Species richness estimation: Estimator performance and the influence of rare species

Katharina Reichert[1]*[†], Karl Inne Ugland[2,†], Inka Bartsch[3], Joaquín Hortal[4], Julie Bremner[2,5] and Alexandra Kraberg[1]

[1]Biologische Anstalt Helgoland, Alfred-Wegener Institute for Polar and Marine Research, 27498 Helgoland, Germany
[2]Marine Biology Research Group, Department of Biology, University of Oslo, P.b. 1066 Blindern, NO-0316 Oslo, Norway
[3]Alfred-Wegener Institute for Polar and Marine Research, 27570 Bremerhaven, Germany
[4]Centre for Population Biology, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK
[5]Centre for Environment, Fisheries and Aquaculture Science, Pakefield Road, Lowestoft, Suffolk, NR33 0HT, UK

Abstract

Species richness estimators (SREs) frequently exhibit widely different results when applied to the same data set. The lack of information on the true species number in an area hampers assessment of the performance of competing SREs. Marine invertebrates have been studied for more than 150 years on Helgoland (North Sea), and a fairly complete species list exists for this assemblage, providing a good opportunity for such comparisons. We evaluated the performance of 12 SREs using data on intertidal marine invertebrates of Helgoland. The Ugland TS estimator provided the most accurate estimate of total richness, whereas MMF, Semilog, and Weibull produced a moderate underestimation. The other SREs seriously under- or overestimated total richness. Although the Ugland TS estimator accounts for changes of species abundance distributions encountered when area increases, most SREs assume a homogeneous distribution. Analysis of separate accumulation curves for rare, intermediate, and common species revealed that the curvature of the randomized empirical accumulation curve of all the species is determined primarily by the occurrence of rare species. Thus, the influence of rare species was assessed through simulations of the Ugland TS. Linear extrapolation will under- or overestimate total richness if the occupancy of rare species is unusually low or high, respectively. The true occupancy of rare species in the Helgoland intertidal is likely to be close to the optimal interval revealed by our simulations, which provide a general framework for assessing the suitability of SREs for a given data set and thus improve the selection process of the optimal method.

Obtaining complete information about the composition of a given species assemblage is a difficult task. Usually, no inventory is so exhaustive that all species are recorded, either locally or regionally (although exceptions exist for a few taxonomic groups, e.g., Condit et al. 2005). As such, species richness estimators (SREs) play a key role in the measurement of biodiversity (Soberón and Llorente 1993; Colwell and Cod-

dington 1994). These estimators can be used to standardize the effects of uneven sampling effort, providing unbiased measures of species richness that minimize measurement errors and improve the analysis of biodiversity patterns (Hortal et al. 2004; Borges et al. 2009).

A number of techniques are available to estimate species richness from a limited collection of samples (*see* reviews in Colwell and Coddington 1994; Gotelli and Colwell 2001; Magurran 2004). Most of these techniques fall into two categories: (1) SREs based on the extrapolation of species accumulation curves and (2) nonparametric estimators (but *see* Hortal et al. 2006). Species accumulation curves are plots of the cumulative number of species recorded with increasing levels of sampling effort, frequently measured as number of samples, which is also a measure of the sampled area. Total richness is estimated by fitting the parameters of a given function to these plots and extrapolating the function to either an infinite sampling effort or the total area studied. Nonparametric estimators, in contrast, are based on particular features of the

samples, such as the number of rare species found only once or twice in all samples. Analytical expressions of these techniques provide estimates of total richness using either presence/absence or abundance data.

To evaluate the performance of different SREs, it is crucial to have an assessment criterion for the total species number from a source that is independent of the investigation and not generated by the models themselves. Such an assessment criterion for the total species number can be established by, e.g., a species reference list. In this case, the comparison of SRE performance can be accomplished by means of a direct comparison of the estimates with the richness value determined from the reference list. In many studies, however, the actual richness values are not known and the values used for comparison are either expert estimates (e.g., Hortal et al. 2006) or obtained from the asymptote of a model fitted to the randomized accumulation curve (e.g., Canning-Clode et al. 2008). These comparisons are conceptually flawed (particularly in the latter case) because an estimate is used to assess the accuracy of others.

Accordingly, the first objective of this work was to generate an independent and convincing estimate of true species number and compare 12 frequently applied SREs with this approximated true species number. For the intertidal rocky platform of the island of Helgoland, a large number of recent and historical investigations provided the opportunity to establish such a reference list. An intense ecological sampling program conducted between 2004 and 2006 (Reichert et al. 2008a) formed the database for the comparison of SRE performance with the reference list.

SREs are used to estimate total species richness from a limited number of subsamples of the studied assemblage and/or territory. However, these subsamples often constitute a surprisingly small fraction of the total area studied, even when the level of sampling effort is high (Hortal et al. 2007). In general, only a tiny fraction of the territory is surveyed, fractions in the range 10^{-6}–10^{-4} being typical; thus an extremely disproportionate ratio between sampled and unsampled area frequently exists. This sampling effect interacts with the spatial distribution of individuals in determining survey success. Because the majority of species are rare, most will not appear in the sampled area and will be surveyed only within larger areas (*see* e.g., Preston 1948; Storch et al. 2003). Therefore, differences in species distributions may compromise the reliability of the SRE. Additionally, heterogeneity, in terms of spatial distribution of habitats, and spatial population dynamics of species related in turn to the distribution of species and their abundances in an area may result in unrealistic estimates of species richness. In previous studies, however, the focus is usually on the sampling strategy (e.g., number of samples, grain size) and how it affects the estimators' performance without incorporating spatial patterns of variation in species abundance distributions (e.g., Hortal et al. 2006; Canning-Clode et al. 2008).

Thus, a second objective of this study was to further explore the influence of species abundance distributions, in terms of the rarity of species, on the best-performing estimator of the intertidal data set. For this purpose, a simplified data set was developed modeling the empirical intertidal data set to show how different occurrence probabilities of rare species may influence the bias of this SRE.

Materials and procedures

Study location and biological data set—Helgoland is a small North Sea island system located 60 km off the German coast (54°11'N, 7°55'E). The island is approximately 1 km² and has a coastline composed mainly of rocky outcrops. The rocky littoral is surrounded by soft sediments, geographically isolated by hundreds of kilometers from other hard substrates in the North Sea.

Our analysis focuses on the intertidal zone of the rocky platform at the island's northern shore. This zone is characterized by a series of channels extending northwest toward the open sea, separated by ridges. The ridges and channels alternate in the alongshore direction, and the substratum is mostly relatively soft natural red sandstone. For more details of the site, *see* Reichert et al. (2008a,b) and Bartsch and Tittley (2004).

We used data from a recent intensive survey of marine invertebrates at the north shore (Reichert et al. 2008a) to test the accuracy of the different SREs. Invertebrates were sampled at 3-month intervals from summer 2004 to spring 2006 (in total eight samplings in time). Five replicated 0.25 m² quadrats (~1 meter apart) were placed randomly at each of 11 plots. The plots, separated by tens of meters, were chosen at random from a larger pool of plots examined during a previous study (Reichert et al. 2008b). Fifty-five quadrats were sampled at each sampling time, yielding a total of 440 samples and 60 species. Given a total area of 32,909 m² (based on geographic information system [GIS] data) and that these quadrats cover 110 m², it follows that the survey covered 0.33% of the total area of the intertidal rock platform of the island's northern shore (the actual focus of the survey).

Analysis 1: Calibration of SREs with reference lists—Species reference list: Helgoland has a long history of environmental and taxonomic marine data collection, particularly since the establishment of the Biologische Anstalt Helgoland in 1892 (*see* Franke et al. 2004). This allows us to establish a comprehensive species reference list and, therefore, determine total richness with accuracy.

A literature search was carried out to establish a species reference list that reflects the total species number of invertebrates on the northern shore intertidal system at Helgoland. Hoffmann (1829) conducted the first investigation of intertidal macrozoobenthos, followed by 17 further works during the 19th century (e.g., Heincke 1894). More recently, Reichert and Buchholz (2006) carried out comprehensive surveys of intertidal invertebrate assemblages. The latest investigation of intertidal invertebrates was undertaken by Reichert et al. (2008a,b), and this data set is stored in a database developed by LargeNet, a responsive mode project of the MarBEF (Marine

Biodiversity and Ecosystem Functioning) EU Network of Excellence (Vandepitte et al. 2010). In total, 36 studies referring to the area under investigation and published between 1829 and 2008 were consulted to create the species reference list. All macrobenthic invertebrates mentioned in these studies for the northern shore intertidal system have been included in the reference list.

After the literature search, we used the EU Register of Marine Species (Costello et al. 2004) to identify the valid species names, solving all synonymies and taxonomic problems. This resulted in a list of 249 species. In addition, all species that could not be sampled through the sampling procedure used in Reichert et al. (2008a) were excluded, following three criteria:

(1) All vagile species were excluded, as these were not sampled by Reichert et al. (2008a). Here, vagiles are defined as species able to move relatively fast (e.g., amphipods). This criterion reduced the list to 158 species, leaving only sessile and slow-moving invertebrates (e.g., snails).

(2) Morphologically similar species combined into a species complex by Reichert et al. (2008a) were also grouped into species complexes in the reference list. This further reduced the list to 137 species/complexes.

(3) Habitat specialists (e.g., species only occurring in the supralittoral zone or living in dead wood) were also excluded from the reference list, as they were not sampled by Reichert et al. (2008a).

This resulted in a list of 134 species or species complexes of sessile and slow-moving macrobenthic invertebrates in the north shore area. We double-checked this list to ensure that no species currently extinct from the island and/or unable to appear in the rocky intertidal north shore was included. All species were recorded in the 1990s and/or later, and none of the listed invertebrates is a specialist of a habitat not present in the studied area. Thus, given the exhaustive nature of all historical and recent investigations at Helgoland and the strict application of all criteria, we assume that the final reference list presents an accurate picture of the species living in the north shore. As a consequence, we conclude that the adapted reference list (134 species) is the "true" richness of the studied area and can thus be used as a baseline for tests of the accuracy of the different SREs.

Species richness estimators: We assessed the performance of 12 different SREs. Five were nonparametric estimators that use presence/absence data: Chao2 (Chao 1984, 1987), ICE (incidence-based coverage estimator; Lee and Chao 1994; Chao et al. 2000), Jackknife1 and Jackknife2 (both developed by Burnham and Overton 1978, 1979), and Bootstrap (Smith and van Belle 1984). Quantitative descriptions of these SREs can be found in Colwell (2006), Magurran (2004), and the original papers. These SREs were calculated using the EstimateS software developed by Colwell (2006). Evaluations of their performance and their comparability are available in, for example, the review from Walther and Moore (2005), Hortal et al.

(2006), and references therein. In these comparative works (among others), nonparametric estimators, particularly the two from Chao and the two Jackknives, perform adequately in many cases, in terms of estimation bias, precision, and accuracy. However, these estimators have also been reported to fail in providing reliable estimates, for, e.g., terrestrial plant communities (Chiarucci et al. 2003) or marine assemblages (Chapman and Underwood 2009). Such assemblages are characterized by numerous rare species. Therefore, it seems apparent that nonparametric estimators perform poorly under these circumstances.

Six curve-based SREs were estimated from the smoothed observed species accumulation curve, obtained using an analytical expression for the randomization of the samples (Ugland et al. 2003). Two of these estimators were non-asymptotic species accumulation curves: Power (the original Arrhenius' power law; Arrhenius 1921) and Semilog (the semilogarithmic plot of Gleason 1922). Another four were asymptotic species accumulation curves: Michaelis-Menten (established originally for enzyme kinetics by Michaelis and Menten 1913 and first used as an SRE by Clench 1979), Negative Exponential (also described by Arrhenius 1921; Soberón and Llorente 1993), Weibull (1951), and Morgan-Mercer-Flodin (MMF) (Morgan et al. 1975; Lambshead and Boucher 2003). The six curve-based SREs (either asymptotic or non-asymptotic) were fitted to the smoothed observed accumulation curve by use of the comprehensive curve fitting program CurveExpert (Hyams 2005), and total species richness was calculated by extrapolating to either the asymptote of the fitted curve or the total area of the north shore. The original references and Tjørve (2003) give the formulations of these curve-based SREs; Tjørve (2003), Hortal et al. (2006), Jiménez-Valverde et al. (2006), and Melo et al. (2007) provide analyses of their performance.

Finally, we also evaluated the Ugland TS estimator (Ugland et al. 2003), which accounts for the degree of environmental heterogeneity (e.g., depth or sediment properties) and the size of the whole area by partitioning the data set of the sampled area into several subsets—a step not required for the other estimators. Total species richness is here estimated in a semilogarithmic plot by extrapolating the linear regression through the end points of the nested accumulation curves.

We assessed accuracy for the 12 different SREs by measuring prediction errors (*see* Walther and Moore 2005). Because we calculated a single estimate per SRE and the "true" number of species was already known (134), we used absolute error, standard error, and percentage error as measures of accuracy. Absolute error is the difference between the "true" and predicted species richness, and percentage error is the absolute error divided by the "true" species richness (e.g., Hortal et al. 2006).

Standard error of the Ugland TS estimator: There are currently no measures of the dispersion of the estimates of the Ugland TS estimator. Here, we applied a modification of the bootstrap technique (Efron 1979) to estimate the standard error of the Ugland TS estimator. Our procedure is best

explained by referring to the Helgoland data set, where the samples were subdivided into 10 subsets: 1–40 (a), 41–80 (b), 81–120 (c), 121–160 (d), 161–200 (e), 201–240 (f), 241–290 (g), 291–340 (h), 341–390 (i), 391–440 (j). Within each of these 10 subsets we made a new subset of samples by sampling with replacement. Take, for example, subset (f): a new representation of this subset is obtained by selecting 40 samples with replacement from the 40 integers {201, 202, ..., 240}. Repeating this technique for all the 10 subsets gave a new data matrix of the resampled original data set. We then ran the Ugland TS method on this new data matrix to obtain the first bootstrap estimate. We repeated this process 100 times and calculated the standard error from these 100 bootstrap estimates.

Analysis 2: Model development for testing the influence of species rarity—Analysis 1 revealed the Ugland TS estimator as the best-performing SRE (*see* "Results"). Therefore, the influence of rare species occurrence on estimator bias was investigated for this SRE only.

Parameters of model data set: To test whether the occurrence of rare species has an impact on the bias of the Ugland TS estimator, we developed a simplified model system comparable in basic parameters to the empirical invertebrate data set of the northern intertidal of Helgoland and calculated 21 alternative values of the average fraction that a rare species occupies in the total area.

The following assumptions for the model data set were made:
1. The total area has the same size as the whole northern intertidal of Helgoland (32,909 m²) and thereby comprises $A = 131,636$ samples of 0.25 m².
2. The total area is homogenous and exhibits only one type of community.
3. The subsets have the same sample size as the subsets of the invertebrate data set (i.e., 40, 80, 120,... 440).
4. The spatial distributions of species are not correlated with each other.
5. All rare species occur in exactly the same number of samples and, therefore, exhibit the same p values.
6. The total number of species $S_{Total} = 150$.
7. The total number of rare species $S_{Rare} = 100$.
8. The total number of common and intermediate species $S_{Com} + S_{Interm} = 50$.

Note that assumptions 2, 4, and 5 are simplified assumptions and do not represent the situation of the empirical data set, which is more complex (several communities, correlated occurrence of species, and different p values for the occurrence of rare species).

As shown by Ugland and Gray (1982), species in an assemblage may be classified as (1) rare, (2) intermediate, or (3) common. Classification boundaries are normally set by expert judgment. For the empirical invertebrate data set, we used the following definitions:
1. Species observed in fewer than six samples (i.e., less than 1.3% of the 440 samples) were classified as rare. There were $S_{Rare} = 26$ species.

2. Species observed in more than 44 samples (i.e., more than 10% of the 440 samples) were classified as common. There were $S_{Com} = 13$ species.
3. All other species were classified as intermediate; in total $S_{Interm} = 21$ species were recorded.

It is essential to note that the distribution of 26 rare, 21 intermediate, and 13 common species in the empirical data set is based on a small collection of 440 samples in an area that would need 131,636 samples to be fully covered. Thus, to estimate species richness in the whole area, it is necessary to perform an extrapolation for each of these subgroups to a huge unsampled area. The extrapolation relies on exhaustive field evidence coming from a large number of recent and historical investigations, representing almost the entire area of the northern intertidal of Helgoland. Such evidence was already integrated in the species reference list.

Thus, we assume that the assumption for S_{Total}, S_{Rare}, and $S_{Com} + S_{Interm}$ in the model data set presents an accurate picture of the number of rare, intermediate, and common species in the whole northern intertidal. The observed number of common and intermediate species (i.e., 13 + 21 = 34) is likely close to the true number of common and intermediate species for the whole northern intertidal. In the simulations, the number of common and intermediate species, $S_{Com} + S_{Interm} = 50$, were chosen to allow for a possible occurrence of 16 more common or intermediate species (i.e., approximately 50% more species in these two groups) in the whole area. Further, the observed 26 rare species is likely quite far from the true number of rare species in the whole northern intertidal, and thus in total 100 rare species were chosen in the model data set.

On the bias of the Ugland TS estimator—Model deduction: Exact expressions for the expected number of species in any sample of a square from the whole area ($A = 131,636$ squares) may be obtained from the hypergeometric distribution (Ugland et al. 2003). For large sample sizes, however, the binomial distribution provides good approximations, i.e., the expected number of species in a random sample of k quadrats is

$$S(k) = S_{Total} \times [1 - (1-p)^k].$$

In our model data set, we assume three subgroups with three different p values. This gives the following three models for the expected number of species of each abundance group in k samples:

$$S_{Rare}(k) = S_{Rare}[1 - (1 - p_{Rare})^k]$$

$$S_{Interm}(k) = S_{Interm}[1 - (1 - p_{Interm})^k]$$

$$S_{Com}(k) = S_{Com}[1 - (1 - p_{Com})^k]$$

where S_{Rare}, S_{Interm}, and S_{Com} are the expected number of rare, intermediate and common species, respectively, and p_{Rare}, p_{Interm}, and p_{Com} represent the average fraction of squares in the whole area where respectively a rare, an intermediate, and a common species is found.

The randomized accumulation curves for rare, intermediate, and common species are shown in Fig. 1A–C. It is seen that there is a perfect model fit in all three groups. However, if we try to fit this type of model, $S(k) = S_{Total} \times [1 - (1 - p)^k]$, to the whole empirical data set, the theoretical expression deviates substantially from the empirical curve (Fig. 2A). This is because the binomial model is not additive; i.e., different curvature of the three accumulation curves leads to a different type of functional form. Therefore, the three groups should be treated individually when calculating the expected number of species in k samples:

$$S(k) = S_{Rare}[1 - (1 - p_{Rare})^k] + S_{Interm}[1 - (1 - p_{Interm})^k] + S_{Com}[1 - (1 - p_{Com})^k]$$

Note that this summation gives a perfect fit (Fig. 2B) simply because each of the three groups exhibits a perfect fit on its own (Fig. 1A–C).

A closer look at Figs. 1 and 2 provides the key to a tremendous simplification. Because the sample size (440) of the empirical invertebrate data set is relatively large and covers the whole environmental gradient and all seasonal aspects, it is assumed that most of the intermediate and common species have been detected in the investigation. Almost all (if not all) of the new species found beyond the already observed 60 species will be rare species, i.e., species with a narrow spatial or seasonal distribution and/or low densities and therefore difficult to sample. Looking at the single curves again (Fig. 1A–C), we may infer from the rapid asymptotic behavior of the intermediate and common group that the accumulation curve beyond 440 samples (i.e., extrapolation to the whole population with its 131,636 quadrats) will be dominated by the occurrence of rare species and will rapidly converge in functional form to the accumulation of the rare species. This is a strong argument for using the binomial model for each of the three subgroups when studying the performance of the Ugland TS extrapolation. Thus, when the sampled number of quadrats is large, the expected number of species may be calculated by the following approximation:

$$S(k) = S_{Com}[1 - (1 - p_{Com})^k] + S_{Interm}[1 - (1 - p_{Interm})^k] + S_{Rare}[1 - (1 - p_{Rare})^k]$$

$$\approx S_{Com} + S_{Interm} + S_{Rare}[1 - (1 - p_{Rare})^k]$$

$$= S_{Com} + S_{Interm} + S_{Rare} - S_{Rare}(1 - p_{Rare})^k$$

$$= S_{Obs} - S_{Rare}(1 - p_{Rare})^k$$

This model for the empirical accumulation curve has only three parameters: the number of observed species (S_{Obs}), the number of observed rare species (S_{Rare}), and the average fraction of samples where a rare species is observed. From Fig. 3 we see that the performance of this approximation is very good when the number of samples passes 40. This surprisingly low boundary for a good approximation is due to the rapid detection of intermediate and common species.

On the basis of the success of the approximation formula for the accumulation curve of the empirical data set (*see* Fig. 3), we applied the same formula to the whole intertidal area of
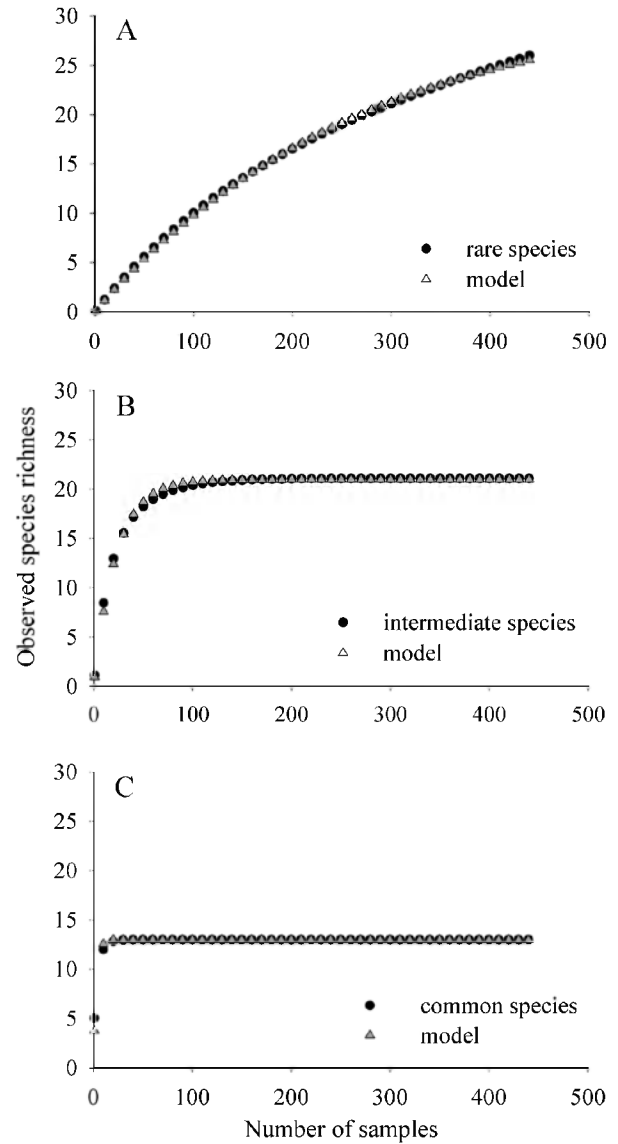

Fig. 1. The randomized empirical accumulation curve and the theoretical expectation of 26 rare species (defined as representation in fewer than six samples) (A), 21 intermediate species (defined as representation in more than six samples and less than 44 samples) (B), and 13 common species (defined as representation in more than 44 samples) (C) in 440 samples on the northern rocky intertidal of Helgoland. Note that the two curves in each graph (A–C) are almost identical.

131,636 squares using the simulation data set. Thus, the expected number of species in k samples from the whole northern intertidal of Helgoland will be

$$S(k) = S_{Total} - S_{Rare}(1 - p)^k,$$

where S_{Total} is the total number of species in the whole northern intertidal of Helgoland, S_{Rare} is the total number of rare species in the same area, and p is the average fraction of squares where a rare species exists. The approximative model rests on the assumption that the accumulation of new species
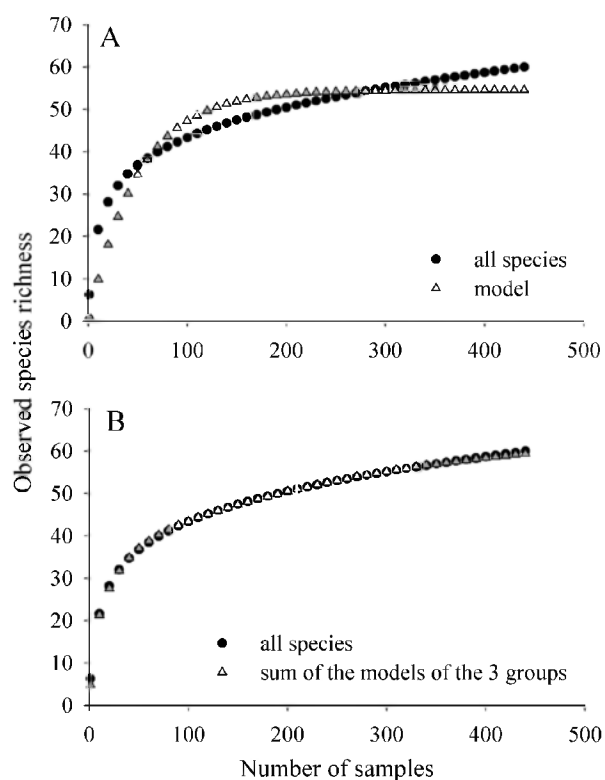
Fig. 2. The randomized empirical accumulation curve and the theoretical expectation of all 60 species observed in 440 samples on the northern rocky intertidal of Helgoland. (A), Note that the model is not able to fit the empirical curve. (B), This model is the sum of the models of the three groups (rare, intermediate, common); note that the perfect fit to the empirical curve is a consequence of the perfect fit of the separate group-models in Fig. 1A–C.
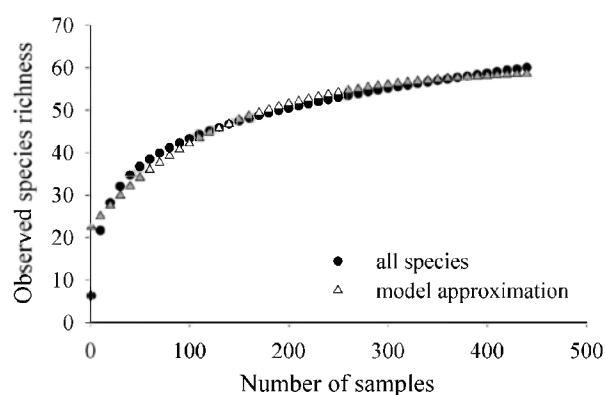

Fig. 3. The randomized empirical accumulation curve and the approximation to the theoretical expectation of all 60 species observed in 440 samples on the northern rocky intertidal of Helgoland. This model is defined as $S(k) = S_{Obs} - S_{Rare}(1 - p_{Rare})^k$ where S_{Obs} is the observed species number in all the samples and is set to 60. The two other parameters, S_{Rare} (number of rare species) and p_{Rare} (average probability of a rare species occurring in a sample), are regarded as unknown and estimated by least squares in comparison with the randomized empirical accumulation curve of all species.

Table 1. Bias of the Ugland TS estimator for the model data set.

p	Squares	Ugland TS estimate	Bias, %
0.00500	658	315	110
0.00400	527	300	100
0.00300	395	273	82
0.00200	263	228	52
0.00150	197	197	31
0.00120	158	175	17
0.00110	145	167	11
0.00100	132	158	5
0.00095	125	154	3
0.00090	118	149	0
0.00085	112	145	−4
0.00080	105	140	−7
0.00075	99	135	−10
0.00070	92	130	−13
0.00065	86	125	−16
0.00060	79	120	−20
0.00050	66	110	−27
0.00040	53	99	−34
0.00030	39	87	−42
0.00020	26	75	−50
0.00010	13	63	−58

Twenty-one alternative values for the parameter p, i.e., the average fraction of squares where a rare species exists and the corresponding number of squares of the whole area, i.e., of the possible 131,636 squares. The corresponding Ugland TS estimate and its bias [(estimate − 150)/150] is shown for each of the 21 alternative values.

is predominantly determined by the addition of rare species when passing 40 samples.

To study how the occurrence of rare species influences the performance of the Ugland TS estimator, we studied 21 alternative values for the parameter p (average fraction of squares that a rare species occupies) given in Table 1. The difference between the given expected number of species in the model data set (S_{Total} = 150) and the calculated number of species using linear extrapolation of the regression line through the 10 end points S(40), S(80), ..., S(440) gives the bias of the Ugland TS estimator.

Assessment

Estimator results and predictive accuracy—Richness estimates and accuracy for the 12 different methods are shown in Table 2. Only the Ugland TS estimator provided a highly accurate estimate. With a predicted richness of 141 (± 5.9 SD), it was the sole SRE yielding less than 10% error (the actual overestimate was 5.2%). In fact, the 95% confidence interval of the Ugland TS estimates is [129, 153], which incorporates the "true" species number 134. The next most accurate estimates, the MMF, Semilog, and Weibull curves, underestimated "true" species richness by 10%–16%. The remaining eight SREs were subject to between 31% and 76% prediction error.

Table 2. Predicted species richness and accuracy for the 12 species richness estimators (SREs) evaluated in this study.

SRE	Predicted richness	Absolute error	Standard deviation	Percentage error
Power	236	102	9.6	76.1
Semilog	116	18	7.5	13.4
Michaelis-Menten mean	61	73	24.0	54.5
Negative exponential	55	79	41.1	59.0
Weibull	113	21	2.6	15.7
MMF	120	14	2.7	10.4
Chao2	93	41	17.3	30.6
ICE	77	57	6.4	42.5
Jackknife1	74	60	3.4	44.8
Jackknife2	85	49	6.8	36.6
Bootstrap	66	68	3.0	50.7
Ugland TS	141	7	5.9	5.2

Calculations are based on data from a survey of sessile and slow-moving invertebrates in the intertidal of northern Helgoland (Reichert et al. 2008a). Accuracy is measured as prediction error from the "true" species richness of 134.

Influence of rare species on the bias of the Ugland TS estimator—The simulation of the Ugland TS method to the model accumulation curve $S(k) = S_{Total} - S_{Rare}(1 - p)^k$ for k = 40, 80, 120, 160, 200, 240, 290, 340, and 440 is given in Table 1. The expected total number of species in the model data set is 150, and the parameter p is the average probability of observing a given rare species in each of the 131,636 quadrats of 0.25 m^2 needed to cover the whole northern intertidal area of Helgoland. For each of the 21 alternative p values, the bias of the Ugland TS estimate has been calculated.

To explain how the bias for the TS estimator has been calculated, we provide an example with actual numerical values. Let us choose p = 0.0008, which corresponds to an average occupancy of pA = 0.0008 × 131,636 = 105 samples among the rare species. Hence, the randomized accumulation curve, corresponding to these parameters, will be $S(k) = 150 - 100 × (1 - 0.0008)^k$ = 150 – 100 × 0.9992^k. We may now calculate the expected number of species in the 10 subsets we use for the Helgoland data set, i.e., k = 40, 80, 120, 160, 200, 240, 290, 340, 390, 440.

For example, $S(40)$ = 150 – 100 × 0.9992^{40} = 53.2, $S(80)$ = 150 – 100 × 0.9992^{80} = 56.2, etc. In this way we get the following empirical species accumulation:

53.2 56.2 59.2 62.0 64.8 67.5 70.7 73.8 76.8 79.7

The corresponding Ugland TS estimate is 140, which is an underestimate of 10 species, as our model includes 150 species as the total species number. For this example, the corresponding bias is calculated as (140 – 150)/150 = –7%.

The Ugland TS method overestimates by more than 30% (50 species) if the rare species on average occupy more than 195 squares, i.e., >0.15% of all squares, and underestimates by more than 30% if the rare species on average occupy less than 55 squares, i.e., <0.04% of all squares. Between these values, there is a window ranging from an occupancy of 92–145 squares (P = 0.0007 to P = 0.00110) where the Ugland TS method produces a bias of ±13% or less.

Discussion

The performance and statistical properties of SREs have been addressed by several authors (e.g., Palmer 1990; Walther and Moore 2005; Hortal et al. 2006; *see* references therein and above for more comprehensive lists of works). However, comparative analyses often provide contradictory results (compare, e.g., Keating and Quinn 1998 and Chazdon et al. 1998), and there is no consensus on the best SRE for general use. We, as previous authors, have shown contradictory results of 12 SREs. Assessment of SRE performance strongly depends on how the "true" species richness (against which the estimations are compared) is obtained. Importantly, our estimations were compared to a complete species list serving as the "true" richness of the area studied. This is in direct contrast to the most common method for calculating the "true" richness, which is to fit the asymptote of a certain model to the randomized accumulation curve calculated from the samples themselves (e.g., Foggo et al. 2003; Lambshead and Boucher 2003; Canning-Clode et al. 2008). The justification of this approach usually relies on the R^2 values close to 1 obtained while fitting these curves; the assumption being that, if the model approximates the observed accumulation curve, its extrapolation must be quite close to the "true" species richness. However, despite a perfect model fit to a specific data set, extrapolation of the randomized accumulation curves will in general not provide information on the true species richness (Ugland et al. 2003). It is therefore difficult to reliably assess which estimator might perform best, (i.e., best represent the underlying assemblage and/or territory by its specific assumptions and statistical properties, alone). Although it is rarely available, it is a great advantage to assess the performance of SREs against a complete species list.

As mechanisms driving the performance of SREs itself, factors related to the particular sampling protocol used (e.g., sample size) and the spatial distribution of species and their abun-

dances (e.g., population dynamics) have been proposed (e.g., Magurran 2004; King and Porter 2005; Walther and Moore 2005). The effect of these factors on a given SRE should be understood before embarking on a study aiming to extrapolate total species richness in an area. However, this is difficult to achieve without an unrealistically detailed knowledge of the system. SREs with the most appropriate assumptions about the underlying species abundance distribution in the unsampled area will produce the most reliable estimates. Nonparametric estimators and asymptotic species-accumulation curves assume almost homogeneous species abundance distribution characterized by a majority of common species and a small number of rare species. Although some estimators may be relatively robust to moderate violations of the underlying species abundance distribution, larger departures are likely to result in increasingly unrealistic estimations (Ugland and Gray 2004, 2005). This poor performance is no exception, particularly in many marine assemblages which are extremely patchy at small spatial scales (e.g., Fraschetti et al. 2005) and are characterized by high numbers of rare species—much like the situation on the northern shore of Helgoland (*see* Reichert et al. 2008a). For example, Chapman and Underwood (2009) describe serious underestimates of true species richness in three different rocky shore marine assemblages. Inadequate performance of nonparametric estimators is also reported for terrestrial plant communities that are characterized by many rare species, i.e., species with a very low probability of encounter (e.g., Chiarucci et al. 2003).

Our approach to separately analyze the species-accumulation curves for rare, intermediate, and common species revealed that the curvature of the randomized empirical species-accumulation curve of all species is determined primarily by the occurrence of rare species. We have shown for the first time, using simulations of the Ugland TS estimator, how the underlying structure of a data set (here the proportion of rare species) can influence the outcome of a prediction. If the probability of encountering a rare species is relatively high, the accumulation of rare species will be too large and the linear extrapolation from the sampled area to the entire unsampled area will overestimate, or vice versa. In contrast, the linear extrapolation will be little biased if the probability of encountering rare species is neither high nor low. For the model data set examined here, the simulations of the Ugland TS method will give a reliable estimate if rare species occupy a fraction between 1.1/1000 and 0.7/1000 of the available area (i.e., the number of squares of size 0.25 m² occupied by a rare species is between 92 and 145).

In our empirical data set, most of the 26 rare species occurred in only one of the 440 samples. For such a singleton, this means an occupancy rate of 1/440 = 0.2% of all sampled squares. Because we do not know the true average number of the occupancy of a rare species over the total area of 131,636 squares, we simulated alternative values between 13 (i.e., 0.01%) and 658 (i.e., 0.5%) squares. In our simulations, the

optimal performance of the Ugland TS method occurred when a rare species on average occupied between 92 (0.07%) and 145 (0.11%) squares. It should be kept in mind, however, that the sampled area represents only 0.33% of the total area. It is therefore likely that the estimate of the occupancy of rare species is highly biased. For example, if the sampling intensity were twice as high as in the empirical sampling program (i.e., 880 instead of 440 samples), but with the same occurrence rate for a rare species (i.e., a species still occurs only once), the data set would have contained a rare species with 1/800 = 0.1%. It is well known from large sampling programs that a large number of species are observed in only one or two samples (Gaston and Blackburn 2000; Ellingsen and Gray 2002; King and Porter 2005; Ugland et al. 2007; Fontana et al. 2008), so the true average occupancy of rare species in the rocky intertidal of Helgoland is likely to be close to the optimal interval revealed by our simulations of the Ugland TS method.

Because we have clearly demonstrated that the probability of encountering a rare species strongly influences the performance of the Ugland TS estimator, the influence of rarity on the performance of other estimators should also be examined. For this purpose, it would have been possible to extend our simulations. In addition, further tests using other data sets would improve our understanding of the influence of rare species in different habitats. However, despite the potentially crucial role SREs play in biodiversity measurement, a sampling regime that detects all rare species is a formidable endeavor. These results also emphasize the difficulty of choosing an appropriate estimator a priori, without a detailed knowledge of the ecosystem concerned and, in turn, the structure of the data set used for the estimation. Nevertheless, our simulations provide a general framework for assessing the suitability of SREs for a given data set, and thus improve the selection process of the optimal method.

References

Arrhenius, O. 1921. Species and area. J. Ecol. 9:95-99. [doi:10.2307/2255763].

Bartsch, I., and I. Tittley. 2004. The rocky intertidal biotopes of Helgoland: present and past. Helg. Mar. Res. 58:289-302. [doi:10.1007/s10152-004-0194-2].

Borges, P. A. V., J. Hortal, R. Gabriel, and N. Homem. 2009. Would species richness estimators change the observed species area relationship? Acta Oecol. 35:149-156. [doi:10.1016/j.actao.2008.09.003].

Burnham, K. P., and W. S. Overton. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. Biometrika 65:927-936. [doi:10.1093/ biomet/65.3.625].

———, and ———. 1979. Robust estimation of population size when capture probabilities vary among animals. Ecology 60:927-936. [doi:10.2307/1936861].

Canning-Clode, J., N. Valdivia, M. Molis, J. Thomason, and M. Wahl. 2008. Estimation of regional richness in marine ben-

thic communities: Quantifying the error. Limnol. Oceanogr. Methods 6:580-590.

Chao, A. 1984. Non-parametric estimation of the number of classes in a population. Scand. J. Stat. 11:265-270.

———. 1987. Estimating the population size for capture-recapture data with unequal catchability. Biometrics 43:783-791. [doi:10.2307/2531532].

———, W.-H. Hwang, Y.-C. Chen, and C.-Y. Kuo. 2000. Estimating the number of shared species in two communities. Stat. Sinica 10:227-245.

Chapman, M. G., and A. J. Underwood. 2009. Evaluating accuracy and precision of species–area relationships for multiple estimators and different marine assemblages. Ecology 90:754–766. [doi:10.1890/08-0650.1].

Chazdon, R. L., R. K. Colwell, J. S. Denslow, and M. R. Guariguata. 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of northeastern Costa Rica, p. 285-309. *In* F. Dallmeier and J. A. Comiskey [eds.], Forest biodiversity research, monitoring and modelling: Conceptual background and old world case studies. Parthenon Publishing.

Chiarucci, A., N. J. Enright, G. L. W. Perry, B. P. Miller, and B. B. Lamont. 2003. Performance of nonparametric species richness estimators in a high diversity plant community. Divers. Distrib. 9:283–295. [doi:10.1046/j.1472-4642.2003.00027.x].

Clench, H. 1979. How to make regional fists of butterflies: Some thoughts. J. Lepid. Soc. 33:216-231.

Colwell, R. K. 2006. EstimateS: Statistical estimation of species richness and shared species from samples. Version 8. User's guide and application published at <http://viceroy.eeb.uconn.edu/estimates> (Accessed May 16, 2010).

———, and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. Philos. Trans. R. Soc. Lond B 345:101-118. [doi:10.1098/rstb.1994.0091].

Condit, R. G., and others. 2005. Tropical tree α-diversity: Results from a worldwide network of large plots. Biol. Skr. 55:565-582.

Costello, M. J., P. Bouchet, G. Boxshall, C. Emblow, and E. Vanden Berghe. 2004. European Register of Marine Species. <http://www.marbef.org/data/erms.php> (Accessed May 16, 2010).

Efron, B. 1979. Computers and the theory of statistics: Thinking the unthinkable. Siam Rev. 2:460-480. [doi:10.1137/1021092].

Ellingsen, K. E., and J. S. Gray. 2002. Spatial patterns of benthic diversity: Is there a latitudinal gradient along the Norwegian continental shelf? J. Anim. Ecol. 71:373-389. [doi:10.1046/j.1365-2656.2002.00606.x].

Foggo, A., M. J. Attrill, M. T. Frost, and A. A. Rowden. 2003. Estimating marine species richness: An evaluation of six extrapolative techniques. Mar. Ecol. Prog. Ser. 248:15-26. [doi:10.3354/meps248015].

Fontana, G., K. I. Ugland, J. S. Gray, and M. Abbiati. 2008.

Influence of rare species on beta diversity estimates in marine benthic assemblages. J. Exp. Mar. Biol. Ecol. 366:104-108. [doi:10.1016/j.jembe.2008.07.014].

Franke, H.-D., F. Buchholz, and K. H. Wiltshire. 2004. Ecological long-term research at Helgoland (German Bight, North Sea): Retrospect and prospect-an introduction. Helg. Mar. Res. 58:223-229. [doi:10.1007/s10152-004-0197-z].

Fraschetti, S., A. Terlizzi, and L. Benedetti-Cecchi. 2005. Patterns of distribution of marine assemblages from rocky shores: Evidence of relevant scales of variation. Mar. Ecol. Prog. Ser. 296:13-29. [doi:10.3354/meps296013].

Gaston, K. J., and T. M. Blackburn. 2000. Pattern and processes in macroecology. Blackwell Science. [doi:10.1002/9780470999592].

Gleason, H. A. 1922. On the relation between species and area. Ecology 3:158-162. [doi:10.2307/1929150].

Gotelli, N. J., and R. K. Colwell. 2001. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. Ecol. Lett. 4:379-391. [doi:10.1046/j.1461-0248.2001.00230.x].

Heincke, F. 1894. Beiträge zur Meeresfauna von Helgoland. II. Die Mollusken. Helgol. wiss. Meeresunters. 1:121-153.

Hoffmann, F. 1829. Einige Bemerkungen über die Vegetation und die Fauna von Helgoland. Verh. Gesellsch. Naturforsch. Freunde Berlin I: 228-260.

Hortal, J., P. Garcia-Pereira, and E. García-Barros. 2004. Butterfly species richness in mainland Portugal: Predictive models of geographic distribution patterns. Ecography 27:68-82. [doi:10.1111/j.0906-7590.2004.03635.x].

———, P. A. V. Borges, and C. Gaspar. 2006. Evaluating the performance of species richness estimators: Sensitivity to sample grain size. J. Anim. Ecol. 75:274-287. [doi:10.1111/j.1365-2656.2006.01048.x].

———, J. M. Lobo, and A. Jiménez-Valverde. 2007. Limitations of biodiversity databases: Case study on seed-plant diversity in Tenerife (Canary Islands). Conserv. Biol. 21:853-63. [doi:10.1111/j.1523-1739.2007.00686.x].

Hyams, D. (2005). A comprehensive curve fitting system for Windows: CurveExpert 1.3. <http://curveexpert.webhop.biz/> (Accessed May 16, 2010).

Jiménez-Valverde, A., S. Jiménez Mendoza, J. Martín Cano, and M. L. Munguira. 2006. Comparing relative model fit of several species accumulation functions to local *Papilionoidea* and *Hesperoidea* butterfly inventories of Mediterranean habitats. Biodiversity Conserv. 15:177-190. [doi:10.1007/s10531-004-5016-y].

Keating, K. A., and J. F. Quinn. 1998. Estimating species richness: The Michaelis-Menten model revised. Oikos 81:411-416. [doi:10.2307/3547060].

King, J. R., and S. D. Porter. 2005. Evaluation of sampling methods and species richness estimators for ants in upland ecosystems in Florida. Environ. Entomol. 34:1566-1578. [doi:10.1603/0046-225X-34.6.1566].

Lambshead, P. J. D., and G. Boucher. 2003. Marine nematode

deep-sea biodiversity: Hyperdiverse or hype? J. Biogeogr. 30:475-485.

Lee, S.-M., and A. Chao. 1994. Estimating population size via sample coverage for closed capture-recapture models. Biometrics 50:88-97. [doi:10.2307/2533199].

Magurran, A. E. 2004. Measuring biological diversity. Blackwell.

Melo, A. S., L. M. Bini, and S. M. Thomaz. 2007. Assessment of methods to estimate aquatic macrophyte species richness in extrapolated sample sizes. Aquat. Bot. 86:377-384. [doi:10.1016/j.aquabot.2007.01.005].

Michaelis, L., and M. L. Menten. 1913. Die Kinetik der Invertinwirkung. Biochem. Z. 49:333-369.

Morgan, P. H., L. P. Mercer, and N. W. Flodin. 1975. General model for nutritional responses of higher organisms. Proc. Natl. Acad. Sci. U. S. A. 72:4327-4331. [doi:10.1073/pnas.72.11.4327].

Palmer, M. W. 1990. The estimation of species richness by extrapolation. Ecology 71:1195-1198. [doi:10.2307/1937387].

Preston, F. W. 1948. The commonness and rarity of species. Ecology 29:254–283. [doi:10.2307/1930989].

Reichert, K., and F. Buchholz. 2006. Changes in the macrozoobenthos of the intertidal zone at Helgoland (German Bight, North Sea): A survey of 1984 repeated in 2002. Helg. Mar. Res. 60:213-223. [doi:10.1007/s10152-006-0037-4].

———, ———, I. Bartsch, T. Kersten, and L. Giménez. 2008a. Scale-dependent patterns of variability in species assemblages of the rocky intertidal of Helgoland (German Bight, North Sea). J. Mar. Biol. Ass. U. K. 88:1319-1329. [doi:10.1017/S0025315408001926].

———, ———, and L. Giménez. 2008b. Community composition of the rocky intertidal of Helgoland (German Bight, North Sea). Helg. Mar. Res. 62:357-366. [doi:10.1007/s10152-008-0123-x].

Smith, E. P., and G. van Belle. 1984. Nonparametric estimation of species richness. Biometrics 40:119-129. [doi:10.2307/2530750].

Soberón, J., and J. Llorente. 1993. The use of species accumulation functions for the prediction of species richness. Conserv. Biol. 7:480-488. [doi:10.1046/j.1523-1739.1993.07030480.x].

Storch, D., A. L. Izling, and K. J. Gaston. 2003. Geometry of the species-area relationship in central European birds: Testing the mechanism. J. Anim. Ecol. 72:509-519. [doi:10.1046/j.1365-2656.2003.00721.x].

Tjørve, E. 2003. Shapes and functions of species-area curves: A review of possible models. J. Biogeogr. 30:827–835. [doi:10.1046/j.1365-2699.2003.00877.x].

Ugland, K. I., and J. S. Gray. 1982. Lognormal distributions and the concept of community equilibrium. Oikos 39:171–178. [doi:10.2307/3544482].

———, J. S. Gray, and K. E. Ellingsen. 2003. The species accumulation curve and estimation of species richness. J. Anim. Ecol. 72:888-897. [doi:10.1046/j.1365-2656.2003.00748.x].

———, and J. S. Gray. 2004. Estimation of species richness: Analysis of the methods developed by Chao and Karakassis. Mar. Ecol. Prog. Ser. 284:1-8. [doi:10.3354/meps 284001].

———, ———, and P. J. D. Lambshead. 2005. Species accumulation curves analysed by a class of neutral models discovered by Arrhenius. Oikos 108:263-274. [doi:10.1111/j.0030-1299.2005.13532.x].

———, Lambshead, P. J. D., B. McGill, J. S. Gray, R. J. Ladle, and R. J. Whittaker. 2007. Modelling dimensionality in species abundance distributions: Description and evaluation of the Gamin model. Evol. Ecol. Res. 9:313-324.

Vandepitte, L., and others (2010). Data integration for European marine biodiversity research: Creating a database on benthos and plankton to study large-scale patterns and long-term changes. Hydrobiologia 644:1-13. [doi:10.1007/s10750-010-0108-z].

Walther, B. A., and J. L. Moore. 2005. The concepts of bias, precision, and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. Ecography 28:815-829. [doi:10.1111/j.2005.0906-7590.04112.x].

Weibull, W. 1951. A statistical distribution function of wide applicability. J. Appl. Math. 18:293-296.