Estimating ecological count-based measures from the point-intercept method

Assaf Zvuloni1, Jonathan Belmaker2,*

1Israel Nature and Parks Authority, PO Box 667, Eilat 88000, Israel
2Department of Zoology and the Steinhardt Museum of Natural History, Tel Aviv University, Ramat Aviv, Tel Aviv, Israel

ABSTRACT: The point-intercept sampling method (PIM) is an efficient, applicable, and common technique for collecting ecological data. With the rise of digital data, its use is considerably increasing for the analysis of images. However, to date the PIM has been solely used to estimate parameters related to coverage. The limitations of the PIM originate from being a plotless technique (i.e. the sampling unit does not define an area), and because it is prone to substantial size-related biases. Based on geometrical considerations, we introduce a simple approach, provided as user-friendly Excel spreadsheets and R functions, to overcome these limitations, providing the sizes of individuals sampled at the sampling points are also recorded. We demonstrate that our approach enables the user to derive unbiased estimations of important ecological measures that could not previously be estimated by the PIM (e.g. population density, size-frequency distribution, average size, and species diversity), and that the improved PIM is even more efficient than conventional plot-based techniques (e.g. the quadrat method).

KEY WORDS: Point-intercept method · Bias · Size-frequency distribution · Spatial sampling · Plotless sampling technique · Stratified sampling

INTRODUCTION

With the worldwide deterioration of diverse terrestrial and marine ecosystems, quantifying ecological parameters related to species demography, community structure, and biodiversity has become a major research priority (IPBES 2013). Such studies rely on accurate ecological information in order to assess and respond to ecosystem changes (Kremen et al. 1994). Augmenting classical sampling methods, such as quadrats or linear transects, various new sampling techniques, including photographic and video methods, have been developed in the last few decades (see reviews by English et al. 1997, Hill & Wilkinson 2004). All visual sampling, however, can be classified into 1 of 2 categories: (1) plot-based techniques, in which the sampling units define bounded areas (e.g. quadrats or belt transects); and (2) plotless techniques, in which the sampling units are lines or points (e.g. linear transect, linear-point intercept, and quadrant-point intercept; see reviews by Hill & Wilkinson 2004). Here, the sampling units do not define areas, but transects or locations across the studied habitat, and all the studied organisms intercepting the lines or points define the sample. Plotless techniques have been recommended due to their ease of deployment in the field (Loya 1978), and are increasingly used in analyses of photographic and video data (e.g. the Coral Point Count with Excel extensions software; Kohler & Gill 2006). Previous comparisons have indicated that the plot-based and plotless techniques provide similar estimates of parameters related to coverage (Bouchon 1981, Dodge et al. 1982).

The point-intercept sampling method (PIM) is an efficient, applicable, and therefore commonly used technique for collecting ecological data (Levy & Madden 1933, Hill & Wilkinson 2004, Kohler & Gill 2006). A specified number of points is distributed, randomly or evenly, over the sampling area, and the features underlying the points are categorized. The cover of each category can then be calculated as the percentage of points overlying it. This technique is
commonly used both in terrestrial and aquatic systems, such as forestry vegetation (e.g. Penttinen & Stoyan 2000), tree canopy cover (e.g. Thomas & Winner 2000), and marine benthic communities (e.g. Carleton & Done 1995). The use of the PIM is increasing with the rising prevalence of digital image analyses and the development of programs that help in extracting ecological measures from such images (Kohler & Gill 2006, Beijbom et al. 2012).

The PIM is a plotless technique, and, as such, the sampled area is not defined by the sampling unit, so that measures relating to density cannot be estimated. Moreover, this method is prone to substantial size-related biases, as the probability of a point intercepting an organism is proportional to its size (discussed in more detail in the 'Methods' section). These inherent limitations and biases of the PIM prevent its use in assessing any ecological count-based measures (i.e. measures which relate to the number of individuals in an area), such as population density, size-frequency distribution, average size, and species diversity. Therefore, to date, the PIM has been restricted to estimating only those parameters related to coverage.

In this study, we describe a simple and user-friendly approach for obtaining unbiased estimations of any count-based measures from PIM data. In practice, we modify the commonly used PIM in order to derive useful ecological parameters of species abundances while preserving the advantages of the method, viz. simplicity, speed, and efficiency. In addition to coverage, this method allows us to obtain size-frequency distributions (SFDs), mean size, and density of the sampled populations and measures related to species composition. SFD, for example, is an important and useful measure as it reflects environmental influences, including the intensity and frequency of disturbances and the degree of environmental degradation (Connell et al. 1997). We use coral colonies as an example throughout the text, although our method will be equally applicable to any sessile organism. We provide Excel spreadsheets and R functions for applying this method in Supplement 1 at www.int-res.com/articles/suppl/m556p123_supp/.

METHODS

Obtaining density, SFD, and mean size

Consider a plane consisting of organisms of different sizes. When randomly throwing a dart onto this plane, the probability $P$ of hitting a specific organism of size $S$ is $P = S/A$, where $A$ is the total area of the plane. Obviously, big organisms covering a large area ($S$) have a greater chance of being hit by the dart than smaller ones, with a relative magnitude that is directly proportional to their sizes. The same rationale holds true when randomly placing a sampling point on a coral reef, for example. The probability of the point intercepting a coral colony is proportional to the coral's size (i.e. the projected area of the coral), and, therefore, the PIM leads to a positive bias toward detecting large corals. It is important to note that the probability of a sampling point intercepting an object is proportional to the object’s area, which is independent of its shape. Hence, an unevenly shaped object will have the same probability of intercepting a sampling point as a round object with a similar area.

If we had a single sampling point, we could take into account the greater probability of sampling a large colony and apply a simple geometric correction to eliminate this bias. For most applications, however, a single point is insufficient for sampling ecological systems, so the sampler places an array of points over the studied area (e.g. repetitive linear, or quadrat-point, intercept; see review by Hill & Wilkinson 2004). In this case, the probability of detecting a coral colony might be much smaller than the sum of the probabilities of all individual points in the array due to potential dependency among the points (this is explained in more detail below). This has thus far hampered the use of PIM for obtaining count-based measures.

We offer an original way to look at the sampling probability of the PIM and solve the size-related biases associated with this method, by taking into account the ‘effectively’ sampled area (ESA) of the sampling unit, i.e. the area where a coral’s center must be located in order for it to be included in the sample (see Zvuloni et al. 2008). Although a point has no size or shape, in practice, when an area is randomly sampled by an array of points, each sampling point has an ESA. As a demonstration, consider a coral $i$ with radius $r_i$ and a sampling point, which is randomly placed at a distance $d_i$ from the center of this coral. Under the assumption that corals have a circular boundary, this coral will be sampled (i.e. intercepted) by the sampling point only if $d_i < r_i$. Therefore, $ESA_i$, the ESA of a single point for sampling a coral $i$ with radius $r_i$ is $\pi \times r_i^2$, which is equal to the projected area of the coral itself (Fig. 1A) and is larger for large corals than for smaller ones. Thus, the sampling effort, which can be represented by the ESA, is dependent on the size of the sampled individ-
Zvuloni & Belmaker: Improving the point-intercept method

When calculating the total ESA of an array of points (Fig. 1B,C), overlap among ESAs of neighboring points must be taken into account. Thus, the probability of a coral being sampled is determined not only by the coral’s size and the number of sampling points, but also by the spatial organization of the points. On one extreme, if the points are very close to each other (i.e. assume that the distances among the points approaches 0), the probability of detecting a coral colony by the array will be almost equal to that of a single point. The other extreme is when the points within the array are very far from each other (i.e. relative to the size of the sampled individuals). Here, the sampling probability of the array will be simply the sum of the probabilities of all the points in the array. Most situations, however, lie in between these extremes so that the degree of overlap between the ESAs of sampling points will be substantial for large individuals but minor, or even 0, for smaller ones (see Fig. 1).

Now, consider corals that are sampled using the PIM with an array of n points. As opposed to plot-based techniques, where the area sampled is clearly defined, when using the PIM the area sampled must be estimated a posteriori using the ESA for each coral size class. Thus, for each coral sampled from size class i, we must record its radius ri in order to obtain its size-specific ESA (ESAi). The ESA can be calculated specifically for each size class of corals as the union of the buffers around individual sampling coordinates:

$$ESA_i = \bigcup_{k=1}^{n} \pi r_i^2_{[k]}$$

Here, k is an index of each sampling point and n is the total number of sampling points. This solution takes into account, in addition to the size of the sampled individuals and the number of sampling points, the spatial organization of the points, as the union of the buffers depends on their overlap. The calculated ESA is exact and assumption-free in relation to the spatial distribution of the sampled population/community.

We have assumed that colonies have a circular boundary, which is a reasonable assumption for many organisms (Weinberg 1981, Marsh et al. 1984, Zvuloni et al. 2008). However, deviations from circularity will not bias the results as, on average, the projected length of a non-circular object being rotated around a centroid will be similar to the diameter of a circular object with the same surface area.

Our ability to calculate the ESA for sampling each size class of corals allows us to estimate the true density Di of corals from size class i:

$$D_i = \frac{N_i}{ESA_i}$$

where Ni is the number of sampled corals from size class i and ESAi is the ESA for corals from size class i,
for this specific array of points (see Eq. 1). Based on Eq. (2), the total density of corals across all m sampled size classes ($D_T$), can be calculated as

$$D_T = \sum_{i=1}^{m} D_i$$

(3)

Thus, in order to obtain the density of the entire population ($D_T$) the density of each size class should be estimated separately, and then all estimated densities should be combined.

The approximately unbiased estimate of the frequency of corals from size class $i$ ($f_i$) can be simply calculated as the fractional contribution of corals from this size class to $D_T$:

$$f_i = \frac{D_i}{D_T}$$

(4)

The above approach can be applied using the Excel spreadsheets and R functions provided in Supplement 1.

Obtaining the corrected SFD (the distribution of $f_i$’s across all size classes) of the population from PIM data is the first step toward calculating other count-based measures. For example, using the corrected frequency of each size class within the population, the mean radius of individuals in the population ($\bar{r}_p$) can be calculated as:

$$\bar{r}_p = \sum_{i=1}^{m} f_i \cdot r_i$$

(5)

Obtaining measures related to species composition

Measures related to species composition (e.g. species diversity) are usually based on the relative abundance or relative coverage of species within the studied community. Relative coverage is not a count-based measure and as such it is not prone to size-related biases when using the PIM. However, relative abundance is based on individual counts and, therefore, the PIM leads to a positive bias toward counting the larger individuals as well as the larger species. We suggest overcoming this bias by using Eqs. (1)–(3) to first calculate corrected relative density across species. The approximately unbiased relative occurrence of species $s$ ($P_s$) can then be calculated as its relative density:

$$P_s = \frac{d_s}{\sum_{s=1}^{q} d_s}$$

(6)

where $d_s$ is the density of species $s$ among the $q$ species found within the studied community. We note, that in order to obtain $P_s$, the density of each species among the $q$ species should first be estimated separately (using the Excel spreadsheets and R functions provided in Supplement 1). All estimated densities should then be inserted into Eq. (6).

RESULTS

Simulation examples

To demonstrate the size-related biases associated with the PIM and the outcome of our suggested approach, we simulated 3 hypothetical populations of corals. Each population was composed of 4500 individuals belonging to 4 different size classes; coral radii were taken to range from 5 to 20 units of length for each size class, with a difference of 5 units between them. We tested 3 different SFDs of corals: (1) uniform distribution, (2) skewed to the right, and (3) skewed to the left. The SFDs were respectively set as

$$\{f_1, f_2, f_3, f_4\} = \{0.25, 0.25, 0.25, 0.25\}$$

(7a)

$$\{f_1, f_2, f_3, f_4\} = \{0.4, 0.3, 0.2, 0.1\}$$

(7b)

$$\{f_1, f_2, f_3, f_4\} = \{0.1, 0.2, 0.3, 0.4\}$$

(7c)

for each size class, as represented by the white columns (‘actual’) in Fig. 2A–C. In the simulation, the corals were randomly scattered over a plane of 3000 by 3000 units, with no overlap between individuals. This area was sampled by 40 randomly placed sampling units, with each unit consisting of an array of 100 points, spatially organized as a matrix of 10 rows and 10 columns with a constant interval of 15 units between neighboring points. The number of individuals of each size class that were intercepted by the points was recorded.

We first calculated the estimated SFDs without treating the biases related to the PIM. Here, we simply considered all individuals intercepted with the 100 sampling points. As can be seen by comparing the black vs. white columns in Fig. 2A–C, the estimated SFDs significantly differed from the true SFD (i.e. actual) in the 3 simulated populations ($p_v << 0.01$ for all tested SFDs; Kolmogorov-Smirnov test). We then applied the correction given in Eqs. (1)–(4) to the observed SFDs. The corrected SFDs are statistically indistinguishable from the original SFDs that were defined in the simulations (compare gray vs. white columns in Fig. 2A–C; $p_v = 0.86$, $p_v = 0.38$, and $p_v = 1$ for uniform, left-skewed, and right-skewed SFDs, respectively; Kolmogorov-Smirnov test). In addition, as can clearly be seen in Fig. 2D,
the estimated densities of the 3 populations after applying our approach were almost identical to the actual one.

**Real data example**

Based on real data from the shallow reef of the Gulf of Aqaba (Red Sea), we compared the SFD and density of the corals as evaluated by the PIM to those evaluated by the conventional quadrat method (QM; see the sampling scheme in Supplement 2 at www.int-res.com/articles/suppl/m556p123_supp/). As shown in Fig. 3A, the PIM without the correction showed a significantly lower abundance of smaller corals and higher abundance of larger corals in comparison to the conventional QM (compare black vs. white columns; \( p_r << 0.01 \), Kolmogorov-Smirnov test). However, by using our mathematical corrections, we retrieved the size-frequency histograms to be almost identical to that evaluated by the QM (compare gray vs. white columns; \( p_r = 0.996 \)). In addition, the density evaluated by the PIM was similar to that evaluated by the QM (Fig. 3B).

**Use as a stratified sampling approach**

As can be seen from the simulation results, the PIM has an inherent bias toward sampling large individuals and naturally puts ‘more effort’ into counting them (Fig. 4). This has 2 major consequences. The first is that for a given number of points, for very small individuals the total area sampled will be too small to get reliable density estimates. Hence, it is advisable to sample only individuals beyond a certain size for which the ESA is sufficiently large or consider increasing the density of sampling points.

The second consequence of this bias towards large coral may be advantageous as it provides a natural stratified sampling approach. As most populations in nature show a high frequency of small individuals and low frequency of large ones (e.g. Enquist & Niklas 2001, Lins de Barros & Pires 2006, Caroselli et al. 2012, Leray et al. 2012, Shaked & Genin 2014), we are in essence putting more effort into uncommon groups (large size classes) and less into the common ones (small size classes).

To test this hypothesis, we sampled the right-skewed SFD population, using 2 sampling ap-
proaches: (1) the conventional QM, in which we sampled the population with randomly placed 100 × 100 unit quadrats. The SFD was calculated only for individuals with centers lying within the quadrats (following the ‘center rule’ scheme of Zvuloni et al. 2008); and (2) the PIM, in which we sampled the population with arrays of 100 points as described above. In this case, we recorded all individuals intercepting the sampling points and calculated the SFD following our proposed correction (Eqs. 1–4). We then tested which approach estimates more accurately the actual SFD. We based our analysis on the assumption, grounded on our experience with corals, that most of the sampling time is devoted to recording and measuring individuals and not to placing the sampling units over the study site. Thus, the sampling effort, expressed by the number of sampled individuals, was taken as the independent variable, while the dependent variable was the error, calculated as the root mean squared error (RMSE) between the assessed and the true frequencies of the 4 size groups.

We used 1000 simulations to generate curves of the mean error as a function of the number of sampled individuals (i.e. sampling effort). As can be seen in Fig. 5, the PIM reaches low errors (better accuracy) earlier than the commonly used QM.

**DISCUSSION**

Numerous survey methods are used for collecting data on the abundance and diversity of organisms in a general effort to develop high-quality conservation and management strategies (see reviews by English et al. 1997, Hill & Wilkinson 2004). The PIM is an efficient and applicable sampling technique, and is therefore extensively used in both terrestrial and marine ecology (Kohler & Gill 2006). However, the weaknesses of the PIM, as a plotless technique that cannot estimate measures relating to density and significant size-related biases, have restricted the use of this method to indices related to coverage.

In this study, we have demonstrated that by following the proposed bias corrections, the PIM can also be used to successfully obtain estimates of various useful count-based measures, such as density and SFD, providing the sizes of individuals sampled by...
the sampling points are also recorded. Moreover, we show that the inherent size-related biases associated with this method may be advantageous for studying most natural populations. In essence, the PIM provides a natural stratified sampling approach, in which more effort is put into uncommon groups (large size classes) and less into the common ones (small size classes). Our simulation results clearly show that for most cases in nature, where populations (e.g. Enquist & Niklas 2001, Shaked & Genin 2014) and communities (Brown & Nicoletto 1991, Gaston & Blackburn 2000) show a high frequency of small individuals and a low frequency of large ones, the PIM reaches low errors (better accuracy) much earlier than the commonly used plot-based techniques. For example, in order to reach an RMSE of 0.04 in assessing the SFD, only 56 individuals have to be sampled by the PIM compared to 100 individuals by the QM (ca. 56% of the sampling effort; Fig. 5).

The approach described in this study can be applied using the Excel spreadsheets and R functions provided in Supplement 1. By inserting the diameters of individuals sampled by the PIM, the user can easily obtain, in addition to coverage, the SFD, density, and mean size of the sampled population. These data can later be used to obtain approximately unbiased measures of relative species abundance and diversity.

It should be noted that the detection of very small individuals, which will hence have very small ESA, is problematic (Fig. 4). For example, when a small individual is sampled, it will substantially increase the total density estimate, as the denominator in the density calculation will be very large. Conversely, due to the low ESA of small individuals, they may often go undetected, leading to an underestimation of total density. Thus, the density estimate of small individuals with ESA close to 0 will be highly variable (depending on whether small individuals where sampled or not). Hence, we recommend that density should be estimated only for individuals above a minimal size threshold. The choice of this threshold can be assessed by examining the ESA in relation to individual size (Fig. 4). Sizes for which the ESA is very small should be excluded. Of course, this threshold can be reduced by increasing the number of sampling points used (Fig. 4). It is important to note that small individuals can also be missed in plot-based methods, and often the sampler must define minimum individual size for sampling.

The use of digital data analyses, such as photography and videography, has been increasing in popularity. These methods can greatly reduce field expenses because they require much less time in the field compared to manual survey methods. For example, the National Coral Reef Institute introduced a popular (3500 estimated users worldwide) freeware, CPCe (Coral Point Count with Excel extensions; Kohler & Gill 2006), that provides a tool for cover estimation from photographs using the PIM. The integration of our approach with similar software (e.g. by providing computer-generated size estimates of sampled individuals) will help to extract the greatest possible amount of information from each photograph.

In sum, we have presented a new way to look at the sampling area of plotless sampling units, viz. the ESA. Through this approach, we have improved the PIM, enabling the user to derive useful ecological measures, while preserving the efficiency of the method. We also show that the inherent biases associated with the PIM may be advantageous when studying most natural populations, as it provides a natural stratified sampling approach. Thus, the PIM with the proposed correction is even more efficient than conventional plot-based techniques, and should be more widely used by marine and terrestrial ecologists alike.

Acknowledgements. We thank N. Paz for editorial assistance and M. Kiflawi and R. Holzman and 4 anonymous reviewers for their insightful comments and suggestions. We also thank the Interuniversity Institute (IUI) in Eilat for the use of their facility.
LITERATURE CITED


Leray M, Béraud M, Anker A, Chancerelle Y, Mills SC (2012) Acanthaster planci outbreak: decline in coral health, coral size structure modification and consequences for obligate decapod assemblages. PLOS ONE 7:e35456


Editorial responsibility: Jana Davis, Annapolis, Maryland, USA

Submitted: February 8, 2016; Accepted: July 25, 2016

Proofs received from author(s): September 1, 2016