

## 1.3 Example: Frierfjord macrofauna

The first example is from the IOC/GEEP practical workshop on biological effects of pollutants ( Bayne, Clarke & Gray (1988) ), held at the University of Oslo, August 1986. This attempted to contrast a range of biochemical, cellular, physiological and community analyses, applied to field samples from potentially contaminated and control sites, in a fjordic complex (Frierfjord/Langesundfjord) linked to Oslofjord ( {F} , Fig. 1.1). For the benthic macrofaunal component of this study ( Gray, Aschan, Carr *et al.* (1988) ), four replicate 0.1m<sup>2</sup> Day grab samples were taken at each of six sites (A-E and G, Fig 1.1) and, for each sample, organisms retained on a 1.0 mm sieve were identified and counted. Wet weights were determined for each species in each sample, by pooling individuals within species.

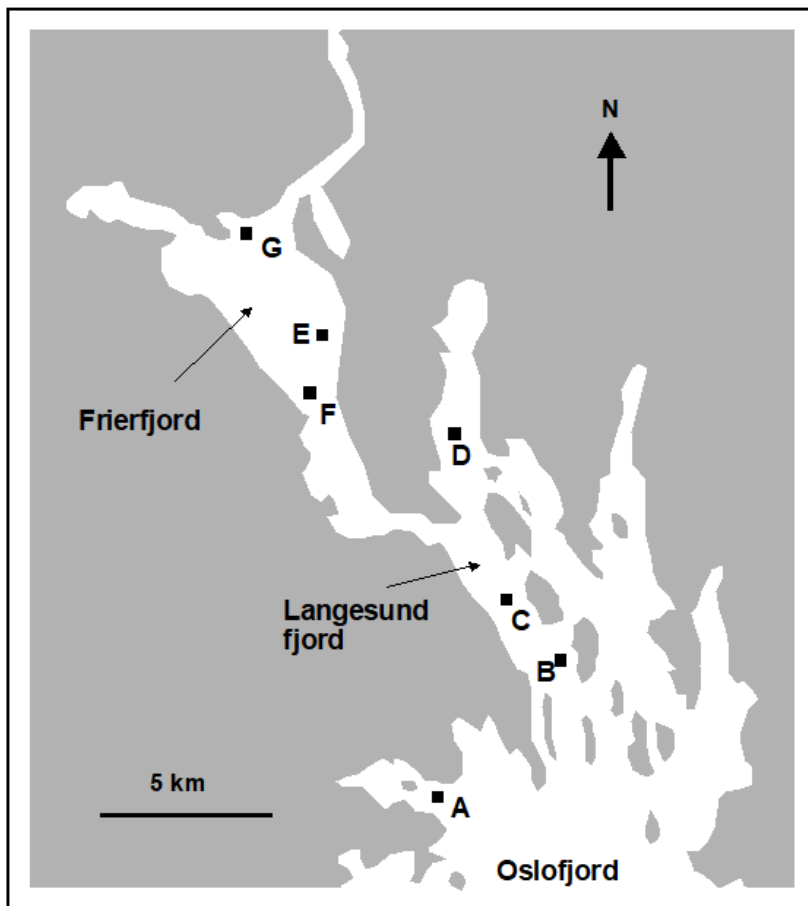


Fig. 1.1. Frierfjord, Norway {F}. Benthic community sampling sites (A-G) for the IOC/GEEP Oslo Workshop; site F omitted for macrobenthos.

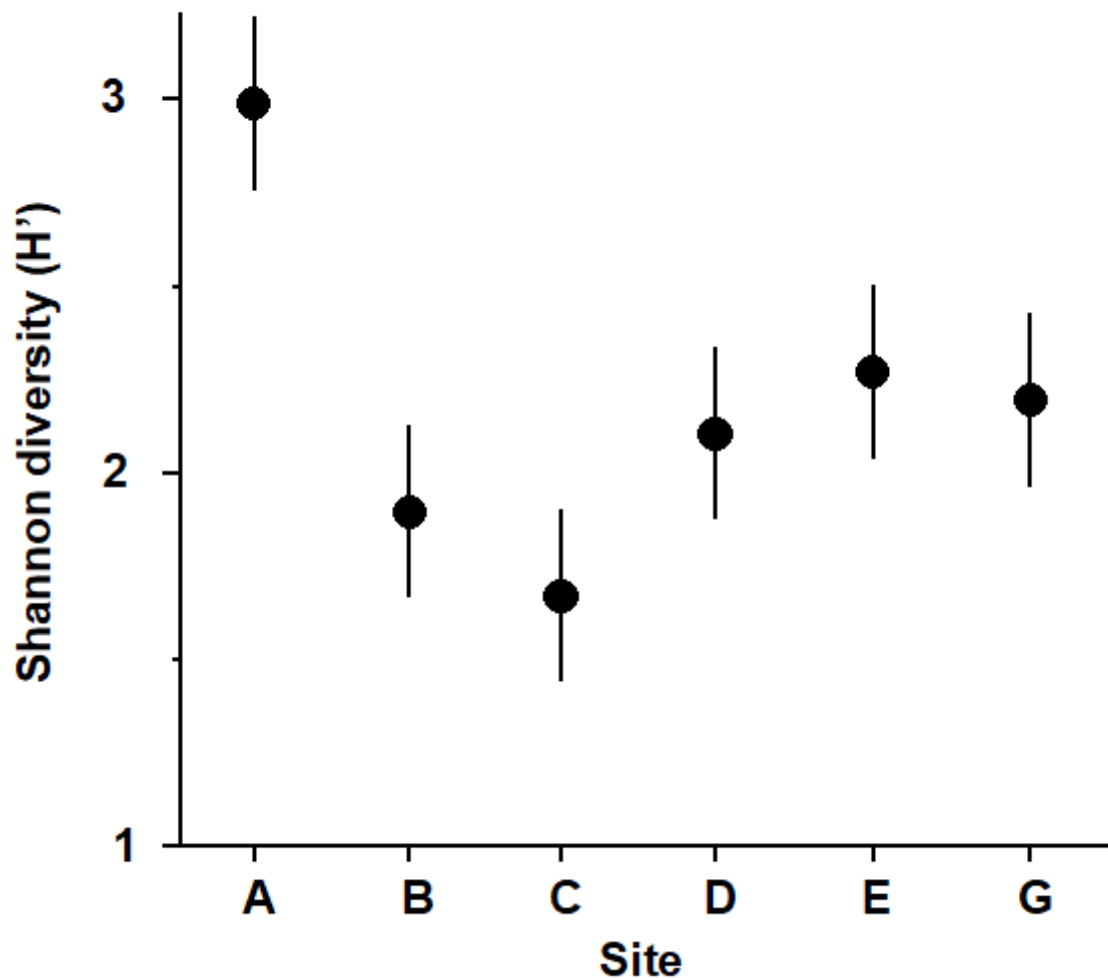


Fig. 1.2. Frierfjord macrofauna {F}. Means and 95% confidence intervals for Shannon diversity ( $H'$ ), from four replicates at each of six sites (A-E, G).

Part of the resulting data matrix can be seen in Table 1.2: in total there were 110 different taxa categorised from the 24 samples. Such matrices (abundance,  $A$ , and/or biomass,  $B$ ) are the starting point for the biotic analyses of this manual, and this example is typical in respect of the relatively high ratio of species to samples (always  $\gg 1$ ) and the prevalence of zeros. Here, as elsewhere, even an undesirable reduction to the 30 ‘most important’ species (see [Chapter 2](#)) leaves more than 50% of the matrix consisting of zeros. Standard multivariate normal analyses (e.g. [Mardia, Kent & Bibby \(1979\)](#)) of these counts are clearly ruled out; they require both that the number of species (variables) be small in relation to the number of samples, and that the abundance or biomass values are transformable to approximate normality: neither is possible.

Table 1.2. Frierfjord macrofauna {F}. Abundance and biomass matrices (part only) for the 110 species in 24 samples (four replicates at each of six sites A-E, G); abundance in numbers per  $0.1\text{m}^2$ , biomass in  $\text{mg}$  per  $0.1\text{m}^2$ .

Species	Samples
---------	---------

	A1	A2	A3	A4	B1	B2	B3	B4
<b>Abundance</b>								
<i>Cerianthus lloydi</i>	0	0	0	0	0	0	0	0
<i>Halicryptus</i> sp.	0	0	0	1	0	0	0	0
<i>Onchnesoma</i>	0	0	0	0	0	0	0	0
<i>Phascolion strombi</i>	0	0	0	1	0	0	1	0
<i>Golfingia</i> sp.	0	0	0	0	0	0	0	0
<i>Holothuroidea</i>	0	0	0	0	0	0	0	0
<i>Nemertina</i> , indet.	12	6	8	6	40	6	19	7
<i>Polychaeta</i> , indet.	5	0	0	0	0	0	1	0
<i>Amaena trilobata</i>	1	1	1	0	0	0	0	0
<i>Amphictetes gunneri</i>	0	0	0	0	4	0	0	0
<i>Ampharetidae</i>	0	0	0	0	1	0	0	0
<i>Anaitides groenl.</i>	0	0	0	1	1	0	0	0
<i>Anaitides</i> sp.	0	0	0	0	0	0	0	0
....								
<b>Biomass</b>								
<i>Cerianthus lloydi</i>	0	0	0	0	0	0	0	0
<i>Halicryptus</i> sp.	0	0	0	26	0	0	0	0
<i>Onchnesoma</i>	0	0	0	0	0	0	0	0
<i>Phascolion strombi</i>	0	0	0	6	0	0	2	0
<i>Golfingia</i> sp.	0	0	0	0	0	0	0	0

	A1	A2	A3	A4	B1	B2	B3	B4
<i>Holothuroi dea</i>	0	0	0	0	0	0	0	0
<i>Nemertina</i> , <i>indet.</i>	1	41	391	1	5	1	2	1
<i>Polychaeta</i> , <i>indet.</i>	9	0	0	0	0	0	0	0
<i>Amaena trilobata</i>	144	14	234	0	0	0	0	0
<i>Amphicteis gunneri</i>	0	0	0	0	45	0	0	0
<i>Amphareti dae</i>	0	0	0	0	0	0	0	0
<i>Anaitides groenl.</i>	0	0	0	7	11	0	0	0
<i>Anaitides sp.</i>	0	0	0	0	0	0	0	0
....								

As discussed above, one easy route to simplification of this *high-dimensional* (multi-species) complexity is to reduce each matrix column (sample) to a single univariate description. Fig. 1.2 shows the results of computing the Shannon diversity ( $H'$ , see [Chapter 8](#)) of each sample<sup>¶</sup>, and plotting for each site the mean diversity and its 95% confidence interval, based on a pooled estimate of variance across all sites from the ANOVA table, [Chapter 6](#). (An analysis of the type outlined in [Chapter 9](#) shows that prior transformation of  $H'$  is not required; it already has approximately constant variance across the sites, a necessary prerequisite for standard ANOVA). The most obvious feature of Fig. 1.2 is the relatively higher diversity at the *control/reference* location, A.

---

<sup>¶</sup> Using the *PRIMER DIVERSE* routine.

---