

An evaluation of the precision of diet description

E. Mümtaz Tirasin^{1,*}, Terje Jørgensen²

¹Department of Fisheries and Marine Biology, University of Bergen, High Technology Centre, N-5020 Bergen, Norway

²Institute of Marine Research, Fish Capture Division, PO Box 1870, N-5817 Bergen, Norway

ABSTRACT: Percentage compositions by weight and by number are the frequently used measures to evaluate the relative importance of different prey types in fish diet studies. The results obtained using these measures are only point estimates, and are often reported without any indication of their precision. Here an attempt is made to set confidence limits to these estimates via normal approximation and bootstrapping. Applications of these approaches are demonstrated with the stomach contents data of mackerel from the North Sea. The precision estimates of the average weight percentages were generally very poor, even when the prey items were grouped into major categories and sample sizes were comparatively large. Intra-haul correlation was found to be an important source of variation in the stomach content composition. A bootstrap method, which incorporated inter- and intra-haul variation, provided more realistic confidence intervals for weight percentages than the normal approximation. Implications of the uncertainty associated with these measures for dietary studies have been discussed. Routine estimation of the precision of dietary measures is recommended.

KEY WORDS: Feeding ecology · Dietary measures · Stomach contents · Bootstrap · Confidence intervals · Intra-haul correlation

INTRODUCTION

The relative importance of prey items in fish diets can be evaluated in a variety of ways (Windel & Bowen 1978, Hyslop 1980, MacDonald & Green 1983). The 3 most common measures used for quantitative description and evaluation are percentage composition by number (%N), percentage composition by volume (%V) or weight (%W), and percentage frequency of occurrence (%O). Percentage composition by number represents the proportion of the number of a particular prey item to the total number of all prey items in the entire stomach contents. Percentage composition by volume or weight is the proportion of the volume or biomass of a particular prey item to the total volume or weight of overall stomach contents. Percentage frequency of occurrence provides information on the proportion of fish stomachs containing a particular prey item irrespective of amount.

Each of these 3 measures provides a different insight into the feeding habits of fish under study. While percentage composition by number furnishes information about feeding behaviour, percentage composition by volume or weight reflects nutritional value of prey (Cailliet 1977, MacDonald & Green 1983, Cortéz 1997). Frequency of occurrence differs from the 2 former measures because it is, in fact, not a quantity of food but of fish qualified by their diet content. It does not describe the diet of an individual fish, but shows how uniformly the whole group of fish selects a particular prey item without actually indicating the importance of the selected prey item in respect to other prey. From this point of view, the percentage frequency of occurrence provides some information on population-wide food habits (Cailliet 1977), as well as some aid to assess the representation of the other 2 measures (Berg 1979).

Detailed discussions regarding the use of these measures as well as their advantages and disadvantages were presented by Windel & Bowen (1978), Berg (1979), Hyslop (1980), MacDonald & Green (1983), Eliassen & Jobling (1985), and Bowen (1996). However,

*E-mail: mumtaz.tirasin@ifm.uib.no

until recently little attention has been paid to the evaluation of the uncertainty associated with these measures (Ferry & Cailliet 1996, Jiang & Jørgensen 1996).

The main objective of this paper is to explore the appropriate means, particularly bootstrap resampling methods, for quantifying the precision of the estimates obtained from the former 2 measures of dietary description, i.e. the percentage diet composition by number or by weight. The second objective is to discuss the implications of the uncertainty associated with these measures for the dietary studies using them either directly or in the form of an index based upon their combination.

MATERIALS AND METHODS

The stomach contents data for this study were extracted from a larger data set of mackerel stomachs collected in the North Sea throughout 1991 in the framework of 'The stomach sampling project 1991' co-ordinated by the ICES (International Council for the Exploration of the Sea). Collection of samples and analysis of stomach contents are described in detail in Tirasin (unpubl.). In this paper, only the data from the 3rd quarter of 1991 (1 July to 5 September) were used. Stomachs which were collected in bulk could not be examined individually, and were therefore excluded along with those which were empty or consisted solely of unidentified food remains. Prey items have been arranged into 7 major groups as Mollusca, Copepoda, Amphipoda, Euphausiidae (Krill), Other Crustacea (i.e. all other crustaceans which are not included in the 3 preceding taxonomic groups or cannot be identified to any lower taxonomic level due to digestion), Fish, and Others (i.e. prey items which are not confined within any other category mentioned above). The numbers of hauls and stomachs sampled per haul for each successive 5 cm mackerel length class and sub-area in the North Sea are presented in Table 1. Division of the North Sea into 3 sub-areas was based upon its bathymetry: the northern North Sea with depths exceeding 100 m, the central North Sea with depths varying from 50 to 100 m, and the southern North Sea with depths of 50 m and shallower (Tirasin unpubl.).

Algebraically the average percentage proportion of a particular prey taxon i in the diet of a group of fish in terms of biomass is expressed as:

$$\% \hat{W}_i = \frac{\sum_{j=1}^{n_j} \bar{W}_{ij}}{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} W_{ij}} \cdot 100 \quad (1)$$

where W_{ij} is the weight of prey taxon i in the stomach contents of individual fish j , n_j is the total number of

prey taxa, and n_i is the total number of non-empty stomachs. If the relative numerical abundance of prey taxon i in the diet is desired then a similar expression can be constructed by substituting the weight for prey counts. The results obtained from the above equation are only point estimates. The standard error (SE) of this estimator is not readily given because the percentage or proportion data are generally not normally distributed (Cochran 1977). Somerton (1991) and Bowen (1996) suggested an angular transformation of the proportion data (Sokal & Rohlf 1995) so that they would conform more closely to the normal distribution. As noted by Somerton (1991), since $\% \hat{W}_i$ is estimated as a pooled proportion rather than the mean of the proportions for individual fish, the variance of this estimator cannot be calculated analytically, but can be approximated by the delta method (Seber 1973). A simpler analytical approach which does not require a transformation is provided by Cochran (1977) to estimate approximate variance of a ratio estimator for large samples. Following Cochran's approach an equation to calculate the approximate SE of the proportion of prey taxon i (\hat{W}_i) in the diet can be arranged as (ignoring the finite population correction):

$$SE(\hat{W}_i) = \frac{\sqrt{\frac{\sum_{j=1}^{n_j} (W_{ij} - \hat{W}_i \sum_{i=1}^{n_i} W_{ij})^2}{n_j(n_j - 1)}}}{\frac{1}{n_j} \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} W_{ij}} \quad (2)$$

where the variables are the same as in Eq. (1). Consequently, in order to get some measure of precision of the estimated average weight percentages, approximate confidence intervals can be set using the SE computed from Eq. (2) and appropriate critical values of the t -distribution.

Application of the above formula to compute the approximate SEs of prey proportions is based on the assumption that individual fish are random samples taken from the entire population. In general, however, fish from one particular haul would tend to have more similar stomach contents than fish from different hauls because prey distributions are often patchy (Hall et al. 1990, Bogstad et al. 1995). Due to this intra-haul correlation, a substantial amount of variation between hauls can be expected. However, Eq. (2) does not account for this variation. To explore its effect on the SE estimation, the individual hauls (i.e. all stomachs in each haul are pooled) are also used as the units in the calculations with Eq. (2).

Another approach, the bootstrap as initiated by Efron (1979), can alternatively be used to estimate SEs or to generate non-parametric CIs for the same estimators (Jiang & Jørgensen 1996). The bootstrap is a com-

Table 1 Frequency of hauls classified by the number of non-empty mackerel stomachs sampled per haul for each successive 5 cm length class (LC) and sub-area in the North Sea. Total number of hauls (H#) and stomachs (N#) per length class

Sub-area	LC	Number of stomachs per haul							H#	N#
		1	2	3	4	5	6	7+		
Northern North Sea	20	9	2	1	0	0	0	2	14	35
	25	8	6	4	5	0	3	15	41	277
	30	5	11	3	6	2	5	33	65	701
	35	15	4	8	3	3	1	9	43	244
	40	10	2	3	3	1	0	3	22	80
Central North Sea	20	5	0	0	0	0	1	0	6	11
	25	2	1	3	2	2	0	6	16	123
	30	9	5	6	2	1	1	13	37	248
	35	7	5	1	3	1	1	8	26	140
	40	4	2	0	1	0	0	0	7	12
Southern North Sea	20	3	0	0	1	0	1	1	8	21
	25	5	5	3	3	2	0	9	27	157
	30	5	4	4	2	0	2	13	30	222
	35	0	3	1	3	0	3	12	22	186
	40	4	2	3	3	0	0	3	15	57

puter-intensive technique that avoids the necessity of deriving formulae via difficult analytical arguments (Efron & Tibshirani 1993). The basic idea is to treat the data set as if it was the population, and to draw a large number of independent random bootstrap samples (resamples) from this set where all observations have an equal probability of being selected. The bootstrap samples are drawn by replacement, and their size usually equals the original number of observations in the data set. From each sample, a bootstrap replicate of the statistic of interest, here the average weight percentages for particular prey groups (Eq. 1), is calculated. Then, for instance, 95% approximate CIs can be defined by simply taking the 2.5th and 97.5th percentiles of the ordered values of the replications. This simple way of constructing non-parametric CIs is termed the percentile (PC) method. It assumes that the frequency distribution of the bootstrap replications is unbiased and fully describes the distribution function of the parameter of interest (Efron & Tibshirani 1993, Mooney & Duval 1993). However, if the original parameter estimate is biased, then the bootstrap replications will generally be biased as well. The bootstrap frequency distribution can be further used to assess the possible bias of the estimator. A good approximation of the bias is simply the difference between the expected value of bootstrap distribution and the observed value of the estimator (Efron & Tibshirani 1993, Mooney & Duval 1993). The bias-corrected (BC) method (Efron & Tibshirani 1993, Manly 1997), which introduces a correction to the PC method to account for the bias, is therefore adopted here to compute better CIs.

Three different resampling procedures have been applied to the stomach contents data in this study. The first procedure was a single-stage resampling considering individual fish as independent sampling units. The second involved the resampling of hauls instead of individual fish. In this procedure, once a haul was selected, then all the fish contained within the haul were sampled as a bulk (pooled). The third procedure was a 2-stage resampling with the intention of taking both the inter- and intra-haul variation into consideration. The first step here was to resample the hauls, and after a haul had been selected, the second step was resampling of the individual fish within the selected haul.

Precision of the bootstrap CIs will depend on the number of resamplings. According to Efron & Tibshirani (1993), about 1000 replications are required in order to make the bootstrap sampling variability acceptably low. However, the authors further noted that

more replications could be needed for a bootstrap statistic which depends on the extreme tails of a bootstrap distribution. In order to determine an appropriate bootstrap sample size to estimate CIs for average weight percentages, 100 interval estimates based on 2000, 3000 and 5000 replications, respectively, have been computed for 2 length classes, 30 and 40 cm, from the northern North Sea. These intervals have then been compared with independent estimates resulting from 100 000 resamplings. The 2 length classes represent different aspects of the present data. The 30 cm length class is rich in terms of number of stomachs sampled per haul, contrary to the 40 cm length class, which is poor in data, containing hauls with mostly 1 or few stomachs (Table 1).

The CIs obtained by both approaches are only applicable to individual prey categories, not to simultaneous distribution of 2 or more, since the proportions of various prey categories in the diet are not independent.

The above bootstrap procedures have been implemented in a BASIC programme which generates an ample number of independent pseudo-random bootstrap samples in any of the 3 resampling strategies. A copy of the programme is available from the first author of this paper (E. Mümtaz Tirasin).

It should be noted that application of the methodology introduced so far assumes that all fish caught in each haul are sampled. If the stomach contents data are collected by sub-sampling, then necessary adjustments must be done by applying appropriate weighting factors prior to the precision estimates.

RESULTS

Average weight percentages (%W) of 7 main prey groups in the diet composition of mackerel in the North Sea, together with their associated CIs estimated by 3 different bootstrap procedures for successive sub-areas and length classes, are presented in Tables 2, 4 & 5. CIs calculated for the same data by employing Eq. (2) and appropriate values of the *t*-distribution are given in Table 3 only for the mackerel from the northern North Sea. Using individual fish as basic units in calculations or in resampling (bootstrap Procedure 1) resulted in consistently smaller intervals than those estimated from both the calculations and the 2 other bootstrap procedures which were based upon hauls as the primary sampling unit. In only 2 cases were the CIs obtained from bootstrap Procedure 1 slightly wider than the intervals computed from bootstrap Procedure 2 (Table 5). The intervals estimated from the 2-stage resampling (Procedure 3) were in general moderately wider (1 to 10%) than those from Procedure 2. However, in 8 cases the differences in interval width were over 20%, and in 5 cases they were slightly smaller than those from Procedure 2 (Tables 2 & 4).

When individual fish were the units in the calculations, the limits and size of CIs computed following Cochran's approach (Eq. 2) often agreed closely with the bootstrap estimates (Tables 2 & 3). However, clear discrepancies in CIs estimated from these 2 different approaches were observed when calculations were based upon the hauls rather than individual fish. The discrepancies were particularly remarkable in length classes containing many hauls with 1 or 2 stomachs, such as 20 and 40 cm mackerel in the northern North Sea (Tables 1, 2 & 3).

CIs resulting from different bootstrap procedures illustrate that the most important source of stomach contents variability is the intra-haul correlation. This inference can be further supported with a demonstration of how the magnitude of variance estimates changes due to the different sources (procedures). Table 6 presents the bootstrap estimates of variance of average weight percentages from 3 different procedures. Only

3 length classes from dissimilar sub-areas were chosen to be shown in order to conserve space. Variance estimates suggest that the relative contribution of the intra-haul variation into overall variability of stomach contents composition is fairly modest.

Variability in CIs, estimated from bootstrap Procedure 3 with varying number of resamplings, is demonstrated for only 2 prey groups in Figs. 1. & 2. Intervals for the 40 cm length class were notably more variable than those for the 30 cm length class. Precision of the intervals increased consistently with increasing bootstrap samples for both length classes and all prey groups. According to the simulations, 3000 to 5000

Table 2. Main prey groups in the diet of mackerel in the northern North Sea in terms of percentage weight (%W) for successive length classes (LC, in cm). The bias-corrected (BC) 95% CIs were estimated by 3 different bootstrap procedures with 5000 resamplings as explained in the text. LowLm and UppLm are lower and upper limits of the intervals. No CIs were calculated for weight percentages amounting to less than 1%

LC	Prey group	%W	Procedure 1 LowLm UppLm		Procedure 2 LowLm UppLm		Procedure 3 LowLm UppLm	
20	Mollusca	0.01						
	Copepoda	20.86	7.26	35.74	0.10	48.79	0.03	49.73
	Amphipoda	0.05						
	Krill	42.20	22.45	59.24	14.42	75.25	13.92	76.94
	Other Crs.	34.73	19.58	55.91	3.49	77.23	3.73	77.38
	Fish	2.12	0.00	8.56	0.00	11.57	0.00	11.08
	Others	0.03						
25	Mollusca	0.08						
	Copepoda	30.55	24.52	37.50	13.31	52.80	12.80	52.56
	Amphipoda	0.23						
	Krill	35.80	27.53	43.84	15.36	56.64	16.17	57.66
	Other Crs.	7.59	5.13	10.61	2.52	17.29	2.34	17.82
	Fish	24.88	16.91	33.42	7.83	47.91	7.68	49.26
	Others	0.87						
30	Mollusca	0.19						
	Copepoda	16.46	13.88	19.27	9.27	26.91	8.95	26.67
	Amphipoda	0.95						
	Krill	47.84	42.73	52.86	29.39	64.24	29.05	64.03
	Other Crs.	5.16	4.03	6.54	2.15	10.28	1.97	10.83
	Fish	28.98	24.54	33.54	17.12	45.18	16.38	45.51
	Others	0.42						
35	Mollusca	0.29						
	Copepoda	7.86	5.35	10.84	2.97	14.84	2.83	15.62
	Amphipoda	0.16						
	Krill	46.26	37.59	55.28	18.36	72.87	17.00	72.44
	Other Crs.	1.63	0.84	2.66	0.47	3.40	0.38	3.77
	Fish	41.25	32.62	49.59	17.17	69.90	16.04	70.78
	Others	2.55	1.52	4.05	0.11	9.83	0.12	9.89
40	Mollusca	0.00						
	Copepoda	6.78	3.37	12.84	2.40	28.33	2.10	29.91
	Amphipoda	0.01						
	Krill	62.84	49.90	74.29	11.22	87.71	11.24	87.74
	Other Crs.	1.80	0.51	3.61	0.48	5.24	0.30	5.12
	Fish	27.62	16.59	39.82	5.94	79.32	5.49	79.29
	Others	0.95						

bootstrap samples are sufficient to construct stable 95% CI estimates for average weight percentages.

DISCUSSION

Precision of data on diet composition

The results provided here clearly demonstrate that precision of the estimated weight percentages was generally very poor, even when the prey items were grouped into major categories and sample sizes were comparatively large. Further evaluation and a comprehensive discussion regarding the results in the context of the food and feeding habits of mackerel in the North Sea is given by Tirasin (unpubl.).

The variability and subsequently the CIs of the average weight percentages were underestimated from both methods when individual fish were used as basic units in calculations. This choice is akin to simple random sampling which disregards any natural stratification in data. However, it is well known that numerous marine organisms often form schools, clusters or swarms due to biological and environmental factors varying over time and space (Steele 1976, Parsons et al. 1984, Valiela 1995). Consequently, predators will often concentrate on these patchy prey distributions. Thus, prey species and size composition in the stomach contents of the predators will tend to be more similar at a local scale (e.g. mackerel caught in the same haul) than in the population at a larger scale (Bogstad et al. 1995). Wider CIs estimated from the calculations based on hauls and 2-stage resampling clearly indicate that intra-haul correlation is a very important source of variation in the stomach contents composition.

These findings seem to have general validity and are not confined to schooling pelagic species. The only other dietary study providing estimates of the precision of prey composition is that of Jiang & Jørgensen (1996) on the diet of haddock (*Melanogrammus aeglefinus* L.) in the Barents Sea. These workers similarly observed wider CIs for the weight percentages of various prey groups when they used hauls as resampling units in bootstrapping, although haddock is a demersal fish with a markedly different ecology and feeding behaviour from mackerel. This inference is also in accordance with the remarks of Pennington & Vølstad (1994), i.e. that presence of intra-haul correlation, even at low levels, could greatly increase the variance of the estimate as compared with that obtained by treating the pooled data as a simple random sample.

Bootstrapping appears to be superior to Cochran's approach to estimate the CIs of average weight percentages. The intervals computed from the latter approach are symmetrical around the parameter esti-

Table 3. Diet composition by percentage weight (%W) for main prey groups for successive length classes (LC, in cm) in the northern North Sea. The 95% CIs were estimated by using Eq. (2) and appropriate values of the *t*-distribution. First group of estimates is based upon individual fish, while the second is based on hauls as explained in the text. LowLm and UppLm are lower and upper limits of the intervals. No CIs were calculated for weight percentages amounting to less than 1%

LC	Prey group	%W	Individuals		Hauls	
			LowLm	UppLm	LowLm	UppLm
20	Mollusca	0.01				
	Copepoda	20.86	5.82	35.90	0.00	48.85
	Amphipoda	0.05				
	Krill	42.20	22.73	61.67	7.54	76.86
	Other Crs.	34.73	15.95	53.52	0.00	82.39
	Fish	2.12	0.00	6.53	0.00	7.18
	Others	0.03				
25	Mollusca	0.08				
	Copepoda	30.55	24.02	37.08	9.70	51.40
	Amphipoda	0.23				
	Krill	35.80	27.48	44.11	14.00	57.59
	Other Crs.	7.59	4.90	10.28	0.43	14.75
	Fish	24.88	16.38	33.38	4.22	45.53
	Others	0.87				
30	Mollusca	0.19				
	Copepoda	16.46	13.71	19.22	7.89	25.04
	Amphipoda	0.95				
	Krill	47.84	42.76	52.92	30.11	65.57
	Other Crs.	5.16	3.92	6.41	1.07	9.26
	Fish	28.98	24.36	33.57	14.47	43.47
	Others	0.42				
35	Mollusca	0.29				
	Copepoda	7.86	5.13	10.59	1.81	13.91
	Amphipoda	0.16				
	Krill	46.26	37.21	55.32	15.97	76.55
	Other Crs.	1.63	0.71	2.54	0.29	2.96
	Fish	41.25	32.39	50.11	12.21	70.30
	Others	2.55	1.31	3.78	0.00	6.84
40	Mollusca	0.00				
	Copepoda	6.78	2.22	11.35	0.00	15.40
	Amphipoda	0.01				
	Krill	62.84	50.51	75.18	17.94	100.00
	Other Crs.	1.80	0.21	3.38	0.58	3.01
	Fish	27.62	15.94	39.30	0.00	68.69
	Others	0.95				

mate, and are not adjusted to account for the skewness in the underlying distribution. Bootstrap BC intervals are asymmetrical, adjusted to the shape of the empirical parameter distribution, and moreover corrected for bias in the estimator. In addition it is not possible to further arrange Eq. (2) to account for both the inter- and intra-haul variation. Thus, although intra-haul variation was found to be a comparatively smaller source in stomach contents variability, the CIs from the bootstrap Procedure 3 should be regarded as having better coverage properties than all other intervals presented here since they incorporate all sources of variability.

Table 4. Main prey groups in the diet of mackerel in the central North Sea in terms of percentage weight (%W) for successive length classes (LC, in cm). The bias-corrected (BC) 95% CIs were estimated by 3 different bootstrap procedures with 5000 resamplings as explained in the text. LowLm and UpLm are lower and upper limits of the intervals. No CIs were calculated for weight percentages amounting to less than 1%

LC	Prey group	%W	Procedure 1 LowLm UpLm		Procedure 2 LowLm UpLm		Procedure 3 LowLm UpLm	
20	Mollusca	0.00						
	Copepoda	28.67	6.37	85.98	0.00	89.34	0.00	90.18
	Amphipoda	58.66	0.03	80.71	0.02	88.69	0.02	89.00
	Krill	0.38						
	Other Crs.	10.33	2.51	23.75	2.78	32.33	2.36	35.73
	Fish	1.96	0.00	12.89	0.00	19.74	0.00	23.64
	Others	0.00						
25	Mollusca	0.02						
	Copepoda	6.53	2.82	13.45	0.93	24.19	0.71	22.70
	Amphipoda	50.78	36.90	70.03	29.58	75.11	30.34	76.41
	Krill	32.03	6.67	53.12	3.34	57.60	3.44	57.21
	Other Crs.	2.65	0.91	5.67	0.90	6.46	0.79	7.04
	Fish	6.71	1.66	20.41	1.59	23.53	1.43	24.56
	Others	1.28	0.68	2.50	0.17	4.90	0.18	5.08
30	Mollusca	2.00	0.69	4.17	0.29	7.80	0.26	7.51
	Copepoda	12.81	8.67	18.02	4.13	23.31	3.69	24.93
	Amphipoda	45.52	37.21	52.47	33.18	57.64	32.76	58.84
	Krill	15.85	9.64	22.07	4.20	27.32	4.42	28.68
	Other Crs.	6.48	4.22	9.84	3.47	10.94	2.81	11.68
	Fish	11.76	5.06	23.03	4.04	28.77	3.75	29.49
	Others	5.58	3.19	8.54	0.18	15.89	0.18	16.12
35	Mollusca	0.44						
	Copepoda	11.46	5.23	21.90	2.54	28.93	2.44	30.24
	Amphipoda	13.81	7.54	24.40	4.53	29.16	4.46	29.94
	Krill	25.50	5.93	45.66	3.13	50.54	2.12	50.56
	Other Crs.	9.32	3.03	14.79	2.47	16.26	2.25	16.43
	Fish	36.52	17.40	60.89	9.00	72.77	9.42	73.96
	Others	2.95	0.60	8.22	0.05	13.37	0.06	15.97
40	Mollusca	2.57	0.00	7.37	0.00	8.77	0.00	9.07
	Copepoda	8.25	0.04	35.09	0.04	36.48	0.04	45.03
	Amphipoda	9.18	2.08	22.81	1.07	26.08	0.92	27.18
	Krill	14.62	0.36	45.12	0.37	55.90	0.19	59.84
	Other Crs.	2.93	0.00	13.98	0.00	14.34	0.00	20.05
	Fish	61.51	15.64	89.44	0.33	93.15	0.22	94.83
	Others	0.94						

Like any approximate method based upon asymptotics, bootstrapping may also yield somewhat erratic estimates when sample sizes are small. Some authors raised concern about the efficiency of bootstrap for samples on the order of 10 to 20 (Schenker 1985, Smith 1985, Léger et al. 1992) while others expressed confidence in estimates resulting from small samples (Stine 1985, Bickel & Krieger 1989). Since the basis of bootstrapping is to resample the original observations, an adequate sample size is needed to obtain reliable measures of precision. This adequate sample size for CI estimation should be considered to be the same size as what is assumed to be representative to make

inferences about the population under study. For instance, the CIs estimated for the 20 and 40 cm length classes from the central North Sea should not be considered reliable since the data available for these length classes are hardly sufficient to be representative. No resampling method, regardless of the computer power, can replace a sound sampling procedure.

Implications for dietary studies

In addition to the quantitative description and evaluation of the relative importance of various prey items in food composition of a population or a certain group of fish, dietary studies also often focus on temporal, spatial, and ontogenetic variations in this composition. Sometimes investigations are extended to include comparisons of the food and feeding habits of different populations or species. Evaluation and comparisons have frequently been based on the average values of 1 or more of the common measures, i.e. %O, %N and %W. Since these averages were solely point estimates, workers have often worried about the correct interpretation of these measures. Some were concerned that the use of %W or %N overemphasized the importance of either very large or small but numerous food items even if they were eaten rarely (Hellawell & Abel 1971, Hyslop 1980, Eliassen & Jobling 1985, Bowen 1996). In such circumstances, the CIs accompanying average %W or %N estimates can facilitate a more objec-

tive assessment of the importance of various prey. Overall significance of any scarce but large or numerous prey items in the diet will be readily detected since their associated CIs would be comparatively very wide and likely include 0.

Several compound indices combining average values of 2 or all of the 3 measures have also been introduced in tabular or graphical forms (Pinkas et al. 1971, George & Hadley 1979, Mohan & Sankaran 1988, Assis 1996, Cortéz 1997, Marshall & Elliott 1997). A typical example, and also one of the more widely used compound indices, in both tabular or graphical forms, is the index of relative importance (Pinkas et al. 1971):

$$IRI_i = \% \hat{O}_i \cdot (\% \hat{N}_i + \% \hat{W}_i) \quad (3)$$

where i denotes a particular prey group (% W may sometimes be substituted by % V as well). It was assumed that different measures contained substantial independent information, and that the use of compound indices would prevent loss of information (Hyslop 1980, MacDonald & Green 1983, Assis 1996). Incorporating the different measures was thought to cancel out the biases in its individual components and seemed to provide a more accurate description (George & Hadley 1979, Cortéz 1997). Summarizing all information in 1 single figure was also supposed to simplify the assessment of the importance of various prey as well as their comparison, and thus was an additional incentive.

Hyslop (1980) questioned the claimed accuracy of such indices on the grounds that they integrated not only average values from different measures but also the errors and variation associated with them. MacDonald & Green (1983) later pointed out that the compound indices were redundant and added little new information to the single measures. The application of compound indices in dietary studies, nevertheless, has continued (Assis 1996, Tavares & Williams 1996, Cortéz 1997, Gray et al. 1997, Marriott et al. 1997). It is obvious from the results of this study that such compound indices should be no longer used. Combination of average percentages with poor precision and a great deal of uncertainty, in any way, will not yield any meaningful figure. On the contrary, it will merely obscure the actual information available from the separate measures.

Certain statistical techniques, rather than the above mentioned indices, were also suggested for use with comparison of diet compositions. For prey weight data, Crow (1979) recommended the application of non-parametric multivariate analysis of variance (MANOVA) to detect dietary differences among fish grouped according to any factor of interest such as location, season, size, habitat, etc. Somerton (1991) developed a non-parametric randomization test (Manly 1997) for a 1-way MANOVA to determine whether the weight proportions

of various prey categories varied between 2 predator groups. In this approach, significance was tested with Hotelling's T^2 , and post-hoc univariate t -tests were run to detect differences due to prey categories. Compared to any other compound index, Somerton's randomization test is certainly more sound and reliable, but restricted to pairwise comparisons. This testing procedure is, therefore, more appropriate for a modest number of predator groups. The presence of all prey categories in the diet of both predator groups is also required. However, a more important limitation is that it cannot be extended to evaluate the intra-haul variation in the data set.

Table 5. Main prey groups in the diet of mackerel in the southern North Sea in terms of percentage weight (% W) for successive length classes (LC, in cm). The bias-corrected (BC) 95% CIs were estimated by 3 different bootstrap procedures with 5000 resamplings as explained in the text. LowLm and UpLm are lower and upper limits of the intervals. No CIs were calculated for weight percentages amounting to less than 1%

LC	Prey group	%W	Procedure 1		Procedure 2		Procedure 3	
			LowLm	UpLm	LowLm	UpLm	LowLm	UpLm
20	Mollusca	0.02						
	Copepoda	3.12	0.41	6.82	0.37	6.27	0.05	8.78
	Amphipoda	3.80	0.00	12.96	0.00	13.37	0.00	22.14
	Krill	0.00						
	Other Crs.	81.83	69.49	91.61	52.61	94.48	48.13	96.87
	Fish	2.96	0.00	8.16	0.00	22.22	0.00	24.89
	Others	8.27	1.80	16.55	0.00	23.69	0.00	26.07
25	Mollusca	3.08	1.41	5.20	1.11	5.12	0.63	5.78
	Copepoda	35.69	25.11	46.32	15.61	53.70	13.43	55.26
	Amphipoda	2.53	0.05	7.87	0.06	9.83	0.03	10.00
	Krill	0.01						
	Other Crs.	22.19	14.80	30.36	13.55	34.01	11.43	37.29
	Fish	10.61	3.58	19.39	3.69	23.75	2.61	26.54
	Others	25.89	16.82	36.02	7.94	44.54	7.99	47.28
30	Mollusca	1.33	0.24	2.88	0.28	3.40	0.17	3.71
	Copepoda	37.12	26.50	48.93	19.60	55.53	18.08	57.34
	Amphipoda	0.60						
	Krill	0.04						
	Other Crs.	10.93	6.56	16.29	3.84	20.95	3.49	22.09
	Fish	33.11	18.74	46.71	12.94	55.05	10.31	57.95
	Others	16.87	9.93	25.06	5.66	29.96	4.75	31.36
35	Mollusca	1.60	0.60	2.99	0.59	3.58	0.42	4.14
	Copepoda	34.58	23.90	47.07	13.36	61.25	12.40	63.53
	Amphipoda	0.30						
	Krill	1.48	0.04	4.92	0.03	5.95	0.02	7.17
	Other Crs.	10.23	5.77	15.82	2.55	23.82	2.18	25.26
	Fish	47.66	31.54	61.30	17.14	73.45	14.81	74.08
	Others	4.15	1.77	7.57	1.21	9.54	0.95	10.53
40	Mollusca	1.64	0.37	4.16	0.31	3.56	0.24	5.04
	Copepoda	5.62	1.70	14.58	0.05	19.68	0.03	23.16
	Amphipoda	0.10						
	Krill	0.98						
	Other Crs.	1.45	0.17	4.04	0.01	5.27	0.00	6.63
	Fish	89.60	76.21	95.70	71.38	97.68	64.78	98.29
	Others	0.61						

Table 6. Approximate variance estimates of average weight percentages, VAR(1 to 3), from 3 different bootstrap procedures (1 to 3) with 5000 resamplings. (SA) sub-areas and (LC) mackerel length classes. R(1,2) are the ratios of variances VAR(2)/VAR(1) and VAR(3)/VAR(2), respectively

SA and LC	Prey group	VAR(1)	VAR(2)	VAR(3)	R(1)	R(2)
Northern North Sea 25	Copepoda	11.12	103.50	106.86	9.31	1.03
	Krill	17.37	110.51	117.24	6.36	1.06
	Other Crs.	1.97	13.69	15.33	6.95	1.12
	Fish	17.78	96.49	105.14	5.43	1.09
Central North Sea 30	Mollusca	0.77	2.99	3.05	3.91	1.02
	Copepoda	5.76	25.32	29.39	4.39	1.16
	Amphipoda	14.96	38.18	43.05	2.55	1.13
	Krill	9.94	35.73	39.62	3.60	1.11
	Other Crs.	2.09	3.70	5.26	1.77	1.42
	Fish	20.16	39.02	41.01	1.94	1.05
Southern North Sea 35	Others	1.80	15.09	15.91	8.39	1.05
	Mollusca	0.41	0.63	0.99	1.52	1.57
	Copepoda	35.14	161.29	179.47	4.59	1.11
	Krill	1.98	2.67	4.60	1.34	1.73
	Other Crs.	6.99	30.51	35.14	4.36	1.15
	Fish	57.71	217.80	244.39	3.77	1.12
	Others	2.24	4.34	6.32	1.94	1.45

This last constraint concerning the intra-haul correlation is equally valid for other statistical techniques, such as the contingency table (log-linear model) analysis, which is an analogous approach to MANOVA and recommended for prey counts data by Crow (1979, 1982) and Cortéz (1997).

CONCLUSIONS

Accompanying precision estimates should be an integral part of all quantitative diet studies because the point estimates from the dietary measures, and hence the relative importance of prey items, can be evaluated reasonably only in relation to associated CIs. The BC CIs estimated from the bootstrap Procedure 3, which incorporates both inter- and intra-haul variation, are suitable tools for a realistic assessment for the reliability of the dietary measures given that sampling is done properly to make valid inferences about the population.

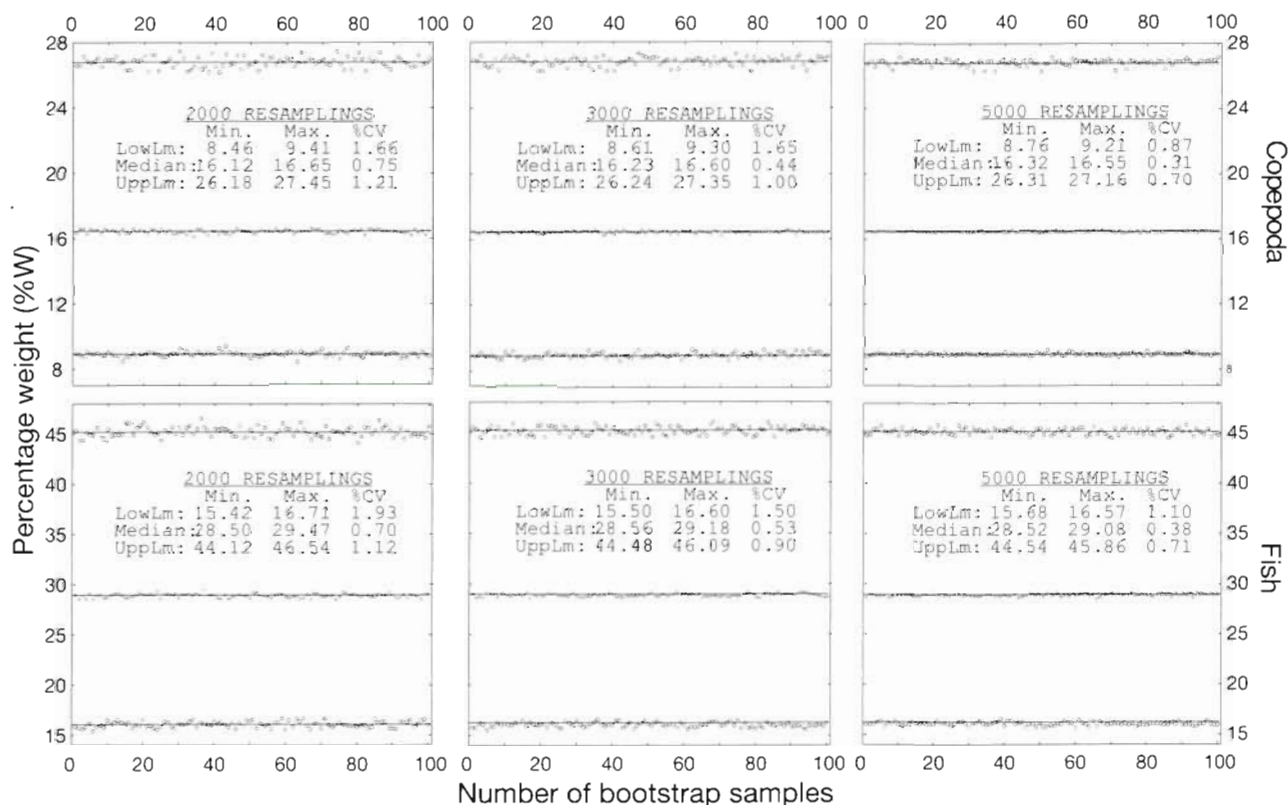


Fig. 1. Middle lines are the average weight percentages of the prey groups Copepoda and Fish in the diet of the 30 cm length class (Table 2). Upper and lower lines are the upper and lower BC 95% confidence limits estimated from 100000 resamplings. Data points plotted around these lines show the upper and lower limits of the BC 95% CIs computed with varying resampling size. Observations around the middle lines are median values of the respective bootstrap distributions. Percentage coefficients of variation (%CV) have been calculated for estimated interval limits and medians as a measure of precision

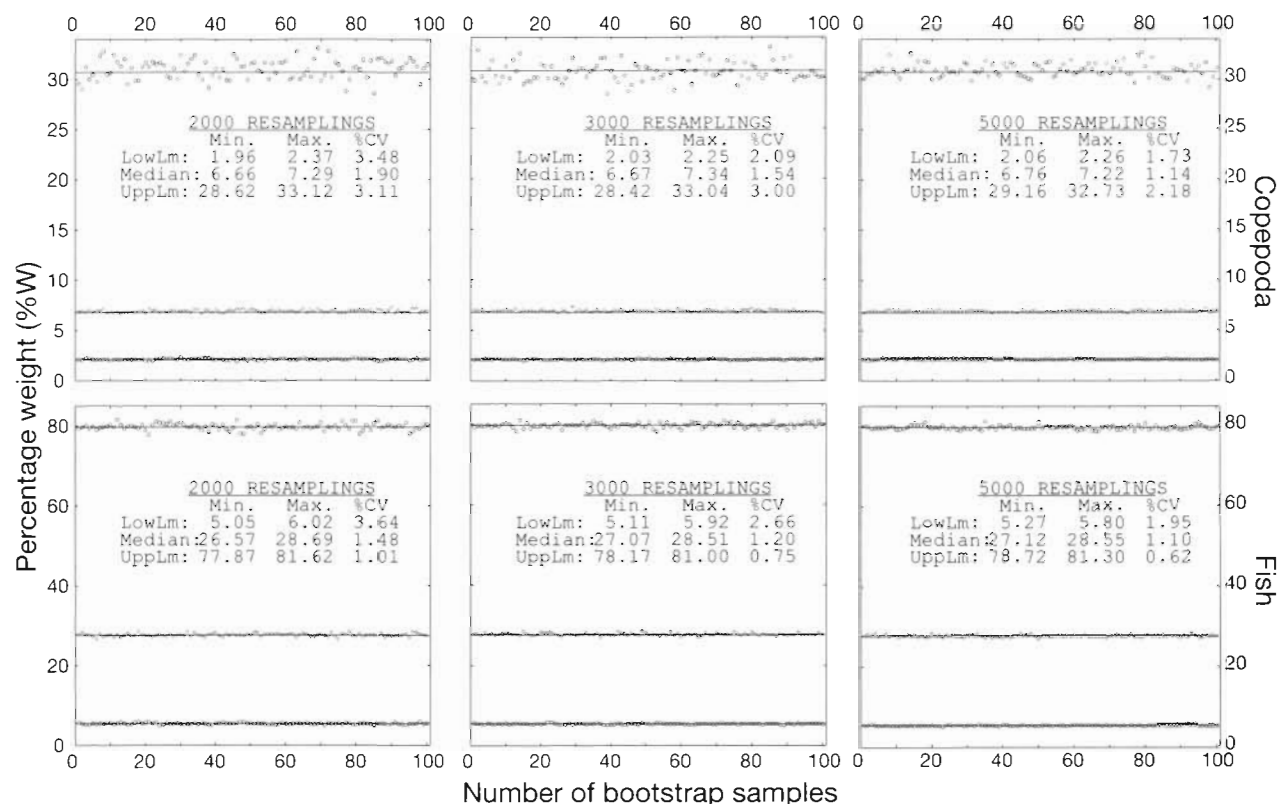


Fig. 2. All lines and the data points plotted around these lines are the same as described in Fig. 1. Computations are based on the dietary data of the 40 cm length class (Table 2)

Acknowledgements. Michael Pennington, Arne Johannessen, Dankert Skagen, Sara Adlerstein and Pedro de Barros are gratefully acknowledged for their discussions and critical comments on the manuscript. Three anonymous reviewers are also thanked for their constructive comments. E.M.T was supported partly by a scholarship from the Institute of Marine Sciences and Technology of Dokuz Eylül University (Turkey), and partly by a grant from the European Commission (Project EC-AIR3-CT94-2410).

LITERATURE CITED

- Assis CA (1996) A generalised index for stomach contents analysis in fish. *Sci Mar* 60:385–389
- Berg J (1979) Discussion of methods of investigating the food of fishes, with reference to a preliminary study of the prey of *Gobiomus flavescens* (Gobiidae). *Mar Biol* 50: 263–273
- Bickel PJ, Krieger AM (1989) Confidence bands for a distribution function using the bootstrap. *J Am Statist Ass* 84: 95–100
- Bogstad B, Pennington M, Vølstad JH (1995) Cost-efficient survey design for estimating the food consumption by fish. *Fish Res* 23:37–46
- Bowen SH (1996) Quantitative description of the diet. In: Murphy BR, Willis DW (eds) *Fisheries techniques*, 2nd edn. American Fisheries Society, Bethesda, Maryland, p 513–532
- Cailliet GM (1977) Several approaches to the feeding ecology of fishes. In: Simenstad CA, Lipovsky SJ (eds) *GUTSHOP* 76 Fish food habit studies. *Proc 1st Pacif Northwest Tech Workshop*. Washington Sea Grant Publ., Seattle, p 1–13
- Cochran WG (1977) *Sampling techniques*, 3rd edn. John Wiley & Sons, New York
- Cortéz E (1997) A critical review of methods of studying fish feeding based on analysis of stomach contents: application to elasmobranch fishes. *Can J Fish Aquat Sci* 54:726–738
- Crow ME (1979) Multivariate statistical analysis of stomach contents. In: Lipovsky SJ, Simenstad CA (eds) *GUTSHOP* 78 Fish food habit studies. *Proc 2nd Pacif Northwest Tech Workshop*. Washington Sea Grant Publ., Seattle, p 80–86
- Crow ME (1982) Some statistical techniques for analysing the stomach contents of fish. In: Cailliet GM, Simenstad CA (eds) *GUTSHOP* 81 Fish food habit studies. *Proc 3rd Pacif Workshop*. Washington Sea Grant Publ., Seattle, p 8–15
- Efron B (1979) Bootstrap methods: another look at the jack-knife. *Ann Statist* 7:1–26
- Efron B, Tibshirani RJ (1993) *An introduction to the bootstrap*. Chapman and Hall, New York
- Eliassen JE, Jobling M (1985) Food of roghhead grenadier *Macrourus berglax*, Lacépède in North Norwegian waters. *J Fish Biol* 26:367–376
- Ferry LA, Cailliet GM (1996) Sample size and data analysis: are we characterizing and comparing diet properly? In: MacKinlay D, Shearer K (eds) *Feeding ecology and nutrition in fish*. *Proceedings of the Symposium on the Feeding Ecology and Nutrition in Fish*, International Congress on the Biology of Fishes. Am Fish Soc, San Francisco, CA, p 71–80
- George EL, Hadley WF (1979) Food and habitat partitioning between rock bass (*Ambloplites rupestris*) and small-

- mouth bass (*Micropterus dolomieu*) young of the year. *Trans Am Fish Soc* 108:253–261
- Gray AE, Mulligan TJ, Hannah RW (1997) Food habits, occurrence, and population structure of the bat ray, *Myliobatis californica*, in Humboldt Bay, California. *Environ Biol Fishes* 49:227–238
- Hall SJ, Raffaelli D, Basford DJ, Robertson MR, Fryer R (1990) The feeding relationships of the larger fish species in a Scottish Sea loch. *J Fish Biol* 37:775–791
- Hellawell JM, Abel R (1971) A rapid volumetric method for the analysis of food of fishes. *J Fish Biol* 3:29–37
- Hyslop EJ (1980) Stomach contents analysis: a review of methods and their application. *J Fish Biol* 17:411–429
- Jiang W, Jørgensen T (1996) The diet of haddock (*Melanogrammus aeglefinus* L.) in the Barents Sea in the period 1984 to 1991. ICES (Int Councl Explor Sea) *J Mar Sci* 53:11–21
- Léger C, Politis DN, Romano JP (1992) Bootstrap technology and applications. *Technometrics* 34:378–398
- MacDonald JS, Green RH (1983) Redundancy of variables used to describe importance of prey species in fish diets. *Can J Fish Aquat Sci* 40:635–637
- Manly BFJ (1997) Randomization, bootstrap and Monte Carlo methods in biology, 2nd edn. Chapman & Hall, London
- Marriott MS, Booth AJ, Skelton PH (1997) Reproductive and feeding biology of the Natal Mountain catfish, *Amphilus natalensis* (Siluriformes: Amphiliidae). *Environ Biol Fishes* 49:461–470
- Marshall S, Elliott M (1997) A comparison of univariate and multivariate numerical and graphical techniques for determining inter- and intraspecific feeding relationships in estuarine fish. *J Fish Biol* 51:526–545
- Mohan MV, Sankaran TM (1988) Two new indices for stomach content analysis of fishes. *J Fish Biol* 33:289–292
- Mooney CZ, Duval RD (1993) Bootstrapping: a nonparametric approach to statistical inference. Sage Publ, CA
- Parsons TR, Takahashi M, Hargrave B (1984) Biological oceanographic processes, 3rd edn. Pergamon, Oxford
- Pennington M, Vølstad JH (1994) Assessing the effect of intra-haul correlation and variable density on estimates of population characteristics from marine surveys. *Biometrics* 50:725–732
- Pinkas L, Oliphant MS, Iverson ILK (1971) Food habits of albacore, bluefin tuna and bonito in Californian waters. *Calif Dep Fish Game Fish Bull* 152:1–105
- Schenker N (1985) Qualms about bootstrap confidence intervals. *J Am Statist Ass* 80:360–361
- Seber GAF (1973) The estimation of animal abundance and related parameters. Hafner Press, New York
- Smith EP (1985) Estimating the reliability of diet overlap measures. *Environ Biol Fishes* 13:125–138
- Sokal RR, Rohlf FJ (1995) Biometry, 3rd edn. WH Freeman Co., New York
- Somerton DA (1991) Detecting differences in fish diets. *Fish Bull* 89:167–169
- Steele JH (1976) Patchiness. In: Cushing DH, Walsh JJ (eds) *The ecology of the seas*. Blackwell, London, p 98–115
- Stine RA (1985) Bootstrap prediction intervals for regression. *J Am Statist Ass* 80:1026–1031
- Tavares CAF, Williams DD (1996) The importance of temporal resolution in food web analysis: evidence for a detritus-based stream. *Ecol Monogr* 66:91–113
- Valiela I (1995) Marine ecological processes, 2nd edn. Springer Verlag, New York
- Windell JT, Bowen SH (1978) Methods for study of fish diets based on analysis of stomach contents. In: Bagenal T (ed) *Methods for assessment of fish production in fresh waters*, 3rd edn. Blackwell Sci Publ, Oxford, p 219–226

Editorial responsibility: Otto Kinne (Editor),
Oldendorf/Luhe, Germany

Submitted: May 19, 1998; Accepted: January 7, 1999
Proofs received from author(s): May 12, 1999