however, unarguable and the search for the optimal designs and analysis methods is ongoing (Keeling et al. 2000, Legendre et al. 2002, Hewitt et al. 2007). Hierarchical

methods are usually more appropriate when a large spatial coverage of the sampling is necessary, whereas geostatistical methods should be considered when detailed information of patterns is required (Underwood & Chapman 1996). However, the hierarchical approach only tells where the variability occurs, whereas with the geostatistical approach, variability can be incorporated seamlessly into the analysis. Hierarchical sampling is also strongly dependent on the scales that are chosen, whereas geostatistical approaches offer better potential to interpolate successfully (Hewitt et al. 2007). When the categories actually form a gradient (i.e. depth, nutrient concentrations), a hierarchical ANOVA can be particularly insensitive (Sommerfield et al. 2002, Hewitt et al. 2007). Any knowledge of scale-dependent distributional patterns, however, may reveal important information about community dynamics in space and time, reflect community stability and be used to investigate the predictability of assemblages (Levin 1992, Thrush et al. 1997). Furthermore, it can be determined at which scales most variability is present and thus possibly most important (Morrisey et al. 1992, Hewitt et al. 2007, Chapman et al. 2010). When unravelling scale-dependent matters, it is also important to pay attention to the role of a single (rare) species in addition to the dominant ones (Cao et al. 1998, Schlacher et al. 1998, Ellingsen et al. 2007, Fontana et al. 2008). Soft-sediment communities often consist of relatively high fractions of rare species compared with other environments (Gray et al. 2005).

Despite low species richness, areas with different habitat quality may still display clear differences in biotic community responses (Barry & Dayton 1991, Bergström et al. 2002), which is true for the nontidal species-poor Baltic Sea. During the last 10 000 yr, this geologically young sea area has undergone several dramatic shifts in salinity, but its present state of low salinity has prevailed for about 3000 yr (Voipio 1981, Gustafsson & Westman 2002). In general, few species are well adapted to living in brackish waters and this, together with the young age of the Baltic Sea, makes its biota unusually poor in species richness (Bonsdorff & Blomqvist 1993, Bonsdorff & Pearson 1999). This in turn makes the Baltic Sea an ideal environment in which to test ecological theory, as the number of possible species interactions is clearly lower than in truly marine environments. However, with the exceptions of a few studies (e.g. Sparrevik & Leonardsson 1995, Rumohr et al. 2001, Bergström et al. 2002, Josefson 2009), there is a clear lack of planned studies on scaling issues from the Baltic Sea, although some scale-related results indeed have been achieved as by-products from studies with other main focuses (e.g. Kraufvelin et al. 2001, Bonsdorff et al. 2003, Laine 2003, Perus & Bonsdorff 2004). Symptomatic for most

of these studies is that the special features of the Baltic Sea, such as clear gradients of salinity, organic enrichment and oxygen levels, seem to be driving much of the observed community patterns over space and time (Bonsdorff & Pearson 1999, Bonsdorff 2006) and this will also apply to some extent to the present study.

In the present study, the variation in soft-sediment macrofaunal communities of naturally low species richness at a range of spatial scales was examined by means of a nested sampling design. A hierarchical sampling grid comprising 64 individual grab samples from each of 3 defined areas (20 to 100 km apart along waterways) on the Åland Islands was used. The samples were taken from successively larger scales (at increasing distances, i.e. 10, 100 and 1000 m). The 3 model areas comprised more or less different physical regimes with regard to exposure, depth, sediment quality, organic content, salinity, oxygen conditions and nutrient levels and/or variability in these variables. The areas were chosen to pinpoint both similarities (generality) and differences (site-specific traits) in scale-dependent distribution patterns of sediment infauna. The central aim was to, by use of univariate and multivariate statistical techniques, describe and analyse the environmental background variables of the areas and the macrofaunal community structure (as both abundance and biomass), while discussing

possibly underlying structuring abiotic and biotic mechanisms at the various study scales. The following hypotheses were tested. (1) There are differences in abiotic and biotic variables both among and within study areas (i.e. between different measurement scales). (2) The scales at which abiotic and biotic variables show maximal amounts of variation are relatively consistent across study areas. (3) The area with the highest environmental variability will have the highest variability in the fauna and vice versa for the area with the lowest environmental variability. In addition, we wanted to compare the influence of sampling effort on biodiversity in the 3 areas with published data from other soft-sediment systems.

## MATERIAL AND METHODS

**Study areas.** The study areas are situated around the Åland Islands at the southwest coast of Finland in the northern Baltic Sea (Fig. 1a). The Åland Islands form a mosaic exhibiting a range of environmental conditions, e.g. zoned archipelagos ranging from sheltered bays to open sea as well as large areas with deep water (>200 m) in an otherwise shallow environment (mean depth 20 to 25 m) with long shorelines and a large proportion of littoral areas. Salinity ranges between 5 and

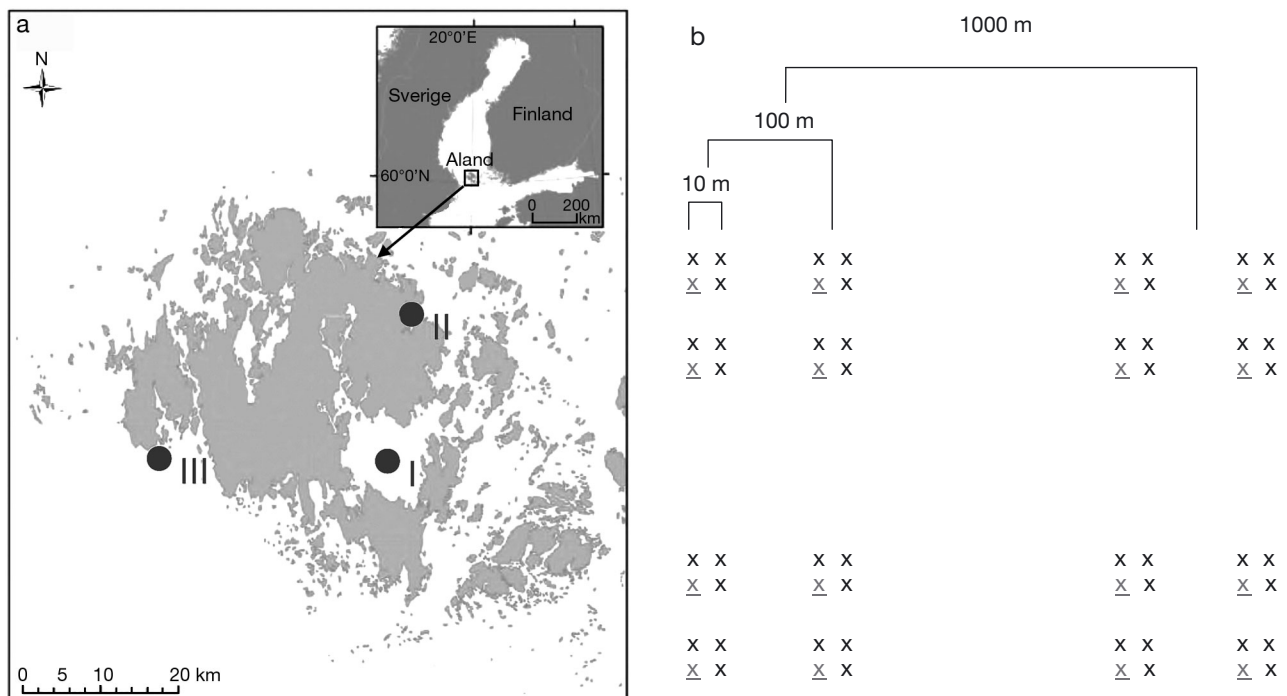


Fig. 1. (a) Location of the sampling areas (Areas I, II and III, indicated with solid circles) on the Åland Islands off the coast of Finland. Inset shows location of study area within the northern Baltic Sea. (b) Schematic view over the total grid set-up for biotic samples in each area with 10, 100 and 1000 m distances indicated (underlined crosses stand for abiotic samples)

8 psu moving from sheltered bays out to the open sea, and the region is influenced by a strong seasonality in hydrographical variables (ice cover is usually present for 30 to 40 d yr<sup>-1</sup> and surface temperatures can rise to above 20°C in summer). More detailed descriptions of the study area can be found elsewhere regarding hydrographical (Seinä & Peltola 1991) and environmental characteristics (Bonsdorff & Blomqvist 1993, Bonsdorff et al. 1997, Perus & Bonsdorff 2004). The different environmental characteristics of the 3 study areas express a gradient of increasing intrasite variability in addition to representing different physical regimes, especially with regard to openness (water renewal rate), depth and organic content (see Table 1).

**Sampling design.** Three separate spatial surveys with identical sampling set-ups were carried out in June 2001. All sampling stations were located with GPS coordinates, which were stored for later revisits. The hierarchical sampling design used is illustrated in Fig. 1b and comprises a total of 64 samples per area. Altogether, 192 benthic samples (species abundance and biomass) were taken in the 3 areas along the following nesting sequence. In each area, the sampling grid was structured at 3 levels: 4 × 4 × 4 = 64 samples with an internal lag of 10 m between the replicates, 4 × 4 subquadrats (10 × 10 m) with an internal lag of 100 m and the full area with a lag of 1000 m between the 4 larger quadrats (100 × 100 m). Hereafter the sampling scales are referred to as 10 m (4 replicates), 100 m (4 levels), 1000 m (4 levels) and area (3 levels, Areas I, II and III). Hydrographical (temperature, salinity, oxygen saturation, pH, total P and total N of the bottom water ~50 cm above the sediment surface) and sediment variables (organic content as loss of ignition) and water depth were recorded from each of the lower left (SW) corner of the 100 m squares, for a total of 16 samples per area. Zoobenthos was sampled with an Ekman-Birge grab sampler (289 cm<sup>2</sup>) and sieved with a 0.5 mm mesh. Samples were stored in 4% buffered formaldehyde (i.e. 10% formalin) and later counted in the laboratory under a dissecting microscope. The organisms were identified to species level (exceptions: Chironomidae, Oligochaeta and *Hydrobia*) and their wet weights measured to the nearest 0.001 g. When applying the spatial aspects of the scale involved in sampling (Legendre & Legendre 1998) to our study, the extent (i.e. the area encompassed by the study) was 100 km, the lags (i.e. the distances between samples) were 10, 100 and 1000 m, the grain (the sampling unit size) was 289 cm<sup>2</sup> and the resolution (or the mesh size of the sieve) was 0.5 mm.

**Statistical analyses.** Data were first analysed by univariate statistical means. Differences in environmental variables were examined by a 2-way nested ANOVA with the 1000 m level nested within Area

and by using the 100 m samples, i.e. the lower left corners of each 10 × 10 m quadrat (100 to 1000 m apart), as replicates. Differences in community and population variables were analysed by a 3-way nested ANOVA by adding the 10 m level replicates to the previous design (Underwood 1981, 1997). Significant overall differences were further broken down by pairwise Student-Newman-Keuls (SNK) tests. Variance components were picked straight from the mean square estimates of the ANOVAs (using untransformed raw data) and were incorporated into the study to elucidate the contribution of each spatial scale (independent of all other scales) to the total variance for every studied environmental or biological community or population variable. Negative components were set to zero and eliminated from the analysis and the remaining components were recalculated as described by Fletcher & Underwood (2002). To obtain a picture of possible species interactions as well as to screen the coupling between environmental and biological data, nonparametric Spearman's rank correlation analyses (chosen because of general problems with the normality) were conducted for all areas at once and specifically for each area.

The number of species in a community for which intermediate yield per sample is compared may be considered as an intermediate result in an infinite series of samples that could be taken from a given site at a given time (Rumohr et al. 2001). Species accumulation plots were generated separately for each area entering the 64 samples in random order 999 times and averaging the resulting curves. Three curves were drawn per Area to illustrate the relationships between sampling and species richness: (1) species observed (Sobs), (2) the Jackknife2 estimate (Burnham & Overton 1979) and (3) Chao2 (Chao 1987).

Nonparametric multivariate techniques (Clarke 1993, Clarke & Ainsworth 1993, Clarke & Gorley 2006, Anderson et al. 2008) were further applied to both environmental and biological data. The environmental data were analysed by a principal component analysis (PCA) based on Euclidean distances. The biological community data were analysed separately for abundance and biomass data by nonmetric multidimensional scaling (NMDS) based on Bray-Curtis similarities and the results were plotted into ordination graphs. The community data were analysed both untransformed (to retain a high influence of the dominant species) and after transformation by the fourth root (to also take the influence of rarer species into consideration); these 2 transformation levels will imply testing of radically different hypotheses. An analysis of multivariate dispersion (MVDISP) was further undertaken to check which areas were the most variable with regard to species abundance and biomass pat-

terns. Furthermore, differences in community structure between the areas and sample scales were tested for by the use of a permutational multivariate ANOVA (PERMANOVA). This analysis provided us with pseudo  $F$ -values and permutational  $p$ -values as well as multivariate variance components revealing at which scale most variability was present in the full community (Anderson 2001, Anderson et al. 2008). Finally, the similarity percentage breakdown procedure (SIMPER) was used for listing the species contributing most to observed similarities within areas and significant dissimilarities between areas, and BIO-ENV was used for linking biotic to abiotic variables.

The univariate analyses were run in GMAV5 or SPSS 17.0, whereas the multivariate statistical analyses and species accumulation curves were run in PRIMER 6.0 and PERMANOVA. Before running the parametric univariate tests, the normality was checked with the Kolmogorov-Smirnov test and homogeneity of variances were determined with Cochran's  $C$ -test. To homogenise variances, it was sometimes necessary to use a square-root transformation,  $\sqrt{(x + 1)}$ , a logarithmic transformation,  $\ln(x + 1)$  or negative inverse transformation,  $-(x + 1)^{-1}$ . For all significance tests, alpha was set at 0.05. For cases where the assumptions of parametric tests still could not be fulfilled despite transformation, a stronger significance level for judging a difference significant was applied (i.e.  $p < 0.001$  instead of  $p < 0.05$ ). Hochberg's sequential Bonferroni correction (Hochberg 1988) was finally used to adjust the overall significance levels to the number of tests performed.

## RESULTS

### Environmental data

The mean values and ranges for environmental variables are listed in Table 1 along with the SD and coefficients of variation (CV, i.e. SD/mean). A homogeneous water depth of 16 to 19 m was a major criterion for choosing Area I, although the topography of the shallower Area II was also fairly homogeneous with its 11 to 15 m depth. Area III, on the other hand, had a clear depth gradient, ranging from 12 to 28 m. The temperature differences partly mirrored the depth differences with generally lower temperatures at deeper parts of the areas and the ranges are clearly wider for Area II (8 to 15°C) and Area III (5 to 17°C) than for Area I (11 to 12°C). A higher salinity (5.9 compared with 5.4 psu), a clearly lower organic content (0.3 to 2.1% compared with 5.2 to 7.5% and 2.9 to 10.1% in Areas I and II, respectively) and, on average, 35% higher total N and 25% higher total P in Area III compared with Areas I and II seemed to represent abiotic differences between the 3 areas (Table 1). However, when analysed with a 2-way nested ANOVA followed by a sequential Bonferroni correction of the  $p$ -values, the 3 investigated areas only differed significantly with regard to salinity, organic content and N concentration (Table 2).

Within the 3 areas themselves, i.e. when comparing the environmental data area-specifically at 1000 m overall in the 2-way ANOVA and separately for each area with SNK-corrected *a posteriori* tests, the ab-

Table 1. Basic descriptions as well as environmental and community data (mean  $\pm$  SD, range and CV) for study areas (Areas I, II and III) in the Åland Islands, northern Baltic Sea

Name Type Bottom Surrounding	Area I Lumparn			Area II Nötvikén			Area III Degersand		
	Semi-enclosed bay, E Åland			Long bay, NE Åland			Open sea area, SW Åland		
	Soft			Soft			Sandy		
	Agricultural land, summer houses			Forests, some summer houses			Forest on the shore, no settlements		
	Mean $\pm$ SD	Range	CV	Mean $\pm$ SD	Range	CV	Mean $\pm$ SD	Range	CV
Depth (m)	17.34 $\pm$ 0.91	16–19	0.05	13.09 $\pm$ 1.45	11–15	0.11	18.09 $\pm$ 4.18	12–28	0.23
Temperature (°C)	11.73 $\pm$ 0.45	10.9–12.2	0.04	11.52 $\pm$ 2.48	8–14.8	0.22	8.74 $\pm$ 3.95	5.2–17.1	0.45
pH	8.02 $\pm$ 0.06	7.9–8.1	0.01	7.96 $\pm$ 0.19	7.7–8.3	0.02	7.87 $\pm$ 0.11	7.7–8.1	0.01
Salinity (psu)	5.41 $\pm$ 0.08	5.1–5.5	0.01	5.44 $\pm$ 0.03	5.4–5.5	0.01	5.91 $\pm$ 0.13	5.6–6.1	0.02
Oxygen (mg l <sup>-1</sup> )	10.21 $\pm$ 0.37	9.5–10.9	0.04	10.38 $\pm$ 0.26	10–11	0.03	10.24 $\pm$ 0.69	8.9–11.2	0.07
Organic content (%)	6.14 $\pm$ 0.68	5.2–7.5	0.11	7.41 $\pm$ 2.00	2.9–10.1	0.27	0.69 $\pm$ 0.42	0.3–2.1	0.61
P ( $\mu$ g l <sup>-1</sup> )	14.76 $\pm$ 2.75	10.1–19.7	0.19	11.69 $\pm$ 2.31	8.6–15.8	0.20	16.68 $\pm$ 2.13	14.1–21.2	0.13
N ( $\mu$ g l <sup>-1</sup> )	248.12 $\pm$ 58.57	166–428	0.24	258.44 $\pm$ 70	167–413	0.27	343.69 $\pm$ 56.73	262–459	0.17
Number of species	5.00 $\pm$ 1.02	3–7	0.20	6.30 $\pm$ 2.02	2–11	0.32	7.80 $\pm$ 3.17	0–13	0.41
Total abundance	25.77 $\pm$ 11.04	7–67	0.43	43.98 $\pm$ 16.08	6–76	0.37	72.30 $\pm$ 63.66	0–303	0.88
Margalef's richness index	1.28 $\pm$ 0.33	0.7–2.3	0.26	1.42 $\pm$ 0.49	0.5–2.54	0.35	1.70 $\pm$ 0.58	0–2.7	0.34
Shannon-Wiener diversity index	1.31 $\pm$ 0.21	0.8–1.7	0.16	1.23 $\pm$ 0.36	0.4–1.9	0.29	1.37 $\pm$ 0.49	0–2	0.36
Total biomass (g wet wt)	1.88 $\pm$ 1.25	0–5.1	0.66	3.80 $\pm$ 1.50	0.1–7.8	0.39	3.39 $\pm$ 4.25	0–19.4	1.25

Table 2. Two-way nested ANOVA of environmental variables. p-values that still are significant at the 0.05 level after a sequential Bonferroni correction are in **bold** text

Source of variation	df	Mean Sq.	F	p
<b>Depth<sup>a</sup></b>				
Area	2	0.41	6.32	0.019
1000 m (Area)	9	0.06	7.27	<b>&lt;0.001</b>
Residual	36	0.01		
<b>Temperature<sup>b</sup></b>				
Area	2	44.46	1.82	0.216
1000 m (Area)	9	24.37	7.96	<b>&lt;0.001</b>
Residual	36	3.06		
<b>pH<sup>b,c</sup></b>				
Area	2	$1.4 \times 10^{-5}$	1.98	0.193
1000 m (Area)	9	$7.3 \times 10^{-6}$	4.64	<b>&lt;0.001</b>
Residual	36	$1.6 \times 10^{-6}$		
<b>Salinity<sup>b,c</sup></b>				
Area	2	$6.3 \times 10^{-4}$	116.05	<b>&lt;0.001</b>
1000 m (Area)	9	$5.4 \times 10^{-6}$	1.31	0.267
Residual	36	$4.1 \times 10^{-6}$		
<b>Oxygen<sup>b,c</sup></b>				
Area	2	$1.0 \times 10^{-5}$	0.25	0.784
1000 m (Area)	9	$4.0 \times 10^{-5}$	4.49	<b>&lt;0.001</b>
Residual	36	$9.0 \times 10^{-6}$		
<b>Organic content<sup>a,b</sup></b>				
Area	2	12.54	182.18	<b>&lt;0.001</b>
1000 m (Area)	9	0.07	1.83	0.096
Residual	36	0.04		
<b>P</b>				
Area	2	101.00	9.07	0.007
1000 m (Area)	9	11.13	2.48	0.026
Residual	36	4.50		
<b>N</b>				
Area	2	44016.27	10.20	<b>0.005</b>
1000 m (Area)	9	4317.29	1.16	0.352
Residual	36	3736.21		
<sup>a</sup> Data transformed to $\ln(x + 1)$ before analysis				
<sup>b</sup> Data could not be transformed to reach heterogeneity of variances (a lower significance level of $p < 0.001$ was demanded)				
<sup>c</sup> $-(x + 1)^{-1}$ transformation				

solute differences were generally smaller. Area I was an especially homogeneous area with regard to most environmental variables, and Area III was by far the most variable, especially with regard to water depth, temperature and organic content (Tables 1 & 2).

The variance components for most environmental variables showed that most variation was present at the level of residuals (Table S1 in the supplement: [www.int-res.com/articles/suppl/m426p013\\_supp.pdf](http://www.int-res.com/articles/suppl/m426p013_supp.pdf)), i.e. the lowest examined level (100 m). This was consistently so for salinity, P and N. Temperature showed consistently more variability at the 1000 m scale, whereas the other variables demonstrated more inconsistent patterns.

A multivariate PCA-ordination, where principal component 1 (PC1) explained 48.2% and PC2 explained 19.5% of the variation in environmental data, illustrated the differences in environmental variables between the 3 areas (Fig. 2). As indicated in Fig. 2, Area III was especially low in organic content and had higher N levels and water depths than did the other areas, which are the 3 variables for which the differences are visualised fairly well along PC1, and Area II generally had higher organic content and lower P concentration than observed in Area I. Also Area III had the highest overall variability in environmental data of the 3 areas (as indicated in Fig. 2 by the more dispersed sampling stations) followed by Area II; Area I was the most homogeneous.

### Population and community data

Out of the 29 species (i.e. taxonomic groups) that were encountered in the 3 areas, only 10 were found in Area I (dominated by Oligochaeta, *Monoporeia affinis* Lindström, Chironomidae and *Macoma balthica* (L.)), whereas 19 were found in Area II (dominated by *M. affinis*, *Hydrobia* spp. and *M. balthica*) and a total of 23 in Area III (dominated by *Marenzelleria neglecta* Sikorski and Bick, Oligochaeta, *M. balthica* and *Mytilus edulis* L.) (Table 3). *M. balthica*, Oligochaeta, *M. affinis* and Chironomidae were found in more than half of the samples in all areas and can be characterised as the dominant species across all investigated areas. *Hydrobia* spp., *Halicryptus spinulosus* von Siebold, *Chironomus plumosus* (L.), *M. neglecta* and *Saduria entomon* (L.) were less dominant, but not rare in any area. The rest of the species were either rare in or absent from at least one of the investigated areas, where rare is defined as a singleton, i.e. a species present in one and only one sample (Krebs 1989). By means of this criterion and translating the species occurrences into percentages, 14% of the species (taxa) found in the whole study were categorized as rare (Table 3). For individual areas, the frequency of rare species was 9% in Area III, 10% in Area I and 21% in Area II.

The species observed (Sobs), Chao2 and Jackknife2 are plotted against the number of samples for all 3 Areas in Fig. 3. For Area I, 4 replicates would be required to include more than two-thirds of the species in the data set that are actually found in 64 replicate samples, and 42 replicates would be needed to include 95% of the species. For Area III, the corresponding numbers were very similar: 5 and 40 replicates, whereas for Area II, they were higher: 9 and 50 replicates (see Rumohr et al. 2001, where 10 and 53 required replicate samples, respectively, were estimated from 70 samples in a very small area in the

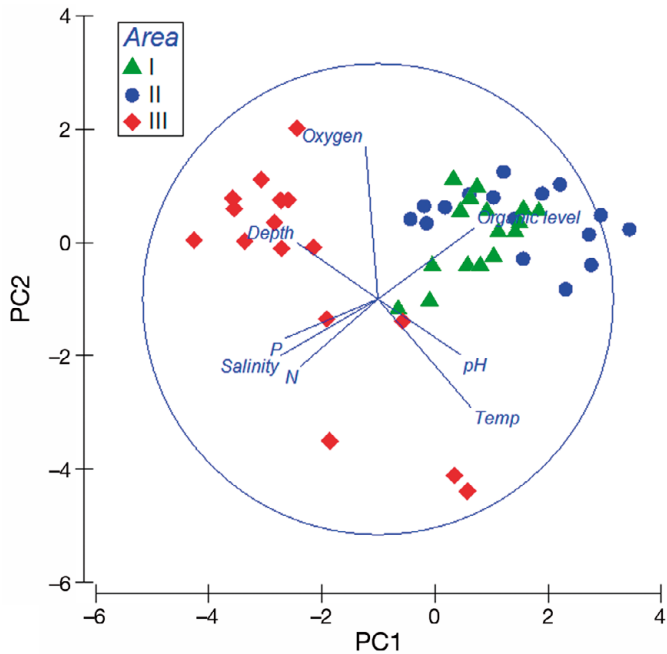


Fig. 2. PCA-ordination on environmental data registered once from all  $10 \times 10$  m grids (at the 100 m level) of all areas (16 environmental samples per area)

western Baltic Sea). The richness extrapolators Chao2 and Jackknife2, in relation to Sobs, tell us that Area II (Sobs end point for 64 samples = 19, Jackknife2 end point = 26 and Chao2 end point = 27) may possibly have contained more species than actually were sampled (there was a greater influence on the Chao2-curve by rare species in this area), whereas the corre-

sponding end point numbers for Area I (10, 11 and 10) and Area III (23, 25 and 24) possibly indicate that most species present in these areas during the sampling occasion also were actually encountered by means of a grid with 64 samples per area (Fig. 3).

With regard to community variables, the average numbers of both species and individuals were clearly higher in Area III compared with the 2 other areas as well as in Area II compared with Area I. Total biomass was also lowest in Area I (about 50% of the others). At the same time, the variability, measured as the CV (and seen as the range), was clearly higher for number of species, number of individuals and total biomass in Area III compared with the 2 other areas, but was not for diversity indices (Table 1).

When analysed by a 3-way nested ANOVA (Table 4) and as illustrated by the bar charts in Fig. 4a–f, the 3 areas differed significantly with regard to many community and population variables. The differences were especially clear for the abundance of *Marenzelleria neglecta* (higher in Area III), *Oligochaeta* (Area III > Area I > Area II) and *Macoma balthica* (Area II > Area I > Area III).

Regarding differences within the 3 areas themselves, i.e. comparing the biotic data area-specifically at the 1000 m level overall in the 3-way ANOVA and separately for each area with SNK-corrected *a posteriori* tests and similarly for 100 m grids nested in 1000 m, the absolute differences were generally smaller, with both Area I and Area II being especially homogeneous areas with regard to most community and population variables. Exceptions to this pattern were Shannon-Wiener diversity indices and the abundance of *Mono-*

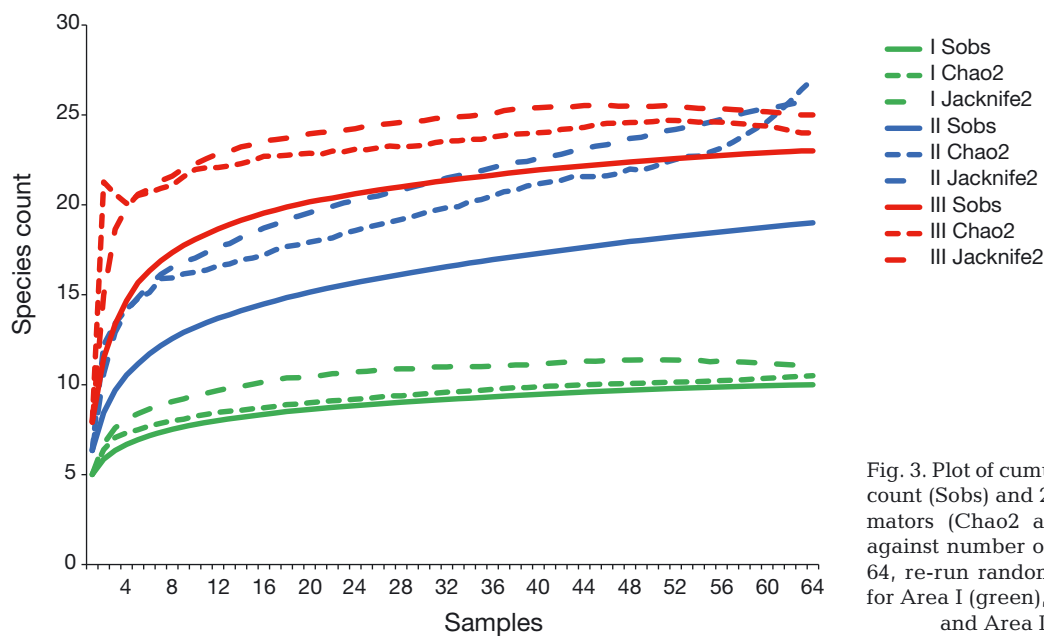


Fig. 3. Plot of cumulative species count (Sobs) and 2 richness estimators (Chao2 and Jackknife2) against number of samples ( $n = 64$ , re-run randomly 999 times) for Area I (green), Area II (blue) and Area III (red)

Table 3. Frequency (no. observations/no. samples) of each species in the whole data set and in the various areas. \*Rare species or singletons on the basis of frequency distribution (the species is only present in one of all samples of a group)

Taxon	Whole data set	Area I	Area II	Area III
<i>Macoma</i>	90.23	100.00	98.43	71.88
<i>Oligochaeta</i>	85.16	93.75	64.06	96.87
<i>Monoporeia</i>	80.47	95.31	87.50	43.75
Chironomidae	69.92	95.31	59.38	51.56
<i>Marenzelleria</i>	54.30	31.25	25.00	64.06
<i>Hydrobia</i>	34.38	3.12	81.25	51.56
<i>Mytilus</i>	29.69	1.56*	35.94	81.25
<i>Halicryptus</i>	28.91	6.25	81.25	26.56
<i>Chironomus</i>	21.09	56.25	14.06	3.12
<i>Saduria</i>	20.70	17.19	29.69	23.43
<i>Pygospio</i>	15.23		4.69	56.25
<i>Fabricia</i>	13.67			54.69
<i>Harmothoe</i>	8.98			35.94
<i>Jaera</i>	8.20			32.81
<i>Manayunkia</i>	7.03			28.12
<i>Paludestrina</i>	5.86		23.44	
Insecta	4.30		3.12	10.94
<i>Corophium</i>	3.91		1.56*	12.50
<i>Nereis</i>	3.52			9.38
<i>Gammarus</i>	3.52			14.06
<i>Mysis</i>	2.73			4.69
<i>Prostoma</i>	2.73		9.38	1.56*
<i>Limapontia</i>	1.56		6.25	
<i>Neomysis</i>	0.78			
<i>Theodoxus</i>	0.78			3.12
Trichoptera	0.39*		1.56*	
Araneida	0.39*		1.56*	
<i>Lymnaea</i>	0.39*		1.56*	
<i>Mya</i>	0.39*			1.56*
Total no. taxa	29	10	19	23
Percent rare species/taxa	13.79	10.00	21.05	8.69

*poreia affinis* and Chironomidae at the 1000 m level in Area I and at the 100 m level in Area II. For *M. affinis* in Area I, these differences were especially interesting since they had a clear east–west pattern with western subareas, numbers 1 to 4 and 9 to 12, having clearly lower abundance than eastern subareas, numbers 5 to 8 and 13 to 16 (Fig. 4e). As for environmental data, Area III was also the most variable area for biological data, especially with regard to total number of species, total abundance, total biomass and the abundance and biomass of *Macoma balthica* (Table 4, Fig. 4).

The contribution of each spatial scale to the total variance was calculated area-specifically for each of the dominating taxa as well as for some community summary variables (Table 5). When analysing these variance components, it was revealed that most variability was present already at the smallest level investigated, i.e. at 10 m or at the level of the residuals, for

Table 4. Three-way nested ANOVA for summary variables and dominating species. p-values that are still significant at the 0.05 level after a sequential Bonferroni correction are in **bold text**

Source of variation	df	Mean square	F	p
<b>Total number of species<sup>a</sup></b>				
Area	2	125.38	10.36	0.005
1000 m (Area)	9	12.11	1.32	0.261
100 m (1000 m [Area])	36	9.17	2.55	<b>&lt;0.001</b>
Residual	144	3.60		
<b>Total abundance<sup>a,b</sup></b>				
Area	2	118.56	10.69	0.004
1000 m (Area)	9	11.10	1.58	0.157
100 m (1000 m [Area])	36	7.00	1.59	0.030
Residual	144	4.41		
<b>Total biomass<sup>c</sup></b>				
Area	2	0.78	4.70	0.040
1000 m (Area)	9	0.17	3.68	<b>0.002</b>
100 m (1000 m [Area])	36	0.04	1.06	0.387
Residual	144	0.04		
<b>Shannon-Wiener diversity index<sup>a</sup></b>				
Area	2	0.31	1.72	0.233
1000 m (Area)	9	0.18	0.58	0.807
100 m (1000 m [Area])	36	0.31	3.33	<b>&lt;0.001</b>
Residual	144	0.09		
<b><i>Marenzelleria</i> abundance<sup>d</sup></b>				
Area	2	10.68	20.17	<b>&lt;0.001</b>
1000 m (Area)	9	0.53	0.92	0.518
100 m (1000 m [Area])	36	0.58	1.51	0.046
Residual	144	0.38		
<b><i>Oligochaeta</i> abundance<sup>d</sup></b>				
Area	2	47.15	35.20	<b>&lt;0.001</b>
1000 m (Area)	9	1.34	0.93	0.508
100 m (1000 m [Area])	36	1.44	1.65	0.021
Residual	144	0.87		
<b><i>Monoporeia</i> abundance<sup>d</sup></b>				
Area	2	16.84	4.72	0.040
1000 m (Area)	9	3.59	3.75	<b>0.002</b>
100 m (1000 m [Area])	36	0.95	2.76	<b>&lt;0.001</b>
Residual	144	0.34		
<b>Chironomidae abundance<sup>d</sup></b>				
Area	2	20.24	5.26	0.031
1000 m (Area)	9	3.85	4.42	<b>&lt;0.001</b>
100 m (1000 m [Area])	36	0.87	2.48	<b>&lt;0.001</b>
Residual	144	0.35		
<b><i>Macoma</i> abundance<sup>a,c</sup></b>				
Area	2	2.32	21.36	<b>&lt;0.001</b>
1000 m (Area)	9	0.11	1.01	0.451
100 m (1000 m [Area])	36	0.11	2.71	<b>&lt;0.001</b>
Residual	144	0.04		
<b><i>Macoma</i> biomass<sup>c</sup></b>				
Area	2	3.88	35.64	<b>&lt;0.001</b>
1000 m (Area)	9	0.11	2.33	0.035
100 m (1000 m [Area])	36	0.05	1.14	0.290
Residual	144	0.04		

<sup>a</sup>Data could not be transformed to reach heterogeneity of variances (a lower significance level of  $p < 0.001$  was demanded)

<sup>b</sup>Data were transformed to  $(x + 1)^{0.5}$  before analysis

<sup>c</sup> $-(x + 1)^{-1}$  transformation

<sup>d</sup> $\ln(x + 1)$  transformation



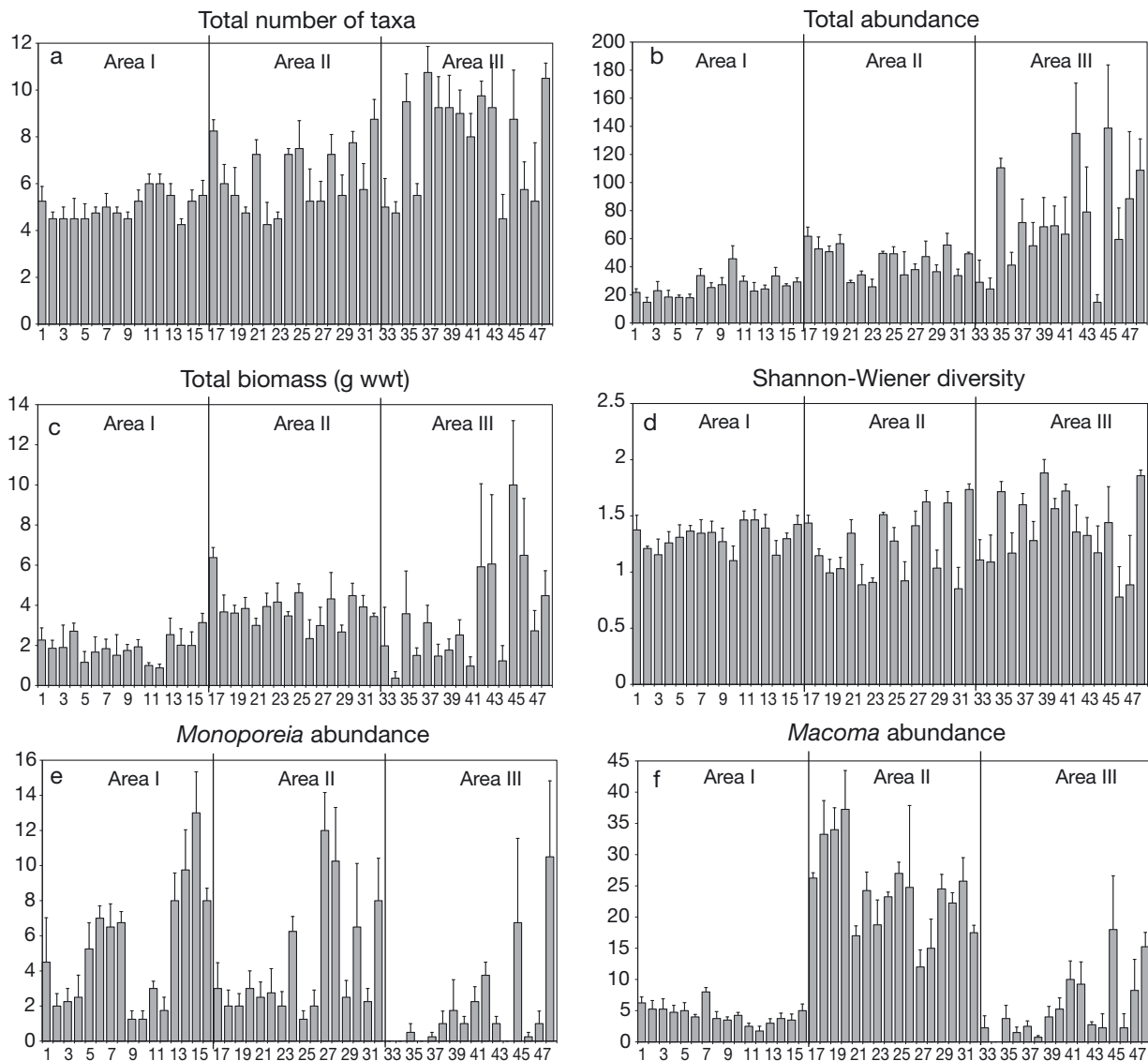


Fig. 4a–f. Mean  $\pm$  SE for number of taxa and individuals, total biomass, Shannon-Wiener diversity index and abundance of *Monoporeia affinis* and *Macoma balthica* per core ( $n = 4$ ) at the 10 m level (variability seen as SE), at the 100 m level (variability seen as bars) and at the 1000 m level (variability seen as groups of 4 bars) in each of 3 areas (Area I to the left, Area II in the middle and Area III to the right)

most of the variables and areas (28 out of 30). The only exceptions were abundance of *Monoporeia affinis* and Chironomidae in Area I (higher variability at the level of 1000 m). For *M. affinis*, this was well in accordance with the clear east–west differences in the distribution patterns shown in Fig. 4e and the significance values presented in Table 4.

Spearman's rank correlation analyses on abundance of dominant species over the entire area ( $N = 192$ ) and area-specifically ( $N = 64$ ) revealed many significant positive and negative correlations. Significant positive correlations over the entire area were found among *Macoma balthica*, *Monoporeia affinis* and *Hydrobia*

sp., between *M. affinis* and Chironomidae and among Oligochaeta, *Marenzelleria neglecta* and *Mytilus edulis* (Table S2 in supplement). Significant negative correlations were found between *M. affinis* and *M. edulis* as well as between Oligochaeta and *Hydrobia* spp. Area-specifically, there were very few highly significant correlations outside Area III. In Area II, Chironomidae showed highly significant positive correlations with *M. affinis* and *M. edulis*, whereas within Area I, there were no significant correlations.

Non-parametric multivariate techniques further confirmed the differences in community structure among the 3 areas and provided information about where most

Table 5. Variance estimates derived from 2-way nested ANOVAs of summary variables and dominating species in the respective areas. The highest value from each analysis is in **bold text**

Area/source of variation	Total no. of species	Total abundance	Total biomass	Shannon-Wiener diversity index	<i>Marenzelleria</i> abundance	Oligochaeta abundance	<i>Monoporeia</i> abundance	Chironomidae abundance	<i>Macoma</i> abundance	<i>Macoma</i> biomass
<b>Area I</b>										
1000 m	0.06	15.56	0.16	0.00	0.00	2.24	<b>12.39</b>	<b>12.08</b>	0.90	0.16
100 m (1000 m)	0.02	24.56	0.00	0.00	0.04	21.62	0.23	0.00	0.62	0.00
Residual	<b>0.99</b>	<b>86.65</b>	<b>1.44</b>	<b>0.04</b>	<b>1.77</b>	<b>76.46</b>	7.21	8.81	<b>3.65</b>	<b>1.24</b>
<b>Area II</b>										
1000 m	0.00	56.04	0.00	0.00	0.00	0.80	0.19	2.12	29.22	0.00
100 m (1000 m)	0.61	25.65	0.00	0.02	0.01	0.98	8.02	3.00	4.00	0.07
Residual	<b>3.62</b>	<b>191.34</b>	<b>2.31</b>	<b>0.11</b>	<b>0.83</b>	<b>11.14</b>	<b>11.01</b>	<b>4.83</b>	<b>84.22</b>	<b>2.01</b>
<b>Area III</b>										
1000 m	0.83	68.82	2.13	0.00	0.00	33.49	2.12	2.35	11.62	0.23
100 m (1000 m)	2.63	406.30	1.20	0.07	0.00	49.96	3.72	1.88	8.93	0.03
Residual	<b>6.93</b>	<b>3613.59</b>	<b>15.28</b>	<b>0.23</b>	<b>13.65</b>	<b>389.84</b>	<b>11.97</b>	<b>8.02</b>	<b>37.95</b>	<b>1.95</b>

variability was present as well as the influence of rare species. NMDS ordinations of untransformed abundance and biomass data (most weight put on the dominant species) plotted out the samples from Area I to the upper right of the ordination scheme, whereas the samples from Area II fell out to the lower right and were a bit less dispersed in space (i.e. had internally more similar community structures). The samples from Area III, in turn, all clustered out to the left of the plot and were clearly more widely dispersed internally than the samples from the other areas (Fig. 5a,c). This was verified by MVDISP analyses with relative indices of multivariate dispersion for abundance data at 0.891 (Area I), 0.7 (Area II) and 1.422 (Area III) and for biomass data at 0.989 (Area I), 0.598 (Area II) and 1.427 (Area III). After a fourth root transformation (which put more weight on rare species), the relative multivariate dispersion shifted and the samples from Area II, the most prominent area with regard to the amount of rare species, were now clearly more dispersed than the samples from Area I, both when visualised in the ordination scheme (Fig. 5b,d) and when the results from the MVDISP analysis are listed—abundance: 0.545 (Area I), 1.019 (Area II), 1.45 (Area III); biomass: 0.788 (Area I), 0.820 (Area II), 1.404 (Area III). At the same time, differences among areas also seemed to be accentuated with the transformation (Fig. 5b,d compared with Fig. 5a,c).

PERMANOVA revealed significant differences for most investigated levels for both species' abundance and biomass data, regardless of transformation, while also demonstrating most variability at the smallest level (10 m) for the variance components, i.e. in 11 out of 12 cases (results not shown).

SIMPER-analyses of untransformed data listed the species contributing most to the similarities within

areas (Table S3a in supplement) and the species contributing most to dissimilarities between different areas (Table S3b). For all areas, the most common species, *Macoma balthica*, was an important contributor to within similarity, whereas the other major contributors to similarity were different among the 3 areas, i.e. Oligochaeta, *Monoporeia affinis* and Chironomidae in Area I, *Hydrobia* spp. in Area II and Oligochaeta and *Mytilus edulis* in Area III (Table S3a). From Table S3a, we can see that the internal homogeneity was much lower for Area III (33%) compared with the other areas: Area I (54%) and Area II (59%). With regard to differences between areas (Table S3b), *M. balthica* was clearly the most important discriminator (Area II > Area I > Area III). The average dissimilarities in the spatial analysis went from 82% (Area II to Area III, most different), to 77% (Area I to Area III) and to 70% (Area I to Area II, most similar), which closely reflects the relative differences in abiotic variables.

#### Linking of biological patterns to the environment

Overall (N = 192) and area-specific (N = 64) Spearman's rank correlations revealed significant correlations between environmental variables and biological community and population variables (Table S4a,b in the supplement). The strongest positive correlation in the overall analysis was found between the salinity and the abundance of *Mytilus edulis* and the strongest negative correlation was found between water depth and the abundance of *Hydrobia* spp., but many other investigated correlations were also highly significant at this level of precision (Table S4a). Area-specific analyses gave a more detailed view of possible interre-

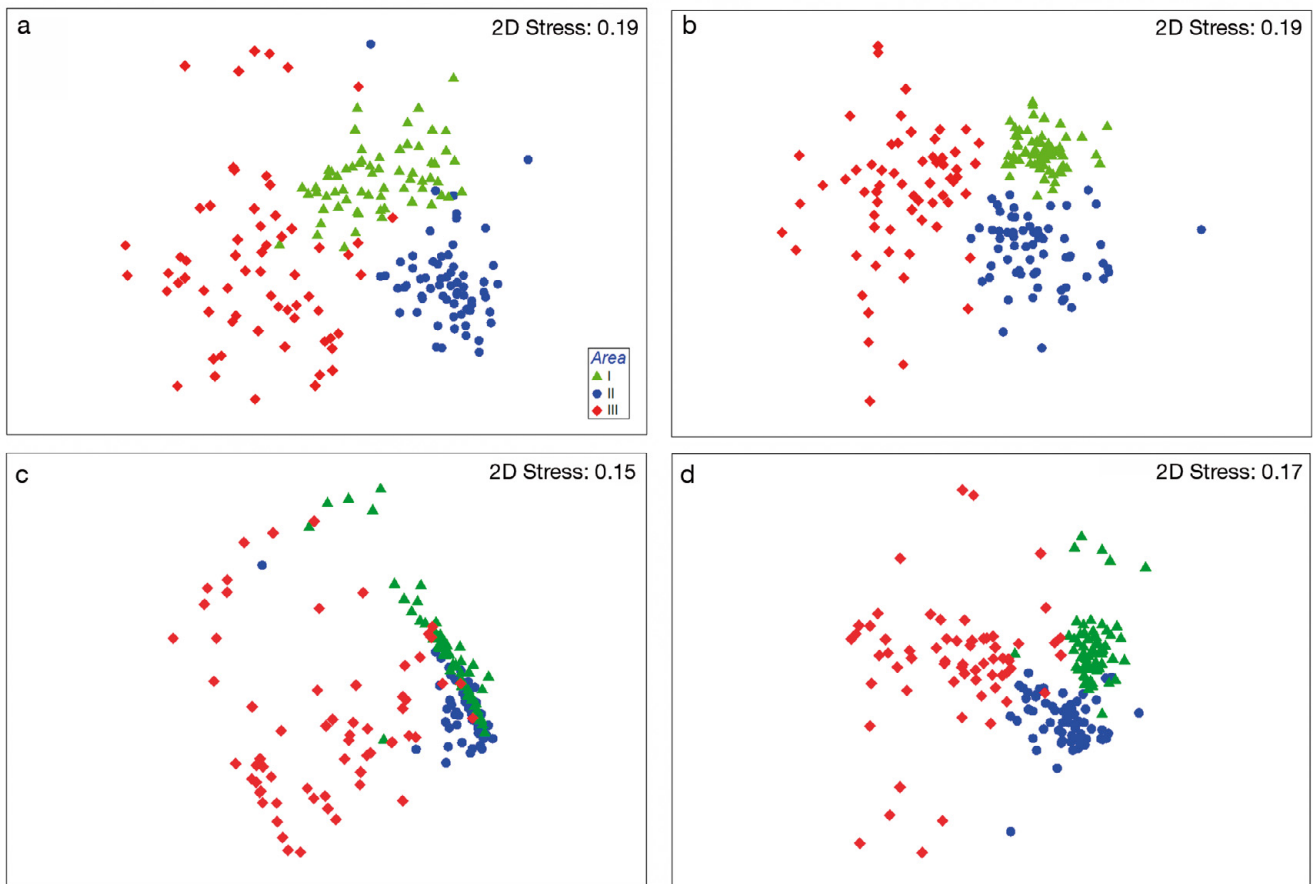


Fig. 5. NMDS-ordination of (a) untransformed macrofauna abundance data, (b) fourth root transformed macrofauna abundance data, (c) untransformed macrofauna biomass data and (d) fourth root transformed macrofauna biomass data using all samples from all areas

relationships (Table S4b). In Area I, negative correlations between the abundance of *Monoporeia affinis* and water depth and organic level stand out. In Area II, the total abundance of organisms was positively correlated with temperature, pH and oxygen and negatively correlated with P. In Area III, there were strong positive correlations between water depth and the abundance and biomass of *Macoma balthica* and between P and number of species as well as strong negative correlations between temperature and pH and the abundance of *M. affinis* and *M. balthica*.

To further explore the connectivity between the biotic and abiotic data in a multivariate sense, and because none of the abiotic variables expressed a very strong correlation with each other (Spearman's rank correlation coefficient always  $< 0.63$ ), several BIO-ENV analyses were performed (results not shown) that included all environmental variables. These BIO-ENV analyses were first run separately on abundance and biomass data, as well as untransformed and fourth root

transformed data. For abundance data, the variables water depth, salinity, organic level and temperature appeared as especially important structuring variables with reasonably high correlation coefficients ( $r_s > 0.50$  when comparing the abiotic and biotic similarity matrices). For biomass data, the oxygen levels seemed to be more important than water depth and temperature. There were no clear differences due to the level of data transformation. Area-specific BIO-ENV analyses on abundance data showed generally lower Spearman coefficients for the matching of abiotic and biotic data than did the total analysis. In Area I, depth, temperature and oxygen were most influential when untransformed biotic data were used, and temperature and pH were most influential when fourth-root transformed data were used. In Area II, the most influential environmental variables seemed to have been water depth, P and N. In Area III, the best subsets of variables describing the observed biotic community structure included temperature, pH, P and N.

## DISCUSSION

These areas of low diversity in the northern Baltic Sea show clear species distribution patterns driven by the local environmental conditions, but a substantial biotic influence on the pattern is also evident. The latter is seen as the high variability present at the smallest sampled level, i.e. at 10 m (as shown in the analyses of components of variance) or at the level at which most organisms complete their life cycle. Our understanding of spatial variability (or 'stability') thus seems to be influenced by both the habitat and the species composition and maybe by the latter factor to a greater extent than first expected. In terms of our original hypotheses, there were clear differences among and within study areas. The influence of the habitat can best be seen in the most exposed area, Area III, where the highest heterogeneity in abiotic variables also demonstrated the highest heterogeneity and biodiversity in zoobenthos. The reverse was seen for the least variable area, Area I, in addition to somewhat contradictory results at various scales for some study variables. The presence of other, more fine-scale structuring aspects may be exemplified by the high variability found at the smallest investigated scale, i.e. at 10 m. This large amount of variation at the smallest scale also indicates that many processes affecting biota are operating at very fine scales. With the exception of the abundance of *Monoporeia affinis* and Chironomidae in Area I, which showed the highest variability at the 1000 m level, all other dominating species as well as all the investigated community variables pointed to the smallest sampled scale (10 m) as being the level that contributed most to the total variability. However, when specifically focusing on the 3 areas, the differences in average values of environmental variables clearly also shaped the communities, i.e. different environmental characteristics caused different zoobenthic community structures (Bonsdorff et al. 2003, Laine 2003). A differing variability in environmental variables is reflected in both the animal community and the population data (Area III > Area II > Area I). The study also showed that many replicates may be needed to get reliable pictures of the biodiversity and species composition of an area, despite a relatively low overall biodiversity (here the role of the rare species is important).

A total number of 29 species (taxonomic groups) were encountered in 192 samples from 3 areas (23 in Area III, 19 in Area II and only 10 in Area I), which is a typical number for the region in question (Bonsdorff et al. 2003, Laine 2003, Perus & Bonsdorff 2004). Only 4 taxa were dominant across all areas, whereas 5 other taxa were not rare in any area, while the rest of the species were either absent from or rare (single-

tons) in at least one of the investigated areas. By directly or indirectly influencing both Jackknife2 and Chao2 estimates of species richness, the number of singletons allows some prediction of the total number of species in the community to be made. Our values, which ranged from 9 to 21% singletons, are well in accordance with the 16% reported by Rumohr et al. (2001), but clearly lower than the 27% reported by Ellingsen (2001), the 30% reported by Cao et al. (1998) and the >30% reported by Holme (1953). These numbers are not fully comparable, since rare species have a low probability of being recorded, and thus their characterisation and observed distribution is directly linked to sampling intensity (Brown 1984, Gray et al. 2005). With the exception of the 70 samples in Rumohr et al. (2001), the sampling intensity in the other studies mentioned was clearly lower than in our study (total sample numbers ranged between 4 and 24). Therefore, a number of those 'rare' species most probably occur at more sites (Glover et al. 2002). With regard to the species richness estimators, it was only in Area II where the number of rare species made a remarkable influence on Jackknife2 and Chao2 (Fig. 3). The area-specific differences in species richness and the role of less common species are also accentuated when comparing Area I with Area II in the NMDS-ordinations (Fig. 5) and the change taking place in the sequence of the indices of multivariate dispersion with transformation intensity. When the more severe fourth-root transformation is used (i.e. more weight is put on rare species), a switch occurs, and Area I has the lowest index value.

An area-specific analysis of the community and population variables showed clear differences in the averages and in the variability. For zoobenthic abundance data, for example, the coefficients of variation were 0.43 (Area I), 0.37 (Area II) and 0.88 (Area III) or fully comparable to a CV of 0.39 reported in Rumohr et al. (2001) and a range of 0.34 to 0.96 by Kraufvelin (1998). Generally, the internal variability in biotic variables also followed the same sequence as for abiotic variables: Area III (highest), Area II (medium) and Area I (lowest). These differences in zoobenthos were highly significant both in the ANOVAs and PERMANOVAs and can probably be related largely to environmental differences. When attempting to elucidate possible biotic structuring forces by examining single species, most species correlated positively with each other, which again supports the existence of common environmental preferences. There were no significant negative correlations between species in any of the area-specific analyses and most probably very few strong species interactions were operating. Biotic forces other than strict competition or predation, such as differences in mobility, recruitment success and intra-

specific factors, may thus have played a part. Also, the correlation analyses were only performed for dominant species and significant correlations can be hard to verify for rare species, a fact that should not be used to rule out their specific importance.

Most variability in the biotic data was present at the smallest investigated level (10 m or the residuals), suggesting that additional variability exists at even smaller spatial scales than the ones that were sampled (Table 5). Perhaps sampling by SCUBA diving would have been preferable to enable sampling at the 1 m scale; this precision is impossible to reach by means of remote sampling (Bergström et al. 2002). Such a measure would at the same time have multiplied the workload both in connection with sampling and with sample analyses, but since the recruitment of many animal species may vary at scales of centimetres to even millimetres (e.g. Olafsson et al. 1994), such an extra input might have been worthwhile. However, interpretation of variance components is not straightforward and should be done with caution. It is, for example, difficult to compare the relative importance of different spatial scales across taxa when the proportion of the total variance contributed by the residual variances in the spatial analysis differed by a factor of almost 3 (24 to 68 %) (see Morrisey et al. 1992).

An existence of considerable spatial variability in biota at even the smallest measured level, (in our case at 10 m), the fact that little additional variation could be attributed to the larger scales (100 to 1000 m) and that this pattern was fairly consistent across variables and areas, has also been reported from other unpolluted marine benthic habitats. These studies include soft sediments (Thrush 1986a, Volckaert 1987, Edgar & Barrett 2002, Stark et al. 2003, Norén & Lindegarth 2005, Commito et al. 2006, Chapman et al. 2010), hard-bottom intertidal (Underwood & Chapman 1996, 1998, Chapman 2002, 2005), hard-bottom subtidal (Thrush 1986b, Glasby 1998, Anderson et al. 2005, Dethier & Schoch 2005, Fraschetti et al. 2005) and fine sands (Kendall & Widdicombe 1999). A registration of most variability at the smallest investigated scales can probably be attributed to behavioural responses to small-scale patches of microhabitat (Underwood & Chapman 1996, Chapman et al. 2010). However, considerable variability has also been found at the scale of 50 to 100 m in intertidal soft sediments (Ysebaert & Herman 2002) and on intertidal hard bottoms (Underwood & Chapman 1996) as well as in heavily bioturbated sandy mud (Kendall & Widdicombe 1999) and edges between sandy and vegetated habitats (Boström et al. 2010). Such findings correspond well to our results on the differences in abundance of *Monoporeia affinis* in Area I and they may be due to both variations in the physical environment and to variation in recruitment, species

interactions and/or mortality or avoidance (e.g. Underwood & Chapman 1996).

The Spearman's rank correlation analyses for linking biotic to abiotic data were not capable of providing many answers, despite quite a few significant correlations over the total data set as well as area-specifically (Table S4). Perhaps the most interesting significant correlation is the negative one between the abundance of *Monoporeia affinis* and organic level in Area I, which can partly explain the strong east–west patterns in the distribution of this species. At the same time, this seems to exemplify a habitat driven biotic response at an intermediate scale level, i.e. within less than 1000 m, and may contribute to the discussion about the recent decrease in *M. affinis* in many parts of the Baltic Sea (Wiklund et al. 2008, 2009).

The BIO-ENV analyses further helped to explore the connectivity between the environment and the biological data in a multivariate context. In the overall analyses, many of the same environmental variables were identified as important contributors as was also seen in the Spearman's rank correlation analysis, i.e. water depth, salinity and organic level stood out as the potentially most important contributors to biotic community pattern. The area-specific analyses showed lower Spearman coefficients for the matching of abiotic and biotic data matrices than did the full analysis, with partly conflicting results. Water depth and nutrients were most important in Area II; temperature, pH and nutrients were most important in Area III; whereas in Area I, many environmental variables (all but N and P) seemed to have contributed to the species patterns.

The interpretation of the relative roles of various abiotic or biotic structuring forces is not straightforward. When interpreting both the correlation analysis and the BIO-ENV analysis one must bear in mind that other, unregistered, environmental variables also might have played a role. In this case, there is no overall or site-specific information about water currents, sedimentation, sediment particle size, microalgae, meiofauna and the possible presence of pollutants, for example. Also, abiotic data were only recorded from 16 sampling points compared with 64 points for the fauna. Therefore, some error was always introduced when a considerable amount of biotic variability from 4 samples was linked with environmental data from only one sample. On the whole, the environmental variables do not seem to have constituted very strong structuring forces over the entire area and even less so within the areas. This indicates that biotic interactions may have been rather important from the level of a single area (1000 m) downwards across the studied scales.

When the main findings are put into perspective, it must be kept in mind that scale is central and different

abiotic and biotic factors are operating at different scales, alone or interactively, to structure benthic infauna. The present study once again emphasizes the importance of cross-scale studies. Nested, replicated sampling designs facilitate interpretation (Underwood & Chapman 1998, Ysebaert & Herman 2002), and species distributions across scales must be considered in environmental monitoring (Morrisey et al. 1992, Hewitt & Thrush 2007), the assessment of the effects of natural and anthropogenic perturbations (Thrush et al. 1997, Menge et al. 2003) and marine conservation (Ellingsen 2001). Also, studies carried out at a particular scale or on a particular component of the benthos are unlikely to be successful in predicting relationships at other scales or for other components of the benthos (Morrisey et al. 1992, Kraufvelin 1999, Somerfield & Gage 2000, Hewitt et al. 2007). Furthermore, it must be stressed that, although generalities do exist, our understanding of communities is site specific. The challenge is to learn to distinguish one from the other. Finally community structure is not driven exclusively by dominant species and, as demonstrated in this study, rare species are also important. In summary, owing to the increased demands of improved interpretation of monitoring programs, detection of anthropogenic impact and ecosystem management in an everchanging variable world, better general refinement of sampling methodology, as well as better incorporation of scaling issues into questions of biodiversity and conservation, are needed (Hewitt et al. 2007). Also, relating biotic patterns to the actual spatial structural aspects of soft-sediment assemblages (Zajac 2008) should not be forgotten. There is still much room for more scale-concerned studies in the field of marine ecology.

*Acknowledgements.* This work was done within the IMAGINE-project funded by the Academy of Finland. Financial support was also received from Åbo Akademi University. Three anonymous reviewers and the editor A. J. Underwood provided many insightful comments that greatly improved the paper.

#### LITERATURE CITED

- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 26:32–46
- Anderson MJ, Connell SD, Gillanders BM, Diebel CE, Blom WM, Saunders JE, Landers TJ (2005) Relationships between taxonomic resolution and spatial scales of multivariate variation. *J Anim Ecol* 74:636–646
- Anderson MJ, Gorley RN, Clarke KR (2008) PERMANOVA+ for PRIMER: guide to software and statistical methods. PRIMER-E, Plymouth
- Andrew NL, Mapstone BD (1989) Sampling and the description of spatial patterns in marine ecology. *Oceanogr Mar Biol Annu Rev* 25:39–90
- Barry JP, Dayton PK (1991) Physical heterogeneity and the organization of marine communities. In: Kolasa J, Pickett STA (eds) *Ecological heterogeneity*. Springer-Verlag, New York, p 270–320
- Bergström U, Englund G, Bonsdorff E (2002) Small-scale spatial structure of Baltic Sea zoobenthos—inferring processes from patterns. *J Exp Mar Biol Ecol* 281:123–136
- Bonsdorff E (2006) Zoobenthic diversity gradients in the Baltic Sea: continuous post-glacial succession in a stressed ecosystem. *J Exp Mar Biol Ecol* 330:383–391
- Bonsdorff E, Blomqvist EM (1993) Biotic couplings on shallow water soft bottoms—examples from the northern Baltic Sea. *Oceanogr Mar Biol Annu Rev* 31:153–176
- Bonsdorff E, Pearson TH (1999) Variation in the sublittoral macrobenthos of the Baltic Sea along environmental gradients: a functional group approach. *Aust J Ecol* 24:312–326
- Bonsdorff E, Blomqvist EM, Mattila J, Norkko A (1997) Coastal eutrophication: causes, consequences and perspectives in the archipelago areas of the northern Baltic Sea. *Estuar Coast Shelf Sci* 44:63–72
- Bonsdorff E, Laine AO, Hänninen JH, Vuorinen I, Norkko A (2003) Zoobenthos of the outer archipelago waters (N Baltic Sea)—the importance of local conditions for spatial distribution patterns. *Boreal Environ Res* 8:135–145
- Borcard D, Legendre P, Drapeau P (1992) Partialling out the spatial component of ecological variation. *Ecology* 73:1045–1055
- Boström C, Törnroos A, Bonsdorff E (2010) Invertebrate dispersal and habitat heterogeneity: expression of biological traits in a seagrass landscape. *J Exp Mar Biol Ecol* 390:106–117
- Brown JH (1984) On the relationship between abundance and distribution of species. *Am Nat* 124:255–279
- Burnham KP, Overton WS (1979) Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60:927–936
- Cao Y, Williams DD, Williams NE (1998) How important are rare species in aquatic community ecology and bioassessment? *Limnol Oceanogr* 43:1403–1409
- Chao A (1987) Estimating the population size for capture–recapture data with unequal catchability. *Biometrics* 43:783–791
- Chapman MG (2002) Patterns of spatial and temporal variation of macrofauna under boulders in a sheltered boulder field. *Austral Ecol* 27:211–228
- Chapman MG (2005) Molluscs and echinoderms under boulders: tests of generality of patterns of occurrence. *J Exp Mar Biol Ecol* 325:65–83
- Chapman MG, Tolhurst TJ, Murphy RJ, Underwood AJ (2010) Complex and inconsistent patterns of variation in benthos, micro-algae and sediment over multiple spatial scales. *Mar Ecol Prog Ser* 398:33–47
- Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* 18:117–143
- Clarke KR, Ainsworth M (1993) A method of linking multivariate community structure to environmental variables. *Mar Ecol Prog Ser* 92:205–219
- Clarke KR, Gorley RN (2006) PRIMER v6: user manual/tutorial. PRIMER-E, Plymouth
- Cole RG, Healy TR, Wood ML, Foster DM (2001) Statistical analysis of spatial patterns: a comparison of grid and hierarchical sampling approaches. *Environ Monit Assess* 69:85–99
- Commito JA, Dow WE, Grupe BM (2006) Hierarchical spatial structure in soft bottom mussel beds. *J Exp Mar Biol Ecol* 330:27–37
- Dethier MN, Schoch GC (2005) The consequences of scale: assessing the distribution of benthic populations in a com-

- plex estuarine fjord. *Estuar Coast Shelf Sci* 62:253–270
- Edgar GJ, Barrett NS (2002) Benthic macrofauna in Tasmanian estuaries: scales of distribution and relationships with environmental variables. *J Exp Mar Biol Ecol* 270:1–24
- Ellingsen KE (2001) Biodiversity of a continental shelf soft-sediment macrobenthos community. *Mar Ecol Prog Ser* 218:1–15
- Ellingsen KE, Hewitt JE, Thrush SF (2007) Rare species, habitat diversity and functional redundancy in marine benthos. *J Sea Res* 58:291–301
- Fletcher DJ, Underwood AJ (2002) How to cope with negative estimates of components of variance in ecological field studies. *J Exp Mar Biol Ecol* 273:89–95
- Fontana G, Uglund KI, Gray JS, Abbiati M (2008) Influence of rare species on beta diversity estimates in marine benthic assemblages. *J Exp Mar Biol Ecol* 366:104–108
- Fraschetti S, Terlizzi A, Benedetti-Cecchi L (2005) Patterns of distribution of marine assemblages from rocky shores: evidence of relevant scales of variation. *Mar Ecol Prog Ser* 296:13–29
- Glasby TM (1998) Estimating spatial variability in developing assemblages of epibiota on subtidal hard substrata. *Mar Freshw Res* 49:429–437
- Glover AG, Smith CR, Paterson GLJ, Wilson GDF, Hawkins L, Shearer M (2002) Polychaete species diversity in the central Pacific abyss: local and regional patterns, and relationships with productivity. *Mar Ecol Prog Ser* 240:157–170
- Gray JS (1974) Animal–sediment relationships. *Oceanogr Mar Biol Annu Rev* 12:223–261
- Gray JS, Bjørgesæter A, Uglund KI (2005) The impact of rare species on natural assemblages. *J Anim Ecol* 74:1131–1139
- Gustafsson B, Westman P (2002) On the causes for salinity variations in the Baltic Sea during the last 8500 years. *Paleoceanography* 17:1040–1054
- Hewitt JE, Thrush SF (2007) Effective long-term monitoring using spatially and temporally nested sampling. *Environ Monit Assess* 133:295–307
- Hewitt JE, Thrush SF, Cummings VJ, Pridmore RD (1996) Matching patterns with processes: predicting the effect of size and mobility on the spatial distribution of the bivalves *Macomona liliana* and *Austrovenus stutchburyi*. *Mar Ecol Prog Ser* 135:57–67
- Hewitt JE, Thrush SF, Dayton PK, Bonsdorff E (2007) The effect of spatial and temporal heterogeneity on the design and analysis of empirical studies on scale-dependent systems. *Am Nat* 169:398–408
- Hochberg Y (1988) A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 75:800–802
- Holme NA (1953) The biomass of the bottom fauna in the English channel off Plymouth. *J Mar Biol Assoc UK* 32:1–49
- Huston MA (1999) Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. *Oikos* 86:393–401
- Josefson A (2009) Additive partitioning of estuarine benthic macroinvertebrate diversity across multiple spatial scales. *Mar Ecol Prog Ser* 396:283–292
- Keeling MJ, Wilson HB, Pacala SW (2000) Re-interpreting space, time-lags, and functional responses in ecological models. *Science* 290:1758–1761
- Kendall MA, Widdicombe S (1999) Small scale patterns in the structure of macrofaunal assemblages of shallow sediments. *J Exp Mar Biol Ecol* 237:127–140
- Kraufvelin P (1998) Model ecosystem replicability challenged by the 'soft' reality of a hard bottom mesocosm. *J Exp Mar Biol Ecol* 222:247–267
- Kraufvelin P (1999) Baltic hard bottom mesocosms unplugged: replicability, repeatability and ecological realism examined by non-parametric multivariate techniques. *J Exp Mar Biol Ecol* 240:229–258
- Kraufvelin P, Sinisalo B, Leppäkoski E, Mattila J, Bonsdorff E (2001) Changes in zoobenthic community structure after pollution abatement from fish farms in the Archipelago Sea (N Baltic Sea). *Mar Environ Res* 51:229–245
- Krebs CJ (1989) *Ecological methodology*. Harper & Row, New York
- Laine AO (2003) Distribution of soft-bottom macrofauna in the deep open Baltic Sea in relation to environmental variability. *Estuar Coast Shelf Sci* 57:87–97
- Legendre P (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74:1659–1673
- Legendre P, Legendre L (1998) *Numerical ecology*. Elsevier, Amsterdam
- Legendre P, Dale MRT, Fortin MJ, Gurevitch J, Hohn M, Myers D (2002) The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography* 25:601–615
- Levin SA (1992) The problem of pattern and scale in ecology. *Ecology* 73:1943–1967
- Menge BA, Lubchenco J, Bracken MES, Chan F and others (2003) Coastal oceanography sets the pace of rocky intertidal community dynamics. *Ecology* 100:12229–12234
- Morrisey DJ, Howitt L, Underwood AJ, Stark JS (1992) Spatial variation in soft-sediment benthos. *Mar Ecol Prog Ser* 81:197–204
- Norén K, Lindegarth M (2005) Spatial, temporal and interactive variability of infauna in Swedish coastal sediments. *J Exp Mar Biol Ecol* 317:53–68
- Norkko A, Cummings VJ, Thrush SF, Hewitt JE, Hume T (2001) Local dispersal of juvenile bivalves: implications for sandflat ecology. *Mar Ecol Prog Ser* 212:131–144
- Olafsson EB, Peterson CH, Ambrose WGJ (1994) Does recruitment limitation structure populations and communities of macro-invertebrates in marine soft sediments: the relative significance of pre- and post-settlement processes. *Oceanogr Mar Biol Annu Rev* 32:65–109
- Perus J, Bonsdorff E (2004) Long-term changes in macrozoobenthos in the Åland Archipelago, northern Baltic Sea. *J Sea Res* 52:45–56
- Rumohr H, Karakassis I, Jensen JN (2001) Estimating species richness, abundance and diversity with 70 macrobenthic replicates in the Western Baltic Sea. *Mar Ecol Prog Ser* 214:103–110
- Schlacher TA, Newell P, Clavier J, Schlacher-Hoenlinger MA, Chevillon C, Britton J (1998) Soft-sediment benthic community structure in a coral-reef lagoon—the prominence of spatial heterogeneity and 'spot endemism'. *Mar Ecol Prog Ser* 174:159–174
- Seinä A, Peltola J (1991) Duration of the ice season and statistics of fast ice thickness along the Finnish coast 1961–1990. *Finn Mar Res* 258:1–46
- Somerfield PJ, Gage JD (2000) Community structure of the benthos in Scottish Sea-lochs. IV. Multivariate spatial pattern. *Mar Biol* 136:1133–1145
- Somerfield PJ, Clarke KR, Olsford F (2002) A comparison of the power of categorical and correlational tests applied to community ecology data from gradient studies. *J Anim Ecol* 71:581–593
- Sparrevik E, Leonardsson K (1995) Effects of large-scale *Saduria entomon* (Isopoda) on spatial distribution of their small *S. entomon* and *Monoporeia affinis* (Amphipoda) prey. *Oecologia* 101:177–184

- Stark JS, Riddle MJ, Simpson RD (2003) Human impacts in soft-sediment assemblages at Casey station, East Antarctica: spatial variation, taxonomic resolution and data transformation. *Austral Ecol* 28:287–304
- Thrush SF (1986a) The sublittoral macrobenthic community structure of an Irish sea-lough: effect of decomposing accumulations of seaweed. *J Exp Mar Biol Ecol* 96:199–212
- Thrush SF (1986b) Spatial heterogeneity in subtidal gravel generated by the pit-digging activities of *Cancer pagurus*. *Mar Ecol Prog Ser* 30:221–227
- Thrush SF (1991) Spatial patterns in soft-bottom communities. *Trends Ecol Evol* 6:75–79
- Thrush SF, Hewitt J, Pridmore RD (1989) Patterns in the spatial arrangements of polychaetes and bivalves in intertidal sandflats. *Mar Biol* 102:529–535
- Thrush SF, Schneider DC, Legendre P, Whitlatch RB and others (1997) Scaling up from experiments to complex ecological systems: where to next? *J Exp Mar Biol Ecol* 216: 243–254
- Underwood AJ (1981) Techniques of analysis of variance in experimental marine biology and ecology. *Oceanogr Mar Biol Annu Rev* 19:513–605
- Underwood AJ (1997) Experiments in ecology: their logical design and interpretation using analysis of variance. Cambridge University Press, Cambridge
- Underwood AJ, Chapman MG (1996) Scales of spatial patterns of distribution of intertidal invertebrates. *Oecologia* 107:212–224
- Underwood AJ, Chapman MG (1998) Spatial analyses of intertidal assemblages on sheltered rocky shores. *Aust J Ecol* 23:138–157
- Valanko S, Norkko A, Norkko J (2010) Strategies of post-larval dispersal in non-tidal soft-sediment communities. *J Exp Mar Biol Ecol* 384:51–60
- Voipio A (1981) The Baltic Sea. Elsevier Oceanography Series 30, Elsevier, Amsterdam
- Volckaert F (1987) Spatial pattern of soft-bottom Polychaeta off Nova Scotia, Canada. *Mar Biol* 93:627–639
- Wiens JA (1989) Spatial scaling in ecology. *Funct Ecol* 3: 385–397
- Wiklund AKE, Sundelin B, Rosa R (2008) Population decline of amphipod *Monoporeia affinis* in Northern Europe: consequence of food shortage and competition? *J Exp Mar Biol Ecol* 367:81–90
- Wiklund AKE, Dahlgren K, Sundelin B, Andersson A (2009) Effects of warming and shifts in pelagic food web structure on benthic productivity in a coastal marine system. *Mar Ecol Prog Ser* 396:13–25
- Ysebaert T, Herman PMJ (2002) Spatial and temporal variation in benthic macrofauna and relationships with environmental variables in an estuarine, intertidal soft-sediment environment. *Mar Ecol Prog Ser* 244: 105–124
- Zajac RN (2008) Challenges in marine, soft-sediment benthoscape ecology. *Landscape Ecol* 23:7–18

*Editorial responsibility: Antony Underwood, Sydney, Australia*

*Submitted: June 2, 2010; Accepted: January 13, 2011  
Proofs received from author(s): March 17, 2011*