Marine Environmental Research 154 (2020) 104867

Contents lists available at ScienceDirect

Marine Environmental Research

journal homepage: http://www.elsevier.com/locate/marenvrev

Does grab size influence sampled macrofauna composition? A test conducted on deep-sea communities in the northeast Atlantic

Børge Holte [1,*], Lene Buhl-Mortensen [1]

Institute of Marine Research (IMR), Nordnesgaten 50, 1005, Bergen, Norway

ARTICLE INFO

Keywords:
Benthos
Infauna
Grab size test
Monitoring strategy
Deep sea
North Atlantic

ABSTRACT

In deep-sea surveys, heavy gears are often preferred to effectively collect macro benthos while smaller samplers are sufficient in coastal and shallow areas. However, there are few comparative studies of the samples retained and results gained from different-sized grabs. To study the differences in sampling properties between a small (sampling area of 0.1 m^2) and a large (0.25 m^2) van Veen grab, 1 m^2 of seafloor was sampled with each of the grab sizes at four test sites in the Barents- and Norwegian Sea; one inside a fjord and three offshore, across a depth range of 287–963 m. Overall, the small and large grab collected a comparable number of species and individuals: 248 and 233 species, and 6074 and 6143 individuals, respectively. The large grab retrieved the most species at the deepest location while the small grab collected more species at the other test sites. Based on internationally recommended 0.5 m^2 sampling units, a variation in the total species richness per test site of 7–13% was found while the diversity indices (ES_{100} and H′) varied by less than 10%. Independent of grab size, a cluster and nMDS analysis showed four clearly separated sample-groups that correspond to the four test stations although the multivariate dispersion was consistently higher for the small grab. A SIMPROF test showed no grab size-dependent differences. ANOSIM and PERMANOVA tests showed differences between grab sizes for the deepest station, where a similar difference occurred also between samples of the same grab size. This station displayed the least faunal heterogeneity, indicating that faunal patchiness may influence any grab-size differences. The results indicate that the two van Veen grabs tested deliver comparable quantitative faunal compositions. For the small grab, however, numerous samples were rejected due to poor performance, resulting in increased sampling time, ship costs and a suggested biased sampling towards less heterogeneous sediments at the fjord site.

1. Introduction

Quantitative sediment community composition is often used to document and monitor anthropogenic-derived environmental change in polluted marine areas (reviewed in Gray and Elliot, 2009). Gear size, diversity indices, community structure, indicator species and various multivariate analyses are among the most recommended assessment techniques being used as a background for remedial action that may contribute towards reversing anthropogenic impacts (Gray and Elliot, 2009; ISO, 2013; OSPAR, 2012; Norwegian Environment Agency, 2015; HELCOM, 2017; Noble-James et al., 2018). To secure sampling success when targeting great depths (>500 m) and coarse sediments, heavy and large sampling gears are frequently used instead of smaller gears that may cause uneven hits of the seafloor, inadequate closure of the gear used, and increased field time and sampling costs. The recommended methods used in offshore environmental monitoring are largely focused upon sampling gears that most often work successfully at shelf depths, such as 0.1 m^2 grabs or boxcorers, and 3–5 replicates per locality (station) (e.g. ISO, 2013; Norwegian Environment Agency, 2015; Eggleton et al., 2017; HELCOM, 2017; Przeslawski et al., 2018). In practise, therefore, selecting which sampler size to be used at great depths may become a trade-off between the number of replicates and sample size related field costs.

In contrast to the offshore pollution monitoring conducted by the petroleum industry, the MAREANO baseline mapping program aims to document seafloor nature types in Norwegian seas in general, inclusive of fauna in various sediments (sand, mud, mixed) and depths (40–6000 m) (Buhl-Mortensen et al., 2015). For comparative reasons, a relatively

* Corresponding author.
E-mail address: boerge.holte@imr.no (B. Holte).
[1] The authors have equally contributed to the MS

https://doi.org/10.1016/j.marenvres.2019.104867
Received 22 August 2019; Received in revised form 16 December 2019; Accepted 22 December 2019
Available online 26 December 2019

Table 1
Station- and sample data for small (0.10 m^2) and large (0.25 m^2) van Veen grabs, including sampling effort (no. of accepted and rejected replicate samples), sampled replicate volume, carbon and nitrogen in sediment, grain size, bottom temperature and salinity. Maximum grab volumes are 18 and 80 L, respectively. TOC: Total organic carbon (%, by weight); TN: Total nitrogen (%, by weight).

Station	Region	Depth (m)	Sampling date	Position		Grab size	Accepted samples	Rejected samples	Sample volume l.	TOC %	TN %	Clay + silt %	Sand %	Temp. °C	Salinity
				N	E										
R0	Tanafjord	303	20.04.2014	70°	28°	Large	4	3	50–55	1.1	0.09	94	6	5.2	35.0
R0				49.91	31.05	Small	10	11	12–18						
R1403	Central	287	24.08.2014	72°	33°	Large	4	0	60	1.5	1.9	84	15	3.4	35.0
R1403	Barents Sea			23.38	13.55	Small	10	5	10–18						
R1349	Storegga	767	16.06.2014	63°	05°	Large	4	0	75–80	1.5	2.1	92	8	−0.6	34.9
R1349	(1)			35.64	34.40	Small	10	0	18						
R1350	Storegga	963	20.06.2014	63°	05°	Large	4	0	80	1.8	2.3	99	1	−0.8	34.9
R1350	(2)			37.93	30.04	Small	10	3	18						

heavy 0.25 m^2 grab sampler has been used by MAREANO, not only deeper than 500 m but also in shallower waters. To test whether petroleum-related offshore monitoring, using 0.1 m^2 grabs, may take advantage of MAREANO-stations as future base-line control stations, the MAREANO program in 2013 launched the present study in order to document the quantitative comparability of the samples collected using small vs. large grabs. The major questions were: *Is there any difference in macrofaunal sampling metrics between 0.25 m^2 and 0.1 m^2 van Veen grabs, and how do any such differences influence the fauna being collected using the ISO recommended sampling area of 0.5 m^2 per station?*

The purpose of the present study is to answer this question by comparing the amount and composition of macrofauna sampled by using small (0.1 m^2) and large (0.25 m^2) van Veen grabs, each capturing a bottom area of 1 m^2 at the same localities and times.

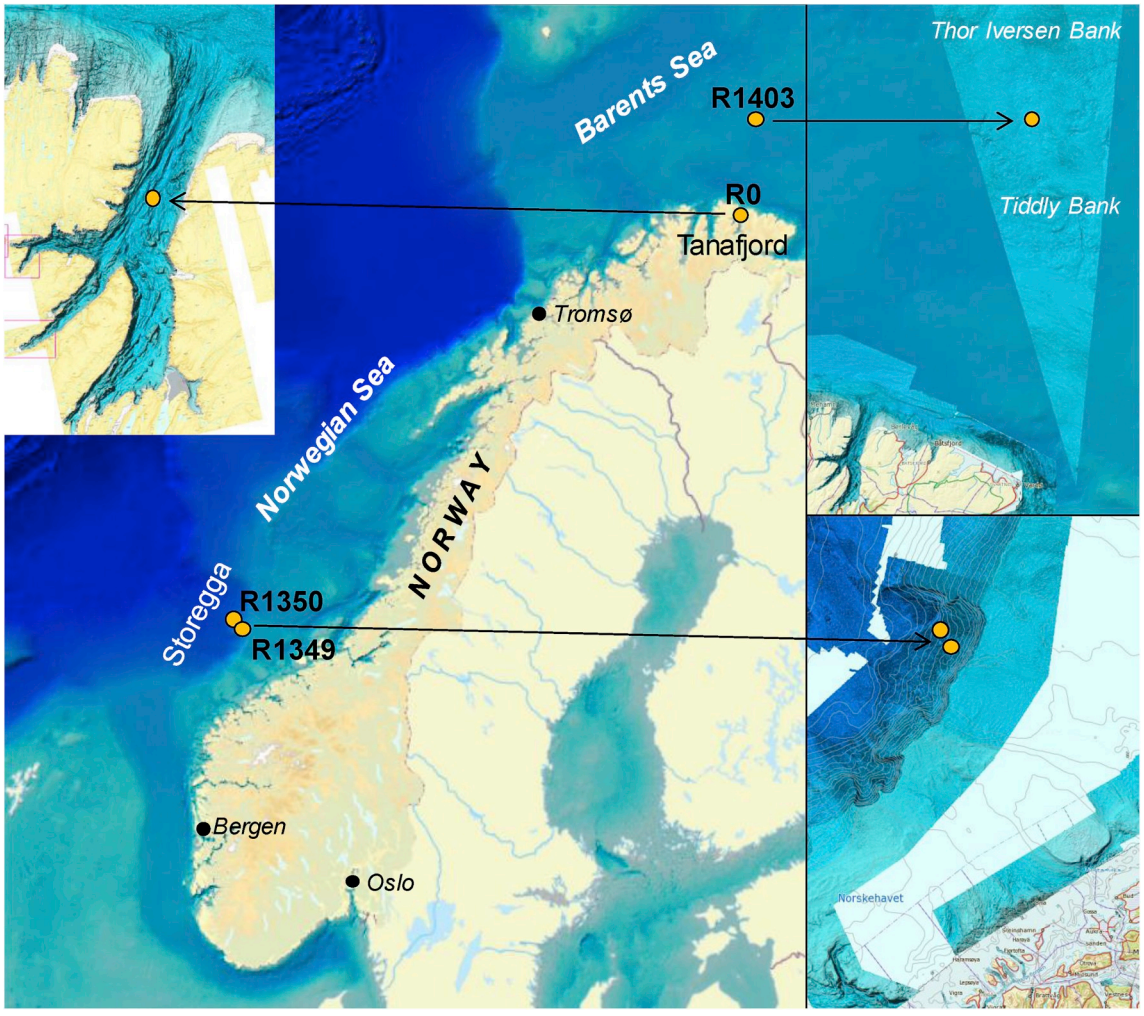

Fig. 1. Grab test sites in the Tanafjord (R0), the central part of the Barents Sea (R1403), and on the Storegga continental slope in the Norwegian Sea (R1349, R1350). Geographical co-ordinates are shown in Table 1. Map source: www.mareano.no.

Table 2
Abundance and diversity for the four test stations as collected by small (0.1 m^2) and large grabs (0.25 m^2). The figures show sampled area per station (2.0 m^2), per grab size (1.0 m^2) and for two 0.5 m^2 area units per grab size. N: No. of specimens; S: No. of taxa; ES100: Hurlbert's diversity index for 100 sampled specimens; H': Shannon-Wiener diversity index.

Station	Grab-size	N		S			ES100		H′	
		1 m^2	2 × 0.5 m^2	2 m^2	1 m^2	2 × 0.5 m^2	1 m^2	2 × 0.5 m^2	1 m^2	2 × 0.5 m^2
R0	Small	1857	920; 937	170	141	100; 107	35	33; 37	4.9	4.6; 5.1
R0	Large	1727	794; 933		113	73; 96	32	28; 34	4.6	4.1; 4.9
R1403	Small	1019	470; 549	141	109	78; 84	39	35; 41	5.0	4.6; 5.0
R1403	Large	679	314; 365		101	68; 77	43	41; 44	5.4	5.1; 5.3
R1349	Small	1037	490; 547	123	101	74; 77	34	31; 35	4.6	4.1; 4.4
R1349	Large	946	307; 639		91	63; 67	31	28; 35	4.5	4.3; 4.5
R1350	Small	2161	1057; 1104	90	62	37; 46	15	13; 16	2.3	2.1; 2.4
R1350	Large	2791	1323; 1468		74	48; 62	15	13; 17	2.3	1.9; 2.6

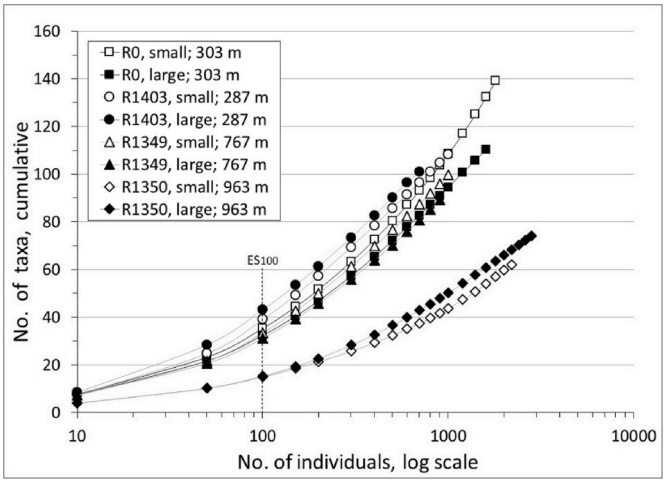

Fig. 2. Rarefaction curves based on Hurlbert's index, showing the cumulative number of taxa vs. the number of collected individuals. The vertical dotted line shows the calculated number of taxa for 100 sampled individuals (ES100; see Table 2).

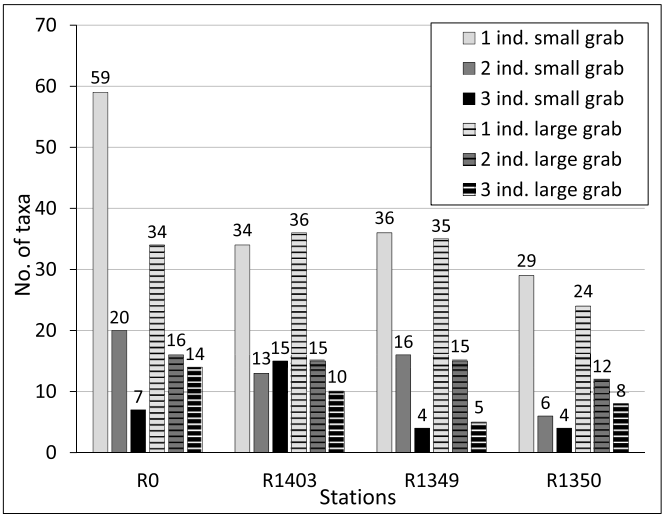

Fig. 3. Number of rare species for small and large grabs occurring as 1, 2 or 3 individuals (for 1.0 m^2 sampled area per grab size and station).

2. Material and methods

2.1. Stations

Four test stations with varying depth and environment were sampled during the 2014 MAREANO field campaign (Table 1). Two stations were located in the Barents Sea, one at 287 m depth in the ice-free central part of the sea (R1403, Fig. 1) and one at 303 m depth in the 65 km long and up to 10 km wide Tanafjord (R0). The other two stations (R1349; R1350) were situated at 767 and 963 m depth in Arctic-originating waters off southwest Norway (Mork and Skagseth, 2010).

2.2. Grabs and sampling

Two van Veen grabs with sampling areas of 0.10 m^2 and 0.25 m^2 were tested comparatively, each grab size sampled a total area of 1.0 m^2 per station, equaling 10 small and 4 large replicate samples covering a total area of 2.0 m^2. The weight of each grab was 50 and 125 kg, with a maximum sampling volume of 18 and 80 L, and a maximum bite depth of 18 and 34 cm, respectively. Hinged inspection windows covered 75% of the grabs' top surface and comprised of 0.5 mm steel mesh and flexible rubber flaps to minimize the bow-wave effect (Rumohr, 2009; ISO, 2013). Apart from their size, both grabs were identically constructed and produced by KC Denmark Ltd, Copenhagen.

To minimize resuspension of surface sediments (the bow-wave effect), the grabs' lowering speed was reduced to 0.2 ms^{-1} just before settling on the seafloor (ISO, 2013). Replicate samples penetrating less than 7 cm into the sediments (in bite depth) or with sediment loss or leakage when retrieved, were routinely rejected and substituted with succeeding approved replicates.

2.3. Sediment parameters

An extra grab sample taken by the large grab was collected for the subsampling of sediments to be analyzed for total organic carbon (TOC), total nitrogen (TN), clay + silt (<0.063 mm) and sand (>0.063; <2.0 mm). Two subsamples were taken through the inspection windows, using Plexiglass corers with an inner diameter of 6 cm and then stored at −24 °C. The upper five cm of the sediment column were analyzed after mixing the two individual samples taken per station. TOC was analyzed after baking in an induction oven at 1300 °C (European Standard, 2012). After being dissolved in sulfuric acid, potassium sulphate and Dayarda's alloy, TN was quantified by titration of ammonium ions using hydrochloric acid. Grain size was measured using mechanical (>0.063 mm) and wet-sieving (<0.063 mm).

2.4. Processing of fauna samples

The fauna samples were sieved through a 1 mm mesh sieve, using a Wilson Autosiever, and stored in Borax-buffered 4–6% formaldehyde

Table 3
The ten most abundant taxa per station and grab size, showing the abundance and percent of the total number of specimens collected per 1 m^2. Act: Actinaria; Amp: Amphipoda; Biv: Bivalvia; Hol: Holothuroidea; Nem: Nemertea; Oph: Ophiuroidea; Pol: Polychaeta; Sip: Sipunculida.

Small grab (10 × 0.10 m^2)				Large grab (4 × 0.25 m^2)			
Taxa		N m^{-2}	%	Taxa		N m^{-2}	%
Tanafjord R0							
Maldane sarsi	Pol	374	20	Maldane sarsi	Pol	427	25
Chirimia biceps	Pol	240	13	*Chirimia biceps*	Pol	197	11
Rhodine gracilior	Pol	189	10	*Rhodine gracilior*	Pol	157	9
Labidoplax buskii	Hol	89	5	*Labidoplax buskii*	Hol	96	6
Ophiuroidea indet. *juv.*	Oph	72	4	*Ophiura sarsii*	Oph	71	4
Terebellides sp.	Pol	55	3	Athenaria indet.	Act	70	4
Yoldiella nana	Biv	55	3	Ophiuroidea indet. *juv.*	Oph	56	3
Parvicardium minimum	Biv	51	3	*Parvicardium minimum*	Biv	53	3
Ceratocephale loveni	Pol	46	2	*Phascolion strombus*	Sip	43	2
Abyssoninoe scopa	Pol	44	2	*Terebellides* sp.	Pol	33	2
				Ceratocephale loveni	Pol	33	2
Barents Sea R1403							
Spiochaetopterus typicus	Pol	200	20	*Galathowenia oculata*	Pol	102	15
Galathowenia oculata	Pol	202	20	*Spiochaetopterus typicus*	Pol	62	9
Terebellides sp.	Pol	42	4	*Portlandia intermedia*	Biv	32	5
Yoldiella nana	Biv	30	3	*Yoldiella nana*	Biv	30	4
Spiochaetopterus sp.	Pol	29	3	*Terebellides* sp.	Pol	27	4
Portlandia intermedia	Biv	27	3	Athenaria indet.	Act	20	3
Parathyasira equalis	Biv	23	2	*Parathyasira equalis*	Biv	18	3
Galathowenia fragilis	Pol	21	2	*Nephtys ciliata*	Pol	18	3
Spiophanes kröyeri	Pol	20	2	*Paramphinome jeffreysii*	Pol	18	3
Cuspidaria subtorta	Biv	18	2	*Cuspidaria subtorta*	Biv	16	2
Storegga R1349							
Paramphinome jeffreysii	Pol	249	24	*Paramphinome jeffreysii*	Pol	180	19
Mendicula ferruginosa	Biv	147	14	*Parathyasira equalis*	Biv	124	13
Parathyasira equalis	Biv	86	8	*Mendicula ferruginosa*	Biv	114	12
Sipuncula indet.	Sip	62	6	*Sipuncula* indet.	Sip	82	9
Yoldiella nana	Biv	56	5	*Yoldiella nana*	Biv	64	7
Aricidea hartmani	Pol	40	4	*Maldane arctica*	Pol	47	5
Maldane arctica	Pol	40	4	*Harpiniopsis similis*	Amp	25	3
Amage auricula	Pol	16	2	*Thyasira obsoleta*	Biv	22	2
Abyssoninoe scopa	Pol	15	1	*Euclymeninae* indet.	Pol	13	1
Glyphanostomum pallescens	Pol	12	1	*Jasmineira schaudinni*	Pol	13	1
Storegga R1350							
Parathyasira equalis	Biv	1378	64	*Parathyasira equalis*	Biv	1810	65
Mendicula ferruginosa	Biv	247	11	*Paramphinome jeffreysii*	Pol	245	9
Paramphinome jeffreysii	Pol	120	6	*Mendicula ferruginosa*	Biv	216	8
Yoldiella nana	Biv	79	4	*Yoldiella nana*	Biv	95	3
Harpiniopsis similis	Amp	39	2	*Harpiniopsis similis*	Amp	65	2
Aphelochaeta sp.	Pol	33	2	*Laonice cirrata*	Pol	34	1
Nemertea indet.	Nem	26	1	*Chaetozone* sp.	Pol	28	1
Aricidea quadrilobata	Pol	26	1	Sipuncula indet.	Sip	25	1
Laonice cirrata	Pol	25	1	*Aphelochaeta* sp.	Pol	24	1
Levinsenia gracilis	Pol	20	1	*Scoloplos* sp.	Pol	18	1

solution. After the cruises, all organisms were sorted from the remaining sediments in accordance with international recommendations (Rumohr, 2009; ISO, 2013). After sorting, the fauna was transferred to 70% ethanol and identified to the lowest possible taxonomical level.

2.5. Statistical analyses

All statistical calculations, analyses and tests were performed using the PRIMER7 software package (Clarke et al., 2014). In accordance with the petroleum-related offshore monitoring guidelines (ISO, 2013; Norwegian Environment Agency, 2015), Nematoda, Porifera, colonial Cnidaria and Bryozoa were excluded from this study. Diversity was calculated using the Shannon-Wiener H' (log_2) and Hurlbert's ES100 indices while rarefaction/diversity and species accumulation curves were made using Hurlbert's index and the re-sampling statistics described by Ugland et al. (2003), respectively.

Cluster- and nMDS analyses were based on square root transformed data and Bray-Curtis similarity index. In addition to analyses of 1.0 m^2 datasets and their replicates, analyses based on the ISO-recommended sampling area of 0.5 m^2 per station were also performed (ISO, 2013). Each replicate in these area units was marked according to their order of field sampling. In the 1.0 m^2-based analyses, the comparability between small and large grab's replicates was optimized by percent standardizing the number of specimens.

In addition to calculating multivariate dispersion distances (PERMDISP; Anderson, 2006), the difference in variance between the grab sizes was assessed using the multivariate permutation tests PERMANOVA (Anderson, 2001; Anderson et al., 2008), ANOSIM (Clarke, 1993) and SIMPROF (Clarke, 1993; Clarke et al., 2016). The input data for these tests was taken from the respective Bray-Curtis matrices. Apart from SIMPROF test, all tests between the 0.5 m^2 sample units of the same grab size were only run for the small grab due to the small number of replicate samples available for the large grab (n = 2 per 0.5 m^2 unit).

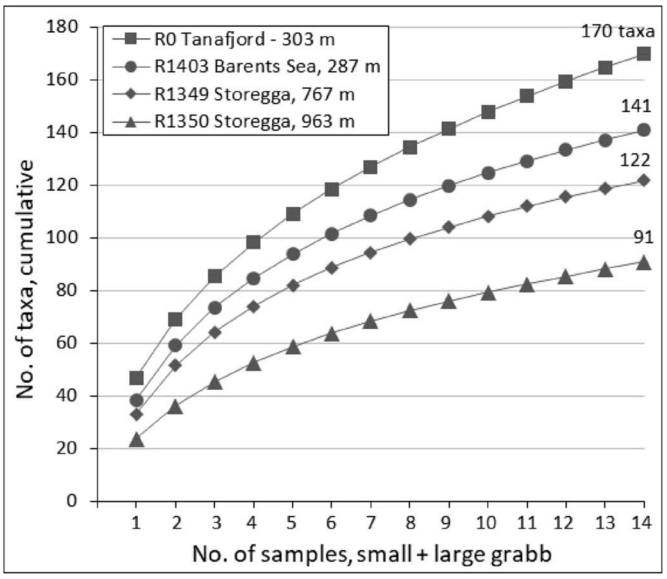

Fig. 4. Species accumulation curves per station based on ten small (0.1 m^2) and four large grab replicates (0.25 m^2). The curves' upper end-points show the total number of taxa collected per station, covering an area of 2.0 m^2.

3. Results

3.1. Sediments and sampling

The sediments at all stations were quite similar, light gray and apparently oxygenated, consisting of 84–99% silt + clay and 1–15% sand with a TOC content of 1.1–1.8% (Table 1). At the Tanafjord station (R0), a gravel component was observed visually in the samples taken by both grab sizes (photo documented on board). Due to gravelly stones that prevented grab closure, 19 small grab replicates and three large grab replicates were rejected, of which 14 were at the Tanafjord station (Table 1). The overall minimum grab bite (sediment depth) was 16 cm for the small grab and 20.5 cm for the large grab.

3.2. Species richness and diversity

304 taxa and 12,217 specimens were recorded from the 56 approved grab samples taken. The small grab collected 248 taxa and 6074 specimens while the large grab collected 233 taxa and 6143 specimens. The highest number of taxa (170) was found at the Tanafjord station (R0) while the lowest number (90) was recorded at the deepest station (R1350; Table 1; Table 2). The large grab collected a higher number of species than the small grab at the deepest station only (Fig. 2; Table 2). The highest difference in number of taxa between the grab sizes was recorded at the fjord station where the small grab collected 28 more taxa than the large grab, equivalent to 16% of the total number at this station (Table 2). Of these, 25 taxa comprised of one single individual (Fig. 3). At the other three stations, the number of singletons between the grab sizes did not exceed five, and the difference between the grab sizes is 8–12 species (Table 2). For the 0.5 m^2 units, the between grab-size diversity indices H′ and ES100 (Table 2) on average varies within 9.8% (SD ± 7.2; n = 22) of their respective maximum values, while the among samples values for the small grab are 11.8% (SD ± 3.6; n = 8) and for the large grab are 15.0% (±8.4; n = 8). For both grab sizes, the diversity indices are clearly lowest at the deepest station (Table 2; Fig. 2), mainly caused by the numerically highly dominant sediment feeding bivalve *Parathyasira equalis* (Table 3). Among the 1.0 m^2 units, the average between-grab diversity metric difference is 5.3% (n = 8; no difference for R1350).

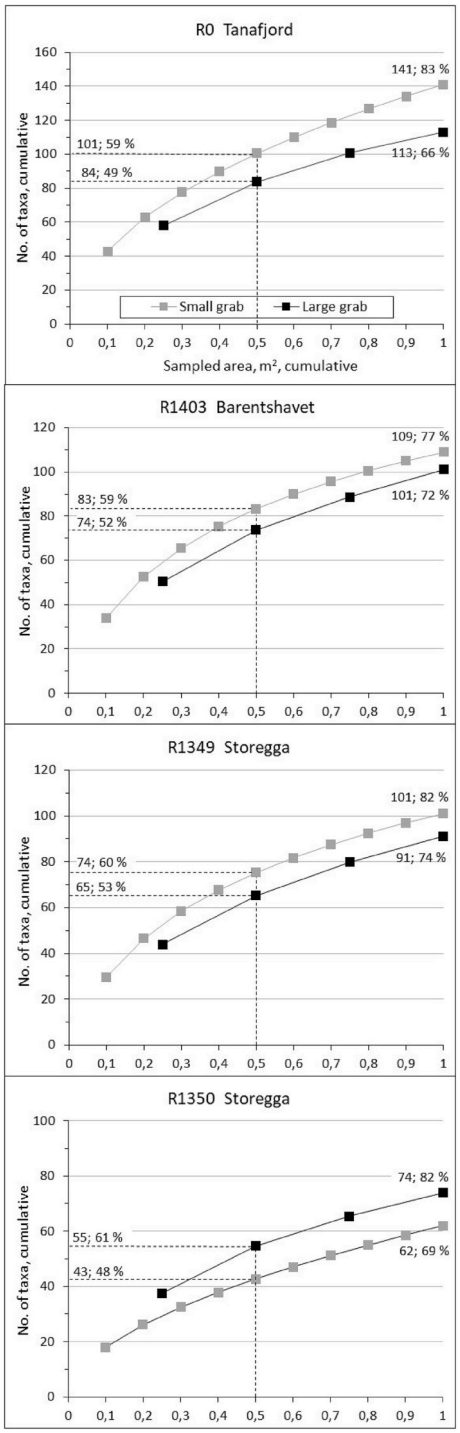

Fig. 5. Species accumulation curves for ten small (0.1 m^2; gray squares) and four large grab samples (0.25 m^2; black squares). The number of collected species and the corresponding proportion of the total number per station (%; see Fig. 4) are shown for 0.5 m^2 and 1.0 m^2.

3.3. Sampled vs. true species richness

Even when all samples from both grab sizes were included in the species accumulation curves, none of the curves reached a horizontal asymptote (Fig. 4; Fig. 5). Thus, the difference between the number of true and collected species is relatively high. For the 0.5 m^2 sampling area both small and large grabs collected quite similar proportions of the total recorded number of taxa per station, i.e. 48–60% vs. 49–61%

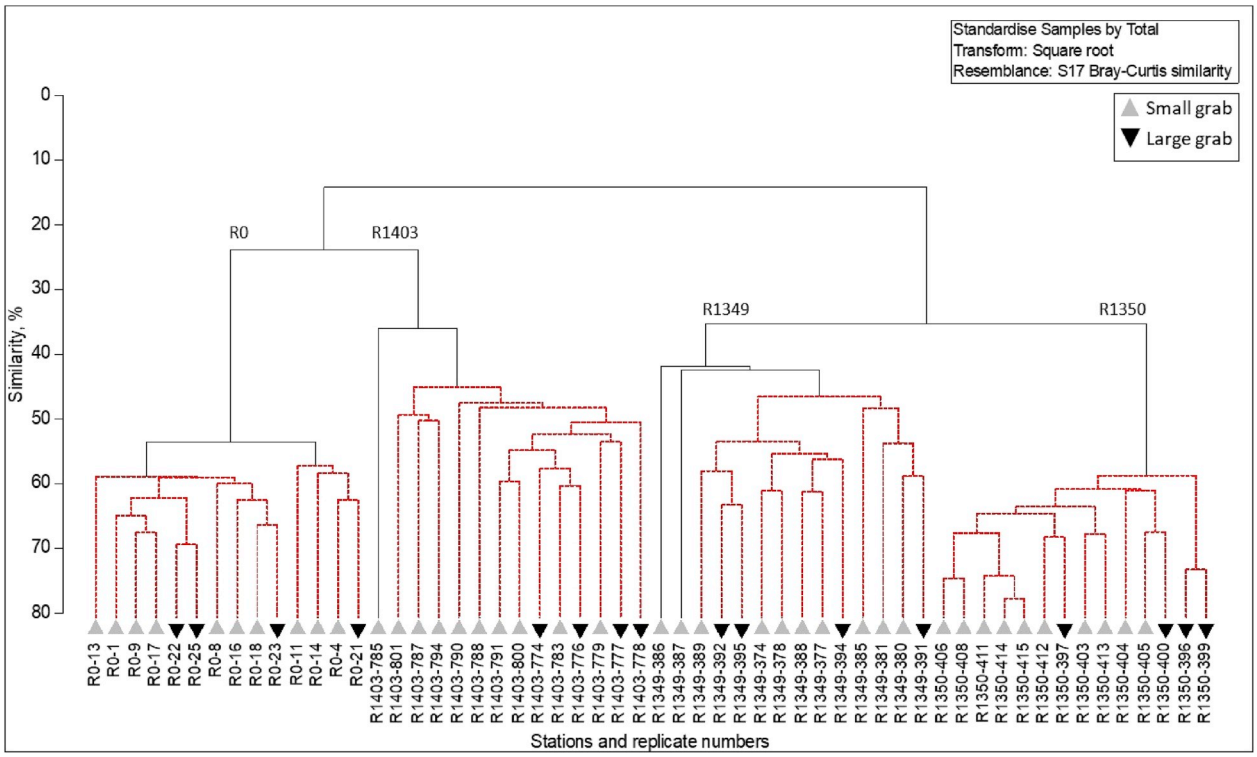

Fig. 6. Cluster analysis for small and large grab replicates collected at four test stations (R0, R1403, R1349, R1350). Solid lines show divisions that are significantly different from other groups (SIMPROF test; 5% significance level). Correlation to the input data: 0.96.

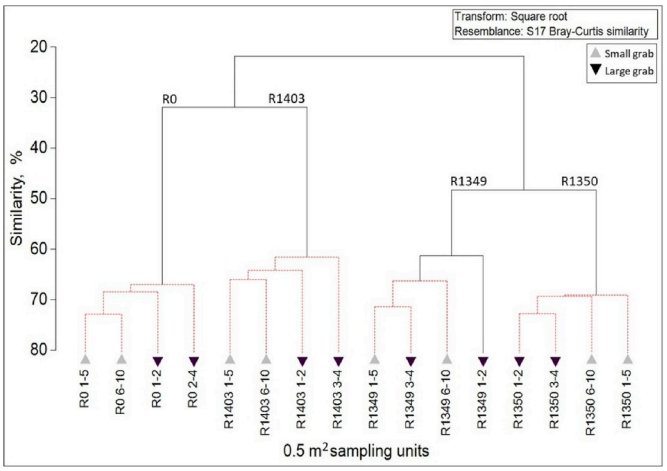

Fig. 7. Cluster analysis of 0.5 m^2 sampling units at four test stations (R0, R1430, R1349, R1350). Each unit comprise five and two replicate samples from small and large grabs, respectively, numbered in accordance with their successive field sampling order. Solid lines show divisions that are significantly different from the other groups of 0.5 m^2 units (SIMPROF test; 5% significance level). Correlation to the input data: 0.97.

respectively. The species accumulation curves for the 0.5 m^2 units (Fig. 5) show a grab-size related difference of 7–13% in the number of taxa collected as a proportion of the total collected per station (Fig. 4), with the largest difference at the deepest station (R1350). For the 1.0 m^2 sampling area, this grab-size difference was largely maintained for the offshore stations but increased from 10% to 17% for the fjord station.

3.4. Community composition; multivariate analyses

The most abundant species (the "top-ten" taxa; Table 3) collected by the two grab sizes, with few exceptions, were similar at the respective stations; 6–8 identical top-ten species occurred in both grabs. Independent of grab size, the cluster analyses formed significantly separated groups of replicates that correspond to the four test stations (Fig. 6; Fig. 7). The replicate nMDS analysis visualizes that the small grab generally has larger faunal variation compared to the large grab (Fig. 8A), which is also indicated by the dispersion distances that are highest (i.e. more heterogeneous) for the small grab at all stations while significant, or close to significant, grab-size related differences occurred at two stations (R0; R1403) (Table 4). Like the replicate-based analyses, all stations were in the directly comparable 0.5 m^2 unit analyses significantly separated from each other independent of grab size. It was noted that one large 0.5 m^2 unit at station R1349 later in the cluster analysis was significantly separated from the other units at this station (Fig. 7; Fig. 8B; Table 4).

In contrast to the SIMPROF test, the ANOSIM and PERMANOVA tests reveal significant grab-size differences for the deepest station (R1350), while also showing significant difference between separate small grab-size units (0.5 m^2) (Table 5).

4. Discussion

The aim of the paper was to document/test the advantages and disadvantages of using large or small grabs when sampling benthic macrofauna at shelf or bathyal depths. This was tested at four locations with varying depth and environment and showed that broadly the results were comparable between grab sizes, but certain aspects could recommend use of one size over another.

4.1. Sampling performance

One reason to choose a larger sampler rather than a small one in deeper water (>200 m) is the better sampling efficiency gained from using heavier gear. Large samplers will generally perform better at retrieving samples from mixed sediments and greater depths due to a

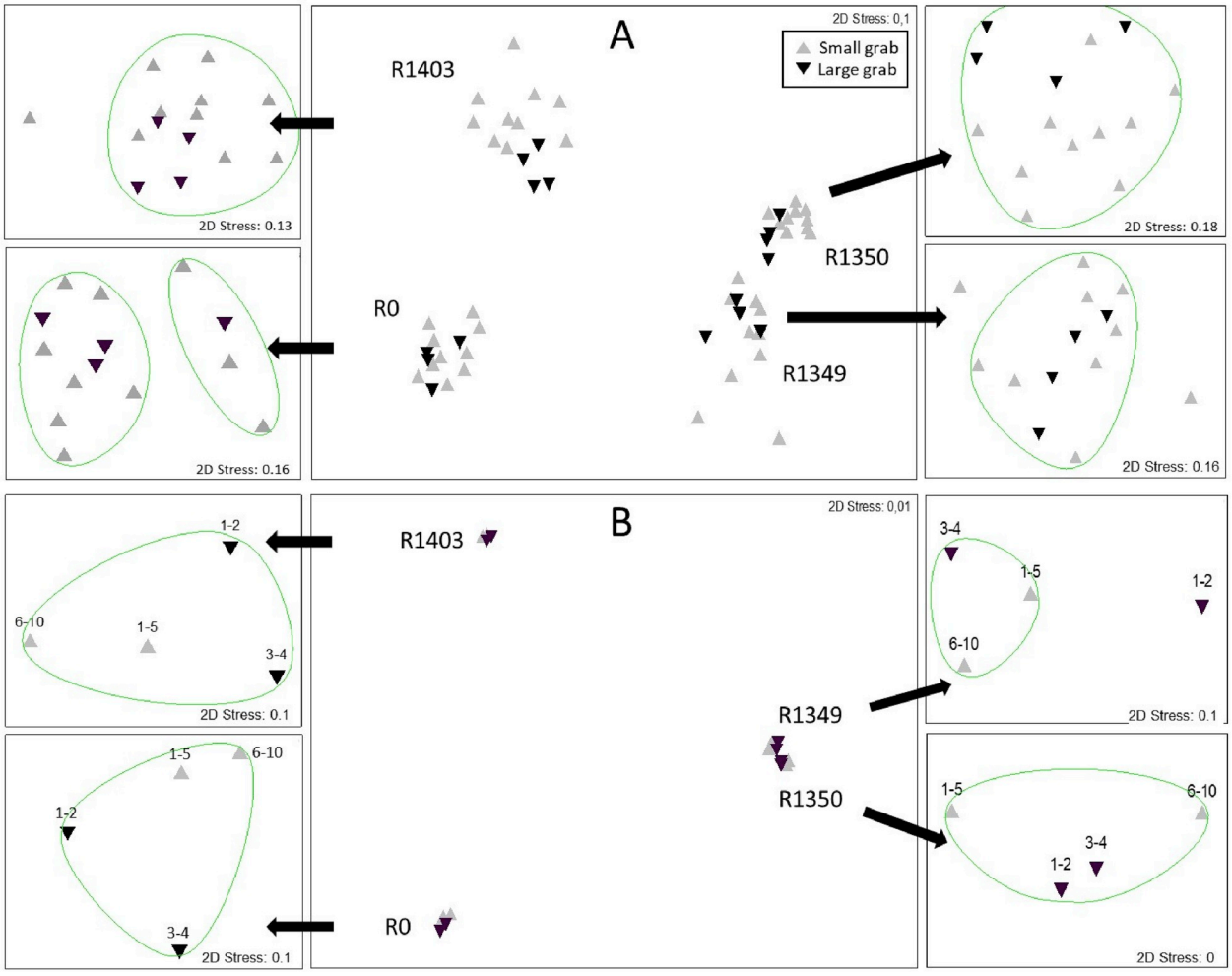

Fig. 8. nMDS-analyses for small and large grab replicates (**A**; see Fig. 6) and 0.5 m^2 sampling units (**B**; see Fig. 7) from four test stations (R0, R1403, R1349, R1350). Each 0.5 m^2 unit comprise five and two replicate samples from small and large grabs, respectively, numbered in accordance with their successive field sampling order. Solid lines show divisions that are significantly different from the other groups of replicates/0.5 m^2 units (SIMPROF test included in the cluster analyses; 5% significance level; see cluster analyses).

Table 4
Results from dispersion tests (PERMDISP) including ANOVA F values. **A.** Tests between small (0.10 m^2; n = 10) and large grabs (0.25 m^2; n = 4); **B.** Control tests between small grab entities alone, based on two 0.5 m^2 area units per station (n = 5; replicate no. 1–5 vs. 6–10). Significantly different dispersions ($p < 0.05$ or just above) are shown in bold text.

A. DISPERSION – Small grab vs. large grab – 1.0 m^2 per grab size				
	R0	R1403	R1349	R1350
ANOVA F:	**11.1**	**8,5**	6.8	0.4
Distance, small grab; SE:	**29.2; 0.8**	**36.6; 1.8**	35.7; 1,4	24.0; 1,6
Distance, large grab; SE:	**24.0; 1.5**	**27.7; 1.3**	28.9; 2,3	22.2; 1,8
p:	**0.06**	**0.02**	0.10	0.59
B. DISPERSION – Small grab vs. small grab – 0.5 m^2 area units				
ANOVA F:	0.0	3.4	0,8	0.4
Distance, small grab no. 1–5; SE:	28,0; 1.0	37.0; 2.8	32.6; 1.0	22.6; 1.7
Distance, small grab no. 6–10; SE:	28,1; 1.7	31.2; 1.3	34.6; 2.1	20.5; 3.0
p:	0.97	0.13	0.42	0.63

more even and forceful hit on the seafloor. Therefore, when sampling in deep water fine-grained sediments with higher ship costs, the use of small sampling devices is often avoided even though they may be statistically advantageous compared to the relatively few replicates using larger samplers (see e.g. Boyd et al., 2006; Montagna et al., 2017). Even though the observed gravel at R0 (Tanafjord) clearly hindered grab closure for both launched grab sizes, the large grab – in contrast to the small grab – showed no rejected samples at the deep-sea stations in the central part of the Barents Sea and the deep slope of the Norwegian Sea.

4.2. Diversity and species richness

The between grab-size variation in abundance and diversity indices for the test stations was slightly lower than that within a single grab size (0.5 m^2 units), indicating that these descriptors of the macrofauna community were not affected by grab size. The overall occurrence of the deepest-burrowing fauna, the maldanid polychaetes *Maldane*, *Chirima* and *Rhodine*, and the capitellid polychaete *Heteromastus* (Clough and Lopez, 1993; Levin et al., 1997), were equally represented in both small and large grab sizes (825 vs. 804 specimens, respectively), thus indicating that the difference in minimum bite depth between the grabs, measured to 16.0 cm and 20.5 cm, respectively, did not have much effect on fauna collection (see e.g. Hines and Comtois, 1985; Dauwe et al., 1998; Rosenberg et al., 2000). Moreover, the most numerically dominant species, contributing with more than 5% to the total abundance per

Table 5
Results from the ANOSIM and PERMANOVA tests. **A.** Tests between small grabs (0.1 m^2; n = 10) and large grabs (0.25 m^2; n = 4) for 1 m^2 sampled area per grab size; **B.** Control test between small grab samples alone, based on two 0.5 m^2 area units per station (n = 5; replicate no. 1–5 vs. 6–10). Significantly different test results ($p < 0.05$ or just above) are shown in bold text.

A. Small grab vs. large grab; 1.0 m^2 sampled area per grab size				
	R0	R1403	R1349	R1350
ANOSIM				
Through R:	−0,9	−0,1	−0,2	**0,3**
Rmin/max (H_0):	−0,3–0,9	−0,4–0,6	−0,4–0,6	**−0,5–0,5**
p:	0,67	0,66	0,88	**0,03**
PERMANOVA				
pseudoF:	1,1	1,3	0,9	**2,0**
p:	0,22	0,12	0,56	**0,02**
Mean of squares:	1047; residual 911	1808; res. 1399	2328; res. 1502	**1333; res. 808**
B. Small grab vs. small grab; 0.5 m^2 area units				
ANOSIM				
Through R:	−0,1	0,1	0,0	**0,2**
Rmin/max (H_0):	−0,3–0,5	−0,2–0,3	−0,3–0,5	**−0,2–0,4**
p:	0,83	0,11	0,39	**0,06**
PERMANOVA				
pseudoF:	0,6	1,2	1,1	**1,8**
p:	0,94	0,18	0,32	**0,04**
Mean of squares:	627; residual 996	1783; res. 1488	1552; res. 1424	**1119; res. 608**

station, were identical for both grab sizes.

In contrast to the other test stations, there was a clear difference in species richness between the small and large grabs at R0 (Tanafjord), and the number of species occurring in one specimen was significantly higher in samples taken using the small grab. Differences in fauna composition captured by various sampling gears have been observed in other studies without finding consistent explanations (Somerfield and Clarke, 1997; Bett and Gauge, 2000; Shen et al., 2012; summarized by Bakke et al., 2016). We speculate that the presence of gravel at R0 could contribute to a relatively high density of micro-niches (Shen et al., 2012) forming faunal patches scaled to better be collected by the small grab. This was supported by Boyd et al. (2006) who found that relatively large faunal patches were a crucial explanatory factor behind the higher species richness found in small grabs compared to large grabs.

4.3. Dispersion and tests of variance

While the small grab tended to collect more species at three of the stations, the large grab sampled the highest number of species at the deepest station (R1350) which was located in subzero temperatures and had the overall lowest faunal dispersion and thus the lowest heterogeneity. It is suggested that the low and similar dispersion for both grabs at this station indicates relatively small-scaled faunal patchiness, which may provide some explanation as to why the large grab collected more species than the small grab (Boyd et al., 2006). For three of the stations, the between-size dispersion differences were not mirrored in the ANOSIM and PERMANOVA tests while, by contrast, a significant grab-size difference was indicated for the deepest station (Table 5A). However, the 0.5 m^2 unit-based tests undertaken within the same grab size (only available for the small grab; Table 5B) also showed a significant difference.

4.4. Sampled vs. true species richness

The species accumulation curves show that the aggregated species richness is clearly lower than the true asymptotic species richness for both grab sizes. Relative to the total number of species per station, a maximum proportion sampledof 59–61 % was recorded by the 0.5 m^2 units, which undoubtedly is an overestimation compared to the probable true (asymptotic) numbers. The maximum proportion sampled is largely in line with observations made by Rumohr et al. (2001) who found that the first five 0.1 m^2 grab replicates out of 70 taken in the Baltic Sea included 53% of the total number of taxa found. Thus, our results support the conclusion by Rumohr et al. (2001) that the number of replicates should not be based on the goal of sampling a larger proportion of the true number of species (see also Gray, 1984; Ellingsen et al., 2007). Furthermore, based on the large gap between the number of sampled species and the estimated true numbers, the differences in the sampled species richness proportion between the grab sizes (7–13%) themselves should not give reason to choose any particular grab size. Thus, the present results from pristine areas, finding the "correct" sampler size seems to be of less importance compared to using the same sampler size throughout a monitoring period.

4.5. Community similarity – cluster and nMDS analyses

Independent of grab size, the cluster and nMDS analyses showed four clearly separated groups of replicates that correspond to the four test stations. Each of the stations revealed similarities within their 14 replicates that coincide with their dispersion values, as, for example, seen for the deepest station (R1350) where the replicates are tightly clustered, closely followed by the Tanafjord station (R0). Paradoxically therefore, these two stations in this study represent "outliers" that showed notable, though non-significant, grab-size differences regarding the capture of rare species (Fig. 3), and aspects of faunal and environmental heterogeneity, which should be further tested. Additionally, the multivariate analyses indicated a consistent grab-size difference by the more scattered distribution of replicates for small grabs, as shown in the nMDS analysis (Fig. 8A). The 0.5 m^2 unit cluster analyses, confirmed the statistically homogenous composition of the respective stations, regardless of grab size.

It was noted that the two significantly separated sub-groups of replicates formed in the cluster analysis for the Tanafjord station, each comprising both grab sizes, showed some differences that may be of general interest related to grab size combined with sediment-induced rejection of replicates. Each of the replicates in the largest sub-group (ten replicates) showed lower diversity, abundance and species richness than any of the replicates of the same grab size in the other sub-group. It cannot, therefore, be excluded that the separation of these two sub-groups was influenced by the 14 gravel-induced rejections of replicates at the Tanafjord station, where repeated sampling may have tended to seek towards less gravelly and thus less heterogeneous sediments ("targeted sampling"). Although no grab-size difference was clear in the multivariate analyses, the large grab had considerably fewer gravel-induced rejections than the small grab (n = 3 vs. n = 11 rejected samples, respectively), thus large grab was less likely to seek towards these more homogeneous sediments.

The overall results from the present study give reason to conclude that the grab sizes tested deliver statistically comparable quantitative faunal results. However, consistent but non-significant differences in the species richness between the grab sizes were found and are likely to be connected to sediment and faunal heterogeneity and varying scales of patchiness.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was financed by the Norwegian Governmental MAREANO program (www.mareano.no) and the Norwegian Oil and Gas Association. We thank these, the crew on board RV G.O. Sars and RV Johan Hjort, our IMR-colleges on board and in the lab, and Akvaplan-niva AS and Fishguard AS for their identifications of polychaetes and amphipods, respectively. The authors also like to thank two anonymous reviewers that greatly contributed to improve the manuscript, and Dr. Rebecca Ross (IMR) for improving the English language.

References

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Aust. J. Ecol. 26, 32–46.
Anderson, M.J., 2006. Distance-based tests for homogeneity of multivariate dispersions. Biometrics 62, 245–253.
Anderson, M.J., Gorley, R.N., Clarke, K.R., 2008. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. PRIMER-E, Plymouth, UK.
Bakke, T., Andrade, H., Ellingsen, K.E., Norling, K., 2016. Vurdering Av MAREANOs Opplegg for Grabbprøvetaking Av Sedimentfauna - Harmonisering Med Prøvetaking Etter Norsk Standard (In Norwegian). NIVA-report No. 6524-2013.
Bett, B.J., Gauge, J.D., 2000. Practical approaches to monitoring the deep-sea environment of the UK Atlantic Margin. In: The Seafloor of the UK Atlantic Margin. Atlantic Frontier Environmental Network, ISBN 09538399-0-7.
Boyd, S.E., Barry, J., Nicholson, M., 2006. A comparative study of a 0.1 m^2 and 0.25 m^2 Hamon grab for sampling macrobenthic fauna from offhore marine gravels. J. Mar. Biol. U.K. 86, 1315–1328.
Buhl-Mortensen, L., Buhl-Mortensen, P., Dolan, M.F.J., Holte, B., 2015. The MAREANO programme – a full coverage mapping of the Norwegian off-shore benthic environment and fauna. Mar. Biol. Res. 11, 4–17.
Clarke, K.R., 1993. Non-parametric multivariate analyses of change in community structure. Aust. J. Ecol. 18, 117–143.
Clarke, K.R., Gorley, R.N., Somerfield, P.J., Warwick, R.M., 2014. Change in Marine Communities: an Approach to Statistical Analysis and Interpretation, 3nd edition. PRIMER-E, Plymouth.
Clarke, K.R., Somerfield, P.J., Gorley, R.N., 2016. Clustering in non-parametric multivariate analyses. J. Exp. Mar. Biol. Ecol. 483, 147–155.
Clough, L.M., Lopez, G.R., 1993. Potential carbon sources for the head-down deposit-feeding polychaete Heteromastus filiformis. J. Mar. Res. 51, 595–616.
Dauwe, B., Herman, P.M.J., Heip, C.H.R., 1998. Community structure and bioturbation potential of macrofauna at four North Sea stations with contrasting food supply. Mar. Ecol.: Prog. Ser. 173, 67–83.
Eggleton, J., Murray, J., McIlwaine, P., Mason, C., Noble-James, T., Hinchen, H., Nelson, M., McBreen, F., Ware, S., Whomersley, P., 2017. Dogger Bank SCI 2014 Monitoring R&D Survey Report. JNCC/Cefas Partnership Report, No. 11.
Ellingsen, K.E., Hewitt, J.E., Thrush, S.F., 2007. Rare species, habitat diversity and functional redundancy in marine benthos. J. Sea Res. 58, 291–301.
European Standard, 2012. Sludge, Treated Biowaste, Soil and Waste Determination of Total Organic Carbon (TOC) by Dry Combustion. EN 15936, Procedure A – Indirect Method.
Gray, J.S., 1984. Ökologie Mariner Sedimente. Springer Verlag, Heidelberg.
Gray, J.S., Elliot, M., 2009. Ecology of Marine Sediments. From Science to Management, second ed. Oxford University Press.
HELCOM, 2017. Manual for Marine Monitoring in the COMBINE Programme of HELCOM. Annex C-8 Soft Bottom Macrozoobenthos, pp. 277–286. HELCOM report.
Hines, A.H., Comtois, K.L., 1985. Vertical distribution of infauna in sediments of a subestuary of central Chesapeake Bay. Estuaries 8, 296–304.
ISO, 2013. Water Quality. Guidelines for Quantitative Sampling and Sample Processing of Marine Softbottom Macrofaunal. International Organization for Standardization. European Standard EN ISO 16665.
Levin, L., Blair, N., DeMaster, D., Plaia, G., Fornes, W., Martin, C., Thomas, C., 1997. Rapid subduction of organic matter by maldanid polychaetes on the North Carolina slope. J. Mar. Res. 55, 595–611.
Montagna, P.A., Baguley, J.G., Hsiang, C.-Y., Reuscher, M.G., 2017. Comparison of sampling methods for deep-sea infauna. Limnol Oceanogr. Methods 15, 166–183.
Mork, K.A., Skagseth, Ø., 2010. A quantitative description of the Norwegian Atlantic Current by combining altimetry and hydrography. Ocean Sci. 6, 901–911.
Noble-James, T., Jesus, A., McBreen, F., 2018. Monitoring Guidance for Marine Benthic Habitats (Revised 2018). JNCC Report No. 598. JNCC, Peterborough.
Norwegian Environment Agency, 2015. Environmental monitoring of petroleum activities on the Norwegian continental shelf. Guidelines No. M408-2015.
OSPAR, 2012. JAMP Eutrophication Monitoring Guidelines: Benthos. OSPAR Commission. Agreement 2012-12.
Przeslawski, R., Berents, P., Clark, M., Edgar, G., Frid, C., Hughes, L., Ingleton, T., Kennedy, D., Nichol, S., Smith, J., 2018. Marine sampling field manual for grabs and box corers. In: Przeslawski, R., Foster, S. (Eds.), Field Manuals for Marine Sampling to Monitor Australian Waters. National Environmental Science Programme (NESP), pp. 172–195.
Rosenberg, R., Nilsson, H.C., Hellman, B., Agrenius, S., 2000. Depth correlated benthic faunal quantity and infaunal burrow structures on the slopes of a marine depression. Estuar. Coast Shelf Sci. 50, 843–853.
Rumohr, H., 2009. Soft-bottom macrofaunal: collection, treatment, and quality assurance of samples. ICES Tech. Mar. Env. Sci. (43), 1–20.
Rumohr, H., Karakassis, I., Jensen, J.N., 2001. Estimating species richness, abundance and diversity with 70 macrobenthic replicates in the Western Baltic Sea. Mar. Ecol.: Prog. Ser. 214, 103–110.
Shen, P.-P., Zhou, H., Zhao, Z., Yu, X.-Z., Gu, J.-D., 2012. Evaluation of sampling sizes on the intertidal macroinfauna assessment in a subtropical mudflat of Hong Kong. Ecotoxicology 21, 1706–1716.
Somerfield, P.J., Clarke, K.R., 1997. A comparison of some methods commonly used for the collection of sublittoral sediments and their associated fauna. Mar. Environ. Res. 43, 145-l56.
Ugland, K.I., Gray, J.S., Ellingsen, K.E., 2003. The species-accumulation curve and estimation of species richness. J. Anim. Ecol. 72, 888–897.