How Many?
(Basics of Bangarang density estimation)

Eric Keen

Abstract
This Backgrounder reviews the theoretical bases for extracting density information from the three kinds of visual surveys we conduct onboard the Bangarang: strip-width surveys, distance sampling, and capture-recapture. Simple, short, and a good start – nothing more. Corgis are mentioned throughout.

Contents
Introduction

Strip-Width Surveys

Distance Sampling
What to measure
Clustered objects
Assumptions
Scenario
Detection function
Estimating density

Capture-recapture
The Simple Petersen Estimate
Common Assumptions
Open populations, discrete sampling
Closed populations, continuous sampling
Superpopulations
Photo-grading

1 Bangarang Backgrounder are imperfect but rigorous reviews – written in haste, not peer-reviewed – in an effort to organize and memorize the key information for every aspect of the project. They will be updated regularly as new learnin’ is incorporated.
Introduction

“Unfortunately, the simple question of ‘how many are there?’ can be difficult to answer robustly. That challenge is particularly great for cetaceans in coastal and riverine habitats.”

Dawson et al. (2008)

\[
\text{Density} = \frac{\text{Number of objects}}{\text{Area}}
\]

Density and population size are related as \( N = D \times A \) where \( A \) is the size of the study area. But when the size of the area in question is not well-defined, it is safer to report density instead of population size. Another option is to use encounter rate, \( n/L \), which is essentially a one-dimensional density, a crude but useful relative density index.

Whether reported densities are absolute or relative, the ability to draw population inferences from transect results depends on both the taxon in question and the method of sampling. The premise of all the sampling methods I mention here is that you cannot adequately survey your entire study area in full. You therefore have to visit only a portion of the area and extrapolate local density from your subsample. There are trade-offs between accuracy, precision, spatial coverage and feasibility.

In the Bangarang Project, my minimum goal is to determine spatial patterns (relative density and aggregations) and associations among marine tetrapods, indices of their prey, and the physical environment. For some taxa I will be able to go beyond this and estimate local abundance. Designing the fieldwork to accomplish this requires a solid understanding in the basic principles at work in my three modes of visual survey: strip-width, distance, and capture-recapture surveys.

As indicated in the quote from Dawson et al. (2008) above, robust survey design is difficult in confined waterways and adequate design is essential to the objectives of the Bangarang Project. To achieve my goals I have incorporated elements of design for strip-width surveys, distance sampling and capture-recapture, but this Backgrounder barely touches on study design at all, except to emphasize the most important points. Design is discussed at some length in the 2014 Methods Backgrounder.

Instead this Backgrounder focuses on the basic principles for the various means of density estimation I hope to invoke in my research. This is not in any way exhaustive or complete. These are the absolute basics, though hints are given about which sophisticated methods will be most relevant for the Bangarang Project down the road. Nearly all of the strip-width and distance sampling sections are based on Buckland et al. (2001)\(^2\), and nearly all capture-recapture information is from Armstrup et al. (2005).

Strip-Width Surveys

In strip transect sampling rectangles are placed randomly in the study area. Observers travel down the centerline of each rectangular strip, counting all objects within the strip as they go. The rectangles can be any variety of length, but their width must be small enough that absolutely no objects are missed by the observers as they travel along.

In a strip-width survey design there are $k$ strips, each of width $2w$ ($w$ is the distance from the centerline to either side boundary of the strip) and length $l$. The total length $L$ is the summed length of all $k$ strips. Therefore the sampled area is $a = 2wL$. If the total number of objects counted is the sum of counts from all $k$ strips, then the estimated density $\hat{D}$ of objects is:

$$\hat{D} = \frac{n}{a} = \frac{n}{2wL}$$

Again, a critical requirement of a strip-width survey is that 100% of events within the strip are observed and noted. To meet this assumption sometimes the strip width has to be very small, which can make this method quite inefficient. But the analysis is intuitive and simple.

Another assumption is that the strip width is easily known in the field and that animals are counted accurately as within or without the strip. Some rangefinding technique is usually used in fieldwork to ensure this assumption is met.

Distance sampling

Line transect methods can be used to estimate absolute abundances of populations, and are common in cetacean surveys (Kinzey et al. 2000). Unlike strip-width sampling (see below), which limits the area you search and assumes that you see everything in that confined search area, line transect sampling requires that you record everything you see, regardless of distance, but that you also record the distance of each sighting from your trackline. This is why it is commonly known as distance sampling.

All objects on or very near the transect trackline need to be detected, but a proportion of objects within a certain distance of the trackline to be missed. This allows a wider strip to be used, providing a larger sample size for the same amount of effort.

Similar to strip-width methods, a line transect design comprises a series of randomly placed lines (or a systematically spaced grid). The difference is that density estimation from line transect surveys has to deal with uncertainty, because only a proportion of the objects in a surveyed area is detected. This unknown proportion is $P_a$.

$$\hat{D} = \frac{n}{2wLP_a}$$

What to Measure

The distances recorded during surveys are the perpendicular distances of objects to your trackline. Because you almost never see objects exactly to your beam during surveys, you have to take the appropriate measurements to calculate the objects distance from the line after the fact. These measurements are 1) the direct distance from your vantage to the object (called a “radial distance” here, measured using some range-finding technique) and 2) the angle created by that radial line and the trackline. With these two data, simple trigonometry will get you to the distance $x$ from the trackline for each sighting (as long as you are not accounting for the curvature of the earth).
Clustered Objects
When an object of study (e.g. whales) can exhibit a clumped distribution (i.e., found in groups) then special practices must be followed. Clusters rather than individuals are counted (though there can be a cluster of 1 animal), meaning groups not individuals are the $n$ in the above equation. Density is therefore expressed as the density of clusters $D_s$ times the average cluster size $E(s)$.

$$D = D_s \cdot E(s)$$

$E(s)$ is usually calculated as the mean size of the $n$ detected clusters, unless detection is a function of cluster size – which it often is. It’s easier to see a group of 10 humpback whales from afar than a group of 2. This introduces a size-bias to your density estimates. There are various ways around this size-biased sampling in the analytical phase, and field practices are not affected.

Another rule: If a cluster falls very near the effective strip ($w$, also called the ‘truncation distance’, beyond which sightings are not incorporated into analyses), then it can only be counted if the center of the cluster occurs within $w$ of the trackline. If the center of the cluster is outside the strip, the sighting cannot contribute to distance sampling analyses. This rule can be used to cull sightings data after fieldwork is done.

Assumptions
The three assumptions essential to unbiased density estimation are:

1. Objects directly on the line are always detected (i.e., they are detected with probability 1.
2. Objects are detected at their initial location, prior to any movement in response to the observer.
3. Distances (and angles where relevant) are measured accurately (ungrouped data) or objects are correctly counted in the proper distance interval (grouped data).

Other important assumptions include the accurate identification of species and estimates of group sizes. The following scenario will show you why these assumptions are so important.

Scenario
Let’s say you’re conducting a line transect survey to estimate the density of corgis in a field. You’re walking down a randomly placed line and you are setting your truncation distance to 80m. Your total effective strip width is therefore 160m, which means you’ll end up sampling approximately 50% of the field. This 80m strip to one side can be divided into, say, 8 even sections. If you detect all corgis in your effective strip, your histogram of observations in each section will be flat: the same number of corgis can be expected in each 10m section of your surveyed area. So you do your survey, taking bearing angles and distances to all the corgis you see (tehee!). But the grass is tall and you might not have seen the corgis that near the edge of your strip. You plot your data and sure enough, the histogram bars in the most distant sections drop off. Hmm. Your histogram demonstrates that you need a correction factor to account for undetected corgi groups.

Therefore, in order to estimate density, our goal is to estimate the probability of detecting all the corgi clusters in your stip, $P_a$. The complication is that you will miss different proportions of objects at different distances. The proportion missed at each distance between the trackline (0) and the strip boundary ($w$) can be defined by a function based on your sightings data, called the detection function. Therefore the probability of detecting an object is:

$$P_a = \frac{\int_0^w g(x)dx}{w}$$

The $g(x)$ function in there is the probability of detecting a corgi at a certain distance $x$. We have to determine what this function is before we can estimate $P_a$. This function $g(x)$ is called the detection function, and it is a pillar of distance sampling theory.

The Detection Function
\[ g(y) = \text{the probability of detecting an object, given that it is distance } y \text{ from the random line} \]

Or, \( g(y) = \text{pr}\{\text{detection } \mid \text{distance } y\} \)

Because \( g(y) \) is a probability, its value always ranges between 0 and 1, no more, no less. It generally decreases with distance from the trackline, and as a rule \( g(0) = 1 \) (see Assumption 1 above). The detection function can also incorporate variables (v) that affect the detection of \( y \), written as \( g(y \mid v) \).

To develop a detection function for our corgi scenario, you would fit a quadratic equation to the midpoints of each of the bars in the histogram of your data. The fitted line will drop off as it nears the 80m mark. This fitted quadratic function is the estimated detection function for your data. So that’s pretty cool: you can estimate the proportion you missed by fitting a function to the portion you saw.

**Estimating Density**

Now that we have \( g(x) \), we can return to the Pa equation. Before we solve for Pa, we can plug its definition into the density equation and cancel out the \( w \)'s (meaning the density estimate is valid whether \( w \) is bounded or unbounded, which is pretty cool):

\[ \hat{D} = \frac{n}{2L \hat{g}(x) dx} \]

We can then compute the integrand using our detection function. We'll call the computed integrand \( \mu \) by convention.

\[ \hat{D} = \frac{n}{2L \hat{\mu}} \]

We can do a couple more steps to make this equation much more convenient and accessible by replacing the term \( 1/\mu \) with a more conventional function. To do so, we consider the probability density function (pdf) of the perpendicular distance data:

\[ f(x) = \frac{g(x)}{\int_0 g(x) dx} \]

Now consider that the expected number of objects at distance \( x \) is independent of what that distance is (this is why random placement of transect lines is so important). That is, the pdf of the distance data is the exact same shape as the detection function. It can therefore be determined by *scaling* the detection function.

Now consider that \( g(0) = 1 \), which is the central assumption of distance sampling. Therefore:

\[ f(0) = \frac{1}{\int_0 g(x) dx} = \frac{1}{\mu} \]

\( f(0) \) can therefore be calculated from measured distances. As a probability density function it is a well-studied feature of mathematics. It would be easier if we used the pdf in our density equation instead of the \( 1/\mu \) term. So our new density equation is:

\[ \hat{D} = \frac{n \cdot \hat{f}(0)}{2L} \]
Capture-Recapture

The premise of capture-recapture methods is that population parameters can be estimated by visiting the study population, “marking” a random sample from the population, re-releasing them, then repeating the process on subsequent visits. The proportion of marked individuals you find in your subsequent samples can be used to estimate population size and other parameters. “Marking” can take on various forms, but in cetacean surveys this usually means taking photographs of unique identifiable features on whales (e.g. ventral surface of humpback flukes) so that you can track an individual’s presence in the study area over time. Photo-identification is a robust and accepted means of capture-recapture analysis for whales3, 4, 5.

The Simple Petersen Estimate

All capture-recapture models are descended from the simple original model developed by Petersen in the 1890s6. Here’s how it works7:

We are interested in estimating the size of a population. We are able to visit the population twice, and we have to make the most of it. The population of interest is closed: there is no birth or immigration.

On our first visit we catch a random sample of individuals, r, from the population. We mark them so as to recognize them later. This sample (r) represents some proportion of the total population size (N), though we don’t know what N or that proportion is yet (if we did our job would be done!). We return the marked individuals to the population. They remix.

Between our visits there is death and there is emigration, but our marked individuals are as likely to be subject to them as the rest of the population. The marked proportion will therefore remain the same.

On our second visit we catch another random sample from the population, n. Some of the individuals (m) in this new sample had been captured before by us – they are already marked. In the same way that death struck marked and unmarked individuals indiscriminately between our visits, the proportion of marked individuals in our new sample should represent the proportion of marked individuals in the entire population (N).

\[
\frac{m}{n} = \frac{r}{N}
\]

From this we can estimate N:

\[
\hat{N} = \frac{r \cdot n}{m}
\]

Simple enough, and this is as simple as it gets. The Simple Petersen estimate. But there are only very strict and unrealistic conditions to which this model can apply (see assumptions below).

Bailey (1951) found that a slight mathematical tweak makes the Petersen estimate more accurate:

\[
\hat{N} = \frac{r(n+1)}{(m+1)}
\]
Adding one to \( r \) and \( m \) improve accuracy when sample sizes are small. When numbers are large, the small addition makes a negligible difference, making Bailey’s modification a more generally applicable simple estimator. Chapman (1951) further modified this as follows:\(^8\):

\[
\hat{N} = \left( \frac{(n_1 + 1)(n_2 + 1)}{(m + 1)} \right) - 1
\]

Here \( n_1 \) is the same as \( r \) above. To estimate abundances from more than two trapping trips, a weighted mean model was developed. The weighted mean model assumes everything that the Petersen model does but utilizes data collected over several days. In averaging the population estimates for each pair visit, weight is given to those in which the sample size is highest (because they tend to be the most accurate).

**Common Assumptions**

The Petersen estimate is simple because it makes a lot of assumptions about the population of interest and our methods of marking them. Subsequent models get complicated in an effort to work around these assumptions.

1. All marks are permanent and noted correctly upon recapture for the duration of the project.
2. Being marked has no effect on individual’s chance of recapture.
3. Being marked has no effect on chances of dying or emigrating.
4. All individuals have an inherently equal chance of being captured.
5. All animals, marked or not, have equal chance of dying or emigrating.

Basically, the marked individuals have to be truly representative of the whole population. Most models either have to assume the above conditions or develop sophisticated methods for working around them. But there are two assumptions of the Petersen model that are side-stepped by many:

6. The Peterson model assumes that the proportion of marked individuals in the population will not change between sampling visits. That is, the population is closed for the period of estimation. This means one of the following must be true:
   a. There are no births or immigrations
   b. There are no deaths or emigrations
   c. There are neither gains nor losses to the population.

7. Sampling periods are short and discrete in time: processes of gains and losses to the population do not occur during the sampling period.

In cetacean studies it is very difficult if not impossible to meet many of these assumptions. From Barlow et al. (2011): “In whale capture-recapture studies, sample heterogeneity can arise from many sources:

1. Individual behavioral differences (e.g. approachability or tendency to fluke up prior to a dive.
2. Geographic preferences of whales combined with non-random geographic coverage, and
3. Biased sex ratios in breeding areas.

Mortality, reproduction, emigration, and immigration during the sampling period can be viewed as heterogeneity affecting the likelihood of an individual being sampled in different sampling periods.”\(^9\)

**Open Populations, Discrete Sampling Periods**

An open population is one that is (or could be) changing during the course of a study due to some combination of births, deaths, immigration or emigration\(^10\). Two common models for open populations are the Cormack-Jolly-
Seber model and the Jolly-Sever model. The Cormack-Jolly-Seber (CJS) model is based solely on recaptures of marked animals and provides estimates of survival and capture probabilities only\textsuperscript{11}. The Jolly-Seber model (JS) incorporates ratios of marked to unmarked individuals and can therefore be used for abundance estimation (as well as capture probability and survival rates)\textsuperscript{12}.

Seber (1982, p. 196) has a detailed description of the Jolly-Seber model\textsuperscript{13}. JS models require that during each “trapping occasion” the number of previously marked and hitherto unmarked individuals be recorded. JS model has the following assumptions:

1. Equal probability of capture.
2. Equal chance of survival.
3. Marked animals do not lose their marks and marks are not overlooked.
4. Sampling periods are short (effectively instantaneous).
5. All emigration from the population is permanent.

To use the JS model, the marked population size ($M_j$) right before each trapping trip must be estimated. We can’t assume it is the total of all individuals marked from previous trips because this is an open population: marked individuals could be dying or emigrating. To estimate this, we have to rely on the assumption of equal catch probability:

When we go out for a trapping trip ($trip_j$), a subset of our population will already be marked ($M_j$). During trip$_j$, we will capture and release a subset of the population ($R_j$). Some of these ($m_j$) will be individuals we’ve already marked in previous trips. But there is a portion of the previously marked population that we will not reencounter ($M_j - m_j$). Because our chances of recapturing individuals is the same for all individuals (regardless of being marked or not), we have the same chances of recapturing our newly released marks ($R_j$) as we do the proportion of previously marked animals we didn’t see last time ($M_j - m_j$). Our recapture rates for both groups will be the same. We can then make a ratio:

$$\frac{Z_j}{(M_j - m_j)} = \frac{r_j}{R_j}$$

$Z_j$ is the recapture rate of the marked population before trip $j$
$r_j$ is the recapture rate of the animals we release after trip $j$

If we rearrange this, we can estimate the marked population just prior to our $j^{th}$ trapping trip.

$$\hat{M}_j = m_j + \frac{R_j Z_j}{r_j}$$

The unbiased version of this is:

$$\hat{M}_j = m_j + \frac{(R_j + 1)Z_j}{r_j + 1}$$

Now, very similar to the original Petersen model, the population size for period $j$ can be determined by equating the sample and population ratios of marked to total animals:

$$\frac{m_j}{n_j} = \frac{M_j}{N_j}$$

Which can be rearranged to estimate $N$:

\[ \hat{N}_j = \frac{M_j \cdot n_j}{m_j} \]

Because the probability of recapture during a trapping period \((p_j)\) is the ratio of re-captured individuals to the total subpopulation of marked individuals, \(m_j / M_j\), the above equation can be rewritten as:

\[ \hat{N}_j = \frac{n_j}{\hat{p}_j} \]

And the unbiased version of this population model is:

\[ \hat{N}_j = \frac{\hat{M}_j (n_j + 1)}{m_j + 1} \]

The major assumption upon which the JS model hinges is equal capture probability. If this can’t be met, the model as is must be considered biased\(^{14}\).

**Closed Populations, Continuous Sampling**

In some circumstances you can only conduct a single trapping trip during which captures/sightings of individuals occur at any time during the data collection. There is thus only 1 capture on each capture occasion. These situations are analyzed using “continuous-time” models and have been particularly helpful in previous whale surveys (e.g., Whitehead and Arnason 1987\(^{15}\)). Continuous time models are akin to the framework for species richness (Fisher et al. 1943, Bunge and Fitzpatrick 1993, Chao 2005, Wilson and Anderson 1995, Wilson and Collins 1993)\(^{16}\).

**Superpopulations**

Crosbie and Manly (1985) and Schwarz and Arnason (1996\(^{17}\), 1999\(^{18}\)) reparameterized the JS model for a new parameter, the size of a superpopulation. The superpopulation can be thought of as either the total number of animals available for capture at any time during the study or, alternatively, as the total number of animals ever in the sampled area between the first and last sampling occasions of the study\(^{19}\). Superpopulation capture-recapture models, called POPAN, are able to account for the fluid nature of annual residency at migratory destinations. The POPAN model has also been modified to account for heterogeneous capture probability linked to reproductive cycles (POPAN – tao)\(^{20}\).

---


Photo Grading

When photographs are used in the estimation of population parameters, they are usually scored for quality control. A scoring process helps to avoid heterogeneity in capture probability between highly and moderately marked individuals. Photo scoring has been used in Hammond (1986), Hammond et al. (1990), Friday et al. (1997), Rugh et al. (1998) and many others. For SPLASH (Structure of Populations, Levels of Abundance and Status of Humpback whales, the landmark study of North Pacific humpbacks), the data coordinator rated each fluke photograph for...

- Pigmentation pattern
  - five ranked categories from all white to all black
- Photographic quality (ranked 1 – 5 on each of five features)
  - Proportion visible
  - Vertical angle
  - Lateral angle
  - Focus/sharpness
  - Exposure
- Distinctiveness of trailing edge (1-5)
- Degree of scarring (1-5)
- Presence of killer whale tooth rake marks (1-5)

To ensure consistency, photographs were rated by the same person throughout this study. “Each season, a catalog was constructed for each region using the single best photograph of each individual seen in that region and season.” Matches were made by a team of 6, and suspected matches were verified by a second matcher. “Friday et al. (2008) showed that the best tradeoff between bias and precision for humpback capture-recapture estimates in the Atlantic was obtained by eliminating photographs of partial or half flukes and photographs in the poorest of four categories of photographic quality. We did not attempt to match photographs with a score of 4 or 5 in any of the five measures of photographic quality (e.g., eliminating flukes with less than 50% visible) and photographs with a score of three in four or more measures of quality. The ranked measures of fluke distinctiveness were not used in analyses.”

The downside is that culling photographs limits sample size and restricts options for analysis. To maintain sample sizes, sophisticated modeling and simulation studies have been developed that mitigate bias.

Photographs can also be used to evaluate error rates. In SPLASH, “Two approaches were used to quantify the number of between-season matches that were missed by our matching process. First, different teams of matchers repeated pair-wise comparisons between nine pairs of catalogs, each representing one season’s sampling in one area. There were a total of 165 known matches in 3,779 possible comparisons, and the second team found 148 of these. In addition, to help estimate match success rate, photographs of 266 individuals that were known to match previous catalogs were added to the 2006 season matching effort. These “seeded” photographs were not the identical images in those catalogs, and the matchers did not know which photographs had been added.”

---

25 Barlow et al. 2011.