Ecological size-frequency distributions: how to prevent and correct biases in spatial sampling

Assaf Zvuloni,1,2* Yael Artzy-Randrup,3 Lewi Stone,3 Robert van Woesik,4 and Yossi Loya1
1Department of Zoology, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel
2The H. Steinitz Marine Biology Laboratory, the Interuniversity Institute for Marine Sciences of Eilat, P.O.B. 469, Eilat, Israel
3Biomathematics Unit, Department of Zoology, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel
4Department of Biological Sciences, Florida Institute of Technology, 150 West University Boulevard, Melbourne, Florida, USA

Abstract

Size-frequency distributions (SFDs) have been used to assess the ecological status of different populations in a variety of ecosystems and recently have become widely used to examine reef corals. SFDs may reflect the response of time-varying influences of the environment, including the intensity and frequency of disturbances and the degree of environmental degradation. Here we elucidate the biases that can arise in the application of popular and traditional sampling methods (e.g. quadrat, belt-transect, and line-intercept). We show that these biases on the estimated SFD arise due to boundary effects of the sampling units. Incorrect evaluations of SFDs may lead to biased estimations of the ecological status of coral populations and may result in, among other things, erroneous nature reserve management policies. Our analysis is based on analytical calculations, simulations, and field observations. We have developed simple mathematical corrections, which provide unbiased estimations for previously collected data acquired by these widely used methods. In addition, we offer alternative sampling methods that do not suffer from these shortcomings. Eliminating these types of sampling errors will not only provide better assessments of the status of a given coral reef, but will also make way for more precise comparisons among coral reefs in different regions. Although we discuss the biases of SFDs in regard to reef coral populations, the work is equally relevant in other ecological contexts.

Introduction

Many areas of ecological research aim toward characterizing the size-frequency distribution (SFD) of populations to assess change across space and across time (e.g., Bak and Meesters 1998, 1999; Meesters et al. 2001). Size is often linked to physiological performance. Most contemporary corals, for example, are colonial organisms that form large superstructures by cloning. Although a coral’s size does not always equate to colony age (Hughes and Jackson 1980), size does reflect many life-history processes, such as maturation, fecundity, and survival (Loya 1976a, Hall and Hughes 1996, Meesters et al. 2001). Size may also reflect the response of corals to time-varying influences of the environment, including the intensity and frequency of disturbances and the degree of environmental degradation (Connell et al. 1997). SFDs are thus able to provide some insight into past environmental events and therefore may provide predictive capacity to infer population change (Bak and Meesters 1998). In addition, SFDs have proved important in many ecological applications ranging from conservation assessments (Bak and Meesters 1998, 1999; Meesters et al. 2001) to disease outbreaks (Kim and Harvell 2004). Numerous studies have used SFDs to assess the ecological status of ecosystems. Such studies usually assume a close relationship between the dynamics of the processes that generate the state and the size structure of the associated organisms. Specifically, it has become common to study the qualitative shape of SFDs (e.g., skewed to the left or to the right) when assessing the status of coral reefs (e.g., van Woesik and Done 1997; Bak and Meesters 1998, 1999; Meesters et al. 2001; Ben-Tzvi et al. 2004; Oigman-Pszczol and Creed 2004; Smith et al.

*E-mail: zvolonia@post.tau.ac.il

Acknowledgments

Research in part (AZ, RVW and YL) was supported by the World Bank and the Global Environmental Facility through the Coral Reef Targeted Research and Capacity Building for Management program, Coral Bleaching and Local Environmental Responses working group. YAR and LS are grateful for the support of the James S. McDonnell Foundation and the Horowitz Center for Complexity Science. We would like to thank Yoni Vortman, Ido Mizrachi, Noa Liscovitch and Yonni Shaked for helpful discussions, as well as Amatizia Genin, Immanuel Noy-Meir, Jonathan Belmaker, Victor China and Moshe Kilawi for reviewing the paper.
Therefore, incorrect evaluation of the SFD may not only lead to the inaccurate evaluation of the status of a population in a particular system, but may also incorrectly portray the environmental impact on that population.

A plethora of survey methods are used to examine sessile-benthic communities in a general effort to develop high-quality conservation and management strategies. Such techniques have often been compared (e.g., Weinberg 1981). Currently, in marine ecology, the line-intercept method (LIM), the quadrat method (QM), and the belt-transect method (BTM) are the most commonly used sampling techniques for estimating various spatial measures (e.g., population densities, size-frequency distributions, species diversity, spatial aggregations, percent cover) of sessile benthic communities. The LIM, introduced by Loya (1972) to coral reef ecology, records all organisms intercepted by a line (e.g., measuring tape). In the QM and BTM, rectangles are placed in the field site, defining bounded areas; all organisms falling within the rectangles define the sample. Some of the measures estimated by these three methods, such as percent cover, are calculated according to the relative coverage of the organisms in the defined area. Other measures, however, such as population densities, SFDs, and species diversity, are count-based measures (measures that relate to the number of individuals in an area). Although these three methods (QM, BTM, and LIM) have become routine tools for estimating various demographic statistics, we show that, in many situations, the count-based measures are prone to bias. For example, estimating SFD of corals by using a reasonable sized quadrat may lead in many situations to a large bias, of even >50%, in over- or underestimation of the frequency of the extreme size groups.

Here we describe the biases that may arise when using each of these three time-honored sampling techniques for studying spatial distributions. Even though the principles behind the biases in all the count-based measures are common, our work is motivated from research assessing the SFDs of coral populations. We offer simple mathematical corrections for an unbiased estimation of SFDs, which can easily be applied to previously collected data, and we present alternative unbiased sampling schemes that do not suffer from these shortcomings. Although we use coral colonies as an example throughout the text, our basic findings may be equally relevant to any other sessile organisms that are not necessarily marine.

Consider the procedure involved when sampling with the BTM and QM. By randomly placing a sampling unit (e.g., a rectangle) on the seabed, any coral colony intersected by an edge of this unit is split into two parts—one part lying within the sampling unit and the other lying externally (see Figure 1). A situation like this does not raise any sampling problems when assessing percent cover, since only the coverage of a coral colony lying within the sampling unit is taken into account. However, when attempting to assess count-based measures, such as SFD, it is not always clear whether to count such an individual or not. In field studies, authors do not often specify the manner in which these borderline cases were treated. Exceptions to this are Nugues and Roberts (2003) and Dikou and van Woesik (2006), who described their coral sampling methods in detail, specifically stating how such borderline corals were treated in their studies. In many cases, however, a sampler will decide in advance to either ignore all such borderline individuals in the general count or to include them throughout the entire field study (O. Ben-Tzvi, S. Oigman-Pszczol, and R. van Woesik, personal communications). As a notation throughout the text, we denote each of these approaches as type I and type II:

- type I: any coral that is covered by an edge is excluded from the sample;
- type II: any coral that is covered by an edge is included in the sample.

The decision of how to deal with such borderline cases may lead to biases in the estimations of count-based measures. A straightforward example for such a case is when attempting to estimate population density, a basic measure that can be strongly biased if the counting scheme is not applied with care. When sampling population densities by the BTM or QM, deciding to consistently count or ignore all the borderline cases will clearly lead to an over- or underestimation (respectively) of the true density. Similarly, when assessing SFD, incorrect treatment of borderline cases can lead to biased estimations, but the reason here is less apparent. The probability that a sampling unit’s edge intercepts a
coral colony correlates to the colony’s size; larger colonies have a higher probability of being covered by an edge than smaller ones. Thus, deciding to consistently ignore all the corals intercepted by an edge leads to a positive bias toward the smaller colonies, and a negative bias toward the larger ones. In this scenario, the estimated frequency of the smaller corals will be higher than the true frequency (Figure 2). On the other hand, deciding to consistently count all borderline corals will result in the opposite bias; the estimated SFD will incorrectly show higher frequencies for the larger corals and lower frequencies for the smaller ones.

A similar bias arises under the LIM, when lines are randomly placed in an area; the probability of a line intercepting a coral colony will again correlate to the colony’s size. Using the LIM would always lead to a positive bias toward counting the larger colonies, and a negative bias toward counting the smaller colonies. Although this method was originally employed to assess only measures relating to relative coverage (measures that would not be influenced by this type of bias), with time the LIM was also applied for the estimation of different count-based measures, such as mean colony sizes (Marsh et al. 1984, Dustan and Halas 1987), SFDs (Loya 1976b, Bruckner, Silverman and Genin 2005, Shaked and Genin 2006), and species diversity (Loya 1972, Loya 1976b, Silverman and Genin 2005, Shaked and Genin 2006). As an additional example, Dornelas et al. (2006) tested the unified neutral theory of biodiversity (Hubbell 2001) on Indo-Pacific coral communities using the LIM. In this case, because species diversity relies on individual counts, using the LIM leads to a positive bias toward counting the larger colonies as well as the larger coral species.

Materials and Procedures

**Estimation and correction of biases**—To address the origins of these biases in a more formal way, let us assume for the sake of simplicity that coral colonies have a circular boundary, a reasonable assumption for these types of populations (see Marsh et al. 1984). Under this assumption, consider a plane consisting of corals of different sizes. When randomly throwing a dart onto this plane, the probability \( P \) of hitting a specific coral of size \( S \) is \( P = S/A \), where \( A \) is the total area of the plane. Obviously bigger corals covering a larger area (\( S \)) have a greater chance of being hit by the dart, with a relative magnitude that is directly proportional to their different sizes.

The same concept holds when randomly placing an imaginary line on a coral reef. In this case, the probability of a line covering a coral colony is proportional to the colony’s diameter. Such imaginary lines can be seen to represent the lines used in the LIM or the edges of rectangles used in the QM and BTM. Taking a closer look at the LIM, the estimated frequency of corals of size-class \( i \) fulfills the following relation:

\[
f_{i,obs} \propto D_i f_i, \tag{1}
\]

where \( f_i \) is the true frequency in the sampled area, \( f_{i,obs} \) is the observed frequency used as an estimate, and \( D_i \) is the average diameter of corals from size class \( i \).

As a simple demonstration of how this relationship can distort the SFD, consider the following example based on the LIM. Suppose a coral population is divided into \( n \) uniformly distributed size classes, i.e., with approximately the same number of corals in each size class. The true normalized frequency of each size class is thus \( 1/n \). As a simple example, suppose there are \( n = 11 \) size classes and set the diameter of corals in the \( i \)th size class to be \( D_i = i \) units. The true frequency of each size class is thus \( f_i = 1/11 = 0.091 = 9.1\% \). However, sampling such a population via the LIM will not reproduce this uniform frequency distribution. Instead, as the example in Figure 3 shows (orange line), far fewer of the smallest corals (1.52%) will be observed than the true 9.1% (black line), while there are many more of the largest corals observed (16.67%), a bias that is predicted by Eq. 1. The bias can be quantified through the following simple calculation. According to Eq. 1, the normalized observed coral frequencies of the \( i \)th size class should be:

\[
f_{i,obs} = \frac{D_i (1/n)}{\left[ D_1 (1/n) + D_2 (1/n) + \ldots + D_n (1/n) \right]} \tag{2}
\]

The ratio between the observed and the true frequency for the \( i \)th size class is then

\[
\frac{f_{i,obs}}{f_i} = \frac{2i}{n(n+1)} = \frac{2i}{n+1} \tag{3}
\]

Thus the frequency of the smallest class \( (i = 1) \) is observed as being only 2/12 (or 16.67%) of the true frequency, whereas the largest class \( (i = 11) \) is observed as being 22/12 of the true frequency (i.e., 83.33% greater). From Eq. 3, one sees that:

---

Fig. 2. Sampling units are represented as black rectangles. In A and B, the type I approach is used, and the corals (circles) counted are those marked in blue. The area that is effectively being sampled (shaded in light blue) is different for each size class, depending on the coral’s radius. The larger corals in A are effectively being sampled from a smaller area than the smaller corals in B. In C and D, the type II approach is used, and the counted corals are those marked in green. Here, the larger corals in C are effectively being sampled from a larger area (the shaded light green area) than the smaller corals in D.
The largest \((n)\)th size class is overrepresented appreciably by \(\frac{1}{2(n+1)}\), which is the true frequency, and this error approaches 100% as the number of size classes \(n\) increases; the smallest size class is substantially underrepresented by \(\frac{1}{2(n+1)}\), which is the true frequency, and this error approaches 100% as the number of size classes \(n\) increases.

Despite these large errors, it is not difficult to correct the observed frequencies of each size class and retrieve an unbiased estimate of the true size-frequency distribution (SFD). The relationship given by Eq. 1 makes clear that the normalized unbiased estimate of each size class's frequency for the line-intercept method (LIM) is

\[
f_i = \frac{f_{(obs)} / D_i}{\sum_{j=1}^{n} f_{(obs)} / D_j}
\]

(4),

where \(i\) is an index of each size class and \(n\) is the total number of size classes. The denominator of Eq. 4 is a normalization factor that ensures \(\sum_i f_i = 1\). Equivalently, we can rewrite Eq. 4 in the following form:

\[
f_i = \frac{\alpha_i f_{(obs)}}{\sum_{j=1}^{n} \alpha_j f_{(obs)}}
\]

(5),

where

\[
\alpha_i = \frac{1}{D_i}
\]

(6)

acts as a correction term inversely proportional to the expected bias of each size class.

In cases of sampling an area by randomly placing rectangles on a reef (i.e., the QM and BTM), defining a correction rule is not as straightforward, but still relies on the same principle. We now return to the two sampling approaches discussed earlier; the type I and type II approaches, where corals covered by an edge are either consistently ignored or consistently counted throughout a field study. Cases in which a combination of these two is used will be discussed latter on, and we now focus our attention only to the implementation of the pure type I and the pure type II approach.

When sampling using only the type I approach, we define a correction term \(\alpha_p\) for each size-class \(i\), as

\[
\alpha_p = \frac{R_w}{(R_w - D_i) \cdot (R_l - D_i)}
\]

(7),

where \(R_w\) and \(R_l\) are a rectangle's width and length. The numerator is thus the area of the rectangle and the denominator is the “effectively” sampled area (i.e., the area where a coral’s center must be located for it to be included in the sample; see Figure 2A and B), giving a ratio that is inversely pro-

![Fig. 3. Uniform (A–C) and nonuniform (D–F) size-frequency distributions plotted as black lines for 11 size classes given by Eqs. 10 and 11. In A and D, expected biased size-frequency distribution (SFD) calculated for the line-intercept method (LIM; orange lines). In A, the largest size class is overrepresented by 83.3% and the smallest size-class is underrepresented by 83.3% (see Eq. 3). B and E show biased SFDs found when sampling randomly with 200 rectangles according to the type I (blue lines) and type II (green lines) approaches. The green (blue) curve is skewed to the left (right). Blue and green dotted lines represent the theoretical predicted biases of the methods for the type I and type II approaches, respectively. In C and F, the blue and green curves are the nonbiased SFDs found in the simulation for the observed type I (blue) and type II (green) approaches, respectively, after correction according to Eqs. 5, 7, and 8.](image-url)
portional to the expected bias. An unbiased estimate of the true SFD in the area would therefore be given by inserting this correction term given by Eq. 7 into Eq. 5.

A type II correction term can be calculated in a similar way:

$$\alpha = \frac{R_i + R_o}{(R_i + D) \cdot (R_o + D) - [D_i^2 - \pi \cdot (D_i/2)^2]}$$ \hspace{1cm} (8).$$

Here again the numerator is the rectangle’s area and the denominator represents the effectively sampled area, which in this case is defined as being a rectangle with rounded corners (see Figure 2C and D). An unbiased estimate of the true SFD when sampling according to the type II approach is obtained by inserting Eq. 8 in Eq. 5.

In contrast to the LIM, the correction terms for the type I and type II approaches, when using the QM or the BTM, depend also on the size of the sampling unit (the rectangle width and length), and not only on the diameter of each size class.

In practice, calculating the corrected SFD from Eqs. 5–8 is straightforward. In the supplementary materials we provide an Excel spreadsheet that corrects any sampled SFD to its unbiased estimate for the LIM, the QM, and the BTM (0144a1.xls).

**Alternative sampling methods**

We now discuss alternative sampling schemes, which have been partly implemented in some field studies. Nugues and Roberts (2003) and Dikou and van Woesik (2006) were conscious of the potential biases that emerge from the boundary effect of their sampling unit. To avoid over- or underestimating the number of colonies, they used two alternative counting schemes based on a combination of the type I and type II approaches. Note that the goal of these authors was to estimate density and mean colony size, and they did not try to estimate SFDs, which is the main measure of interest in this paper. We will give a summary of each approach, describe its shortcomings, and offer an alternative unbiased sampling scheme, the “center rule” scheme.

Dikou and van Woesik (2006) take the type I approach for two parallel edges of a quadrat and the type II approach for the other two parallel edges (as recommended by Andrew and Mapstone 1987). Thus, ca. 50% of the corals being covered by an edge are counted and the other ca. 50% are discarded. Intuitively, it would seem that this approach leads to unbiased estimates, but this is not correct. For a coral of a given size class $i$, let us define a rectangle’s “inner buffer zone” as the area within the rectangle where this coral’s center must be located for it to be covered by an edge. This is the white area lying within the rectangles shown in Figure 2A and B, and referred to as $BZ_{i\text{in}}$. Similarly, we define the outer buffer zone for a coral of size class $i$ as the area outside the rectangle where this coral’s center must be located for it to be covered by an edge. This is the shaded light green area lying outside the rectangles in Figure 2C and D and referred to as $BZ_{i\text{out}}$. For the approach used by Dikou and van Woesik to be unbiased, the following statement must hold: given a coral colony covered by an edge, the probability of its center lying inside the rectangle is equal to the probability of its center lying outside the rectangle.

However, this assumption is incorrect. Because a rectangle’s inner buffer zone is smaller than the matching outer buffer zone, the probabilities are unequal, such that the centers of corals covered by an edge are more likely to be located outside the sampling unit than inside it. This counting scheme, therefore, tends to overestimate population density.

In reference to SFD, the ratio of the outer buffer zone $(BZ_{i\text{out}})$ to the inner buffer zone $(BZ_{i\text{in}})$ depends on the diameter of each size class $i$ ($D_i$) and fulfills the following relation:

$$\frac{BZ_{i\text{out}}}{BZ_{i\text{in}}} = \frac{R_i + R_o + \pi \cdot D_i/4}{R_i + R_o - D_i}$$ \hspace{1cm} (9).$$

This ratio, theoretically, is equal to 1 at its minimum (when $D_i = 0$) and grows, in practice, as a function of coral colony size. For example, $BZ_{i\text{out}}/BZ_{i\text{in}} = 1.6$ for a 1 by 1 m quadrat when $D_i = 0.5$ m. As such, using this scheme will lead to a positive bias toward large colonies in the estimated SFD (i.e., left-skewed).

The second approach used by Nugues and Roberts (2003) is based on counting only colonies with at least 50% of their surface area lying within a rectangle. In this scheme most of the bias arising from using either the type I or the type II approach is overcome, but not entirely. For example, if a coral’s center lies within a rectangle but is positioned very close to the rectangle’s corner, it would not be counted because more than 50% of its surface is outside the rectangle, as demonstrated in Figure 1 by the coral marked a. This type of scenario becomes significant when estimating the frequency of large colonies, which would be discriminated against.

We offer a solution to these problems in what we call the center rule scheme. Here only corals with centers lying within the sampling unit are counted, and all other corals are ignored (see corals with black centers in Figure 1). The advantage of this scheme is that the size of a coral does not play any role in the sampling probability, making this method nonbiased. Indeed, in this scheme, the problem of corals lying close to a corner as discussed earlier is overcome.

It is important to mention, however, that the alternative schemes mentioned above are not applicable for correcting biases in previously collected data. In order to use these schemes, it would be essential to have information about the exact location of each sampled coral, information which is usually not collected. As we demonstrate latter on in the second real data example, to correct biases of past studies it is usually possible and simple to implement the type I or type II correction terms as defined by Eqs. 6-8.

**Assessment**

**Simulation examples**—To demonstrate the outcomes of the different methods together with our suggested corrections, we simulated a hypothetical population of corals composed of 2000 individuals belonging to 11 different size classes; the corals’ radii were taken to range from 1 to 11 units of length for each size class, with a difference of 1 unit between them.
We tested two different frequency distributions of corals in each size class, the first taken to be uniform, set as

\[ \{f_1, f_2, \ldots, f_n\} = \left\{ \frac{1}{11}, \frac{1}{11}, \frac{1}{11} \right\} \]  

for each size class, as plotted in Figure 3A–C (black lines). In the simulation, the corals were randomly scattered over a plane of 1000 by 1000 units, with no individuals overlapping. This area was sampled by randomly placing 200 rectangles of 30 by 70 units, and counting the number of individuals of each size class that were purely within a rectangle and the number of individuals that were covered by an edge. From the results of this simulation, we calculated the estimated SFDs using the standard type I and type II approaches, without the above corrections. The observed distributions found with these approaches were found to be significantly biased in both cases \((P_v < 0.05; \text{Kolmogorov-Smirnov two-sample test})\), as can be seen by the blue and green solid lines plotted in Figure 3B. In fact, these results corresponded closely to the predicted bias, which can be calculated by dividing the true frequency distribution of the corals (Eq. 10) by the correction terms given in Eqs. 7 and 8 and normalizing (see the blue and green distribution of the corals (Eq. 10) by the correction terms). Application of the correction terms, given in Eqs. 7 and 8, retrieved a good estimate for the true frequency distribution, \(P_v < 0.05\), Kolmogorov-Smirnov test) from the true distributions. When testing the combined type I and type II approach used by Dikou and van Woesik, we found that 932 of the 1000 estimated SFDs were significantly different \((P_v < 0.05, \text{Kolmogorov-Smirnov test})\) from the true distributions, giving size-frequency histograms that were skewed to the left. To summarize, it was observed that the bias was so strong in each of these approaches that the null hypothesis, which suggested that the sampled data had the same attributes as the original SFD, was in almost all cases rejected.

**Real data examples**

Based on two field data examples, we examined the extent of the impact these corrections might have on the results and their eventual interpretation. In the first example, we sampled massive corals in the northern tip of the Gulf of Aqaba (Israel, Red Sea). Here, 60 quadrats of 50 by 50 cm were randomly placed along three parallel lines of 40 m each, at a depth of 1.5 m in front of the Interuniversity Institute (IUI) in Eilat. Numbers and sizes of massive corals (which are very abundant in this area) were quantified in each quadrat, and the occurrence or not of each coral falling under an edge was also recorded. The SFD of the recorded corals was evaluated by using the type I and type II sampling approaches, and mathematical corrections were carried out for both of the approaches. In addition, sampling was also done using the alternative center rule scheme.

The total number of corals counted was 602, with 235 found purely within the quadrats (not covered by an edge)
and 367 covered by an edge. No coral had a diameter exceeding the quadrat size. Similar to the simulation results above, we found different SFD estimates for each of the approaches (see Figure 4). The type I approach showed a significantly higher abundance of smaller colonies (right skewed) in comparison with the type II approach ($P << 0.01$, Kolmogorov-Smirnov test). However, by using the two suggested mathematical corrections defined by Eqs. 7 and 8, we retrieved size-frequency histograms that were almost identical (see dashed curves in Figure 4). In addition, the SFD was also estimated by the center rule scheme and was found to be similar to both of the corrected (unbiased) type I and type II approaches (see orange curve in Figure 4).

Our second example is based on data collected by the Israel National Monitoring Program in the Northern Gulf of Aqaba (Eilat). This monitoring program was established as part of a framework plan of the Israeli Ministry of Environment to monitor ecological and environmental changes to the reef and identify possible causes of such changes. The Northern Gulf of Aqaba has undergone rapid deterioration over recent years (Loya 2007), and there has been a need to define a general management policy. Lack of data on the physical, chemical, and biological processes in this area has held up decisions regarding appropriate policies and regulations. As part of the wide monitoring project carried out by the program (Silverman and Genin 2005, Shaked and Genin 2006), one of the study’s objectives was to estimate SFDs of coral populations on the reef using the LIM. In this study, four size classes were defined according to diameter: small (<5 cm), medium (5–15 cm), large (15–30 cm), and huge (>30 cm). As an example, Figure 5 shows that the observed SFD (black bars) of branching corals found at a depth of 10 m in Katza (Eilat) is skewed to the left, with a dominance of large colonies.

We applied the bias correction (Eq. 4) with the diameters of the first three size classes set as the mean diameter of that class, whereas the diameter for the huge size class was taken at its lowest bound (30 cm). Thus the corrected data presented here are in their least extreme form. After using the correction term (Eq. 4), we found that the conclusions change radically and in fact the opposite tendency exists (gray bars); the SFD is strongly skewed to the right and the population is mostly dominated by the smaller colonies, and likely to be a much younger population. This example demonstrates how significant these corrections can be, and how dramatically they can change the interpretation of results, which might have crucial importance in determining nature reserve management policies.

The bias problems identified in this study and their corrections are generally relevant to sampling by the popular methods, especially in marine ecology where the sampler is practically limited with the size of the sampling unit. However, in some generally unusual cases it may turn out that the corrections are not quantitatively large. For example, when the sampled coral colonies are very small relative to the sampling unit’s area, the biases related to the boundary effect will not be very significant.

**Discussion**

The above results demonstrate how widely used sampling methods can lead to significantly biased evaluations of count-based measures (SFD in this case) if not applied with care. In a keynote paper, Bak and Meesters (1999) argue that the status of a coral reef can be determined by the direction in which its SFD is skewed. Hence, it is critical to correct for biased skewness to facilitate accurate interpretations. As we showed in the example from the Israel National Monitoring Program in the Northern Gulf of Aqaba, the correction can radically change ecological interpretations. The observed site appeared to be depauperate in terms of its younger corals; according to the observed SFD the younger corals comprised ca. 12% of the full community. However, the corrected SFD showed an entirely different picture of thriving young corals that were nearly 50% of the entire community.

In general, the type I approach leads to a greater bias than the type-II approach. This is because the ratio between the sampling unit and its effective area (see Figure 2) is more extreme in the type I approach than in the type II approach (see also examples in Figure 3B and E). However, the biases in these evaluations can be removed by using the suggested mathematical corrections. The counting method used by Dikou and van Woesik (2006) provides a better estimation of the SFD in comparison to strictly using an uncorrected type I or type II approach, but this method still suffers from significant overestimation of the larger coral colonies. The offered alternative counting method (center rule), as shown in the simulations and in the real data example, overcomes these shortcomings and provides an unbiased SFD that is almost identical to the actual SFD.

![Figure 5](image-url)
Each of the unbiased sampling methods introduced here has its advantages and disadvantages depending on the particular circumstances and ecological context. For example, an approach that has become popular in coral reef monitoring over recent years is the use of photo-quadrats; that is, quadrats whose edges are defined by the frame of the photo itself or a PVC frame placed on the reef (e.g. Porter and Meier 1992, Kenyon et al. 2006). The corrected type II approach and the suggested center rule sampling methods require the inclusion of corals covered by an edge, which can be difficult to measure using photo-quadrats because the full size of some coral colonies is obscured and outside the range of the photos. Therefore, information on the size of those corals covered by an edge is ignored. In this case only the corrected type I approach would be suitable. However, ignoring corals covered by an edge leads to loss of information, making the type I approach generally weaker than the type II approach. This weakness becomes apparent in situations where there are colonies that are large relative to the frame of reference (i.e., the quadrat), a situation that occurs often in reefs such as in the Western Indian Ocean. In such cases, these large colonies have a smaller effective sampling area, and thus it is possible that although they have been observed, they will not be counted at all. This means that the method of photo-quadrats combined with the corrected type I approach is best for reefs with coral colonies that are small relative to the size of the sampling units.

Comments and Recommendations

As a working recommendation, we suggest the following selection criteria. In general, the corrected type II approach is preferable to the corrected type I approach. This is because the larger effective sampling area (see Figure 2) allows more information to be gathered (e.g., in the first real data example the type I approach recorded only 39% of the corals that type II recorded). Both the corrected type I and type II approaches are preferred over the center rule approach, because of ease of implementation. In practice, determining whether a coral is covered by a quadrant edge is straightforward as opposed to deciding whether its center is located within the quadrant, a decision that can be greatly influenced by a sampler’s subjectivity.

It is important to emphasize that no single optimal method exists, and that the sampling method of each study should be decided according to the characteristics of the sampled area. Depending on the nature of the site, it is possible to implement combinations of these approaches as well. A discussion about such combinations goes beyond the scope of the present paper and will be presented as part of future work.

This study focuses on the problem of borders with respect to SFD evaluations. However, the same problems arise in the spatial sampling of other count-based measures, including population densities, species richness and diversity, and spatial aggregations. For example, large species have a higher probability of being counted when estimating biodiversity or species richness using the LIM. In this case, correction should also be implemented. Similar to the Israeli Monitoring Program example from the Red Sea, the three sampling methods discussed above are very much in use in other monitoring programs worldwide. These sampling methods thus have crucial importance in determining nature reserve management policies. The corrections presented here can be used not only to prevent the introduction of biases but also to correct the findings of previous studies. Eliminating these types of sampling errors provides a firmer basis for ecosystem assessments and facilitates a more precise comparison between similar systems (e.g., coral reefs) with different community structures.

References


