Methods used in the Arctic PRISM surveys

Article *in* Studies in Avian Biology · January 2012

6 authors, including:

Vicky Johnston
Environment Canada
23 PUBLICATIONS **248** CITATIONS

SEE PROFILE

Paul A. Smith
Environment Canada
121 PUBLICATIONS **1,686** CITATIONS

SEE PROFILE

Jennie Rausch
Environment Canada
29 PUBLICATIONS **184** CITATIONS

SEE PROFILE

Stephen Brown
Manomet
61 PUBLICATIONS **1,047** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Predation danger and shorebird ecology View project

Project Avian Mortality at Communication Towers View project

Methods

*Jonathan Bart, Victoria Johnston, Paul A. Smith, Ann Manning,
Jennie Rausch, and Stephen Brown*

Abstract. Detecting declines in population size is one of the highest priorities of the shorebird initiatives in Canada and the United States. The quantitative goal is 80% power to detect a 50% decline, occurring during no more than 20 years, with a significance level of 0.15, using a two-tailed test, and incorporating effects of potential bias into the estimator. The Arctic PRISM program was designed to achieve this goal for arctic-nesting shorebird populations. The survey methods are an application of double sampling. Rapid surveys were made on a large number of plots selected from throughout arctic Alaska and Canada using stratified random sampling. Intensive surveys were made on a subsample of the plots to obtain detection rates, which were used to calibrate results from rapidly surveyed plots. Surveys will be made of the entire arctic region, each lasting several years and producing an estimate of average population size during the survey period. Results from two or more survey periods will be used to estimate change, or trend, in population size.

Key Words: arctic, monitoring, population size, PRISM, shorebirds.

A major goal of Arctic PRISM is to estimate change in shorebird population size occurring during 20 years with power of 80% to detect a 50% decline occurring in no more than 20 years, using a significance level of 0.15 and a two-tailed test, and acknowledging effects of potential bias (Skagen et al. 2003, Bart et al. 2005). This chapter describes the methods being used to achieve the desired power. We discuss the delineation of plots and strata, selection of plots to be surveyed, how the surveys were conducted, and analysis of the resulting data. Methods for analyzing habitat data varied between regions and are described in region-specific chapters (chapters 3–8, this volume). The methods described in this chapter were first developed in northern Alaska during 1997–2001. The basic approach was described by Bart and Earnst (2002). A detailed description of the field methods, including training, was provided by Bart and Earnst (2005) and has been expanded into a comprehensive Arctic PRISM manual (Rausch, Canadian
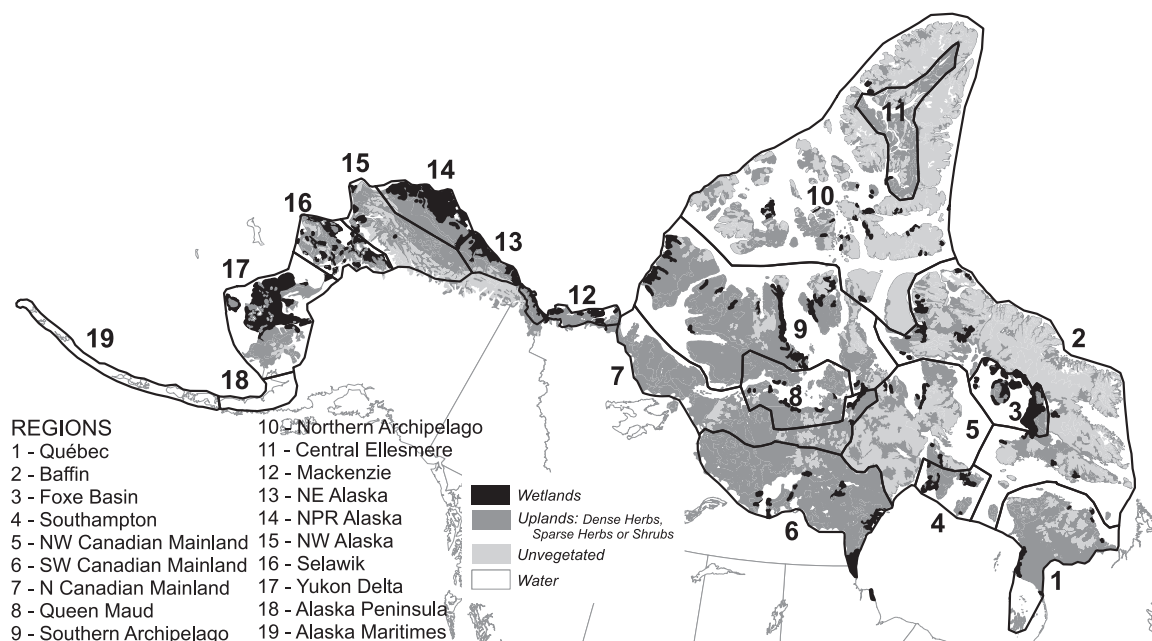
Figure 2.1. Study area, regions, and major habitats as depicted on the Circumpolar Arctic Vegetation Map (CAVM; CAVM Team 2003).

REGIONS
1 - Québec
2 - Baffin
3 - Foxe Basin
4 - Southampton
5 - NW Canadian Mainland
6 - SW Canadian Mainland
7 - N Canadian Mainland
8 - Queen Maud
9 - Southern Archipelago
10 - Northern Archipelago
11 - Central Ellesmere
12 - Mackenzie
13 - NE Alaska
14 - NPR Alaska
15 - NW Alaska
16 - Selawik
17 - Yukon Delta
18 - Alaska Peninsula
19 - Alaska Maritimes

Wetlands
Uplands: Dense Herbs, Sparse Herbs or Shrubs
Unvegetated
Water

Wildlife Service, unpubl. report). In this chapter, we concentrate on the quantitative methods involving the use of double sampling, a topic which has not previously been fully described.

The study area is the arctic portion of North America, as delineated on the Circumpolar Arctic Vegetation Map (CAVM; CAVM Team 2003), with modifications to exclude some mountainous areas (Fig. 2.1). The study area was partitioned into 19 regions (Fig. 2.1) based on logistic considerations. Prior to conducting surveys in each region, the region was further divided into subregions on the basis of how much area could be covered by field crews and anticipated density of shorebirds. Large areas not suitable for nesting shorebirds, such as oceans, lakes, and barren areas, were excluded. Care was taken not to exclude barren areas unless we had convincing evidence that birds did not occur in them.

DOUBLE SAMPLING

A form of double sampling (Cochran 1977) was used to estimate density and population size. A large sample of plots was surveyed with a rapid method of unknown accuracy, and a subsample of these plots was surveyed intensively to determine actual numbers present. The ratio of the results from the rapid and intensive surveys was used to estimate the "detection ratio" and to adjust the results from the rapid surveys (Bart et al. 1998).

Compared to many bird survey methods, double sampling requires only a few assumptions. Specifically, double sampling produces unbiased estimates of population size, and thus trend in population size, if the nominal sampling plan is followed and if the counts on intensive plots yield unbiased estimates of numbers present. No other assumptions are required. Note that counts on the intensive plots do not each have to be completely accurate; they only have to be unbiased (overestimates must just balance underestimates). Note, too, that failure to follow the nominal sampling does not necessarily cause any bias in estimates; it only makes such bias possible. Finally, bias in population size estimates also does not necessarily cause any bias in trend estimates; it only makes such bias possible. Bias in the trend estimate is algebraically equal to trend in the detection ratio, 'expected value of the estimated population size/ actual population size'. Thus, if the expected value of the estimated population size was 15% too low during the first set of surveys and was 10% too low during the second set of surveys, then bias in the trend estimate would be 5%. If the population actually declined by 50%, then the expected value of the estimated change would be a 45% decline. Thus, small changes in the detection ratio, due

either to non-random selection of intensive plots or to errors on the intensive plot surveys, cause little bias in the trend estimate in relation to the large change (50% decline) that the surveys were designed to detect. Stated a different way, a bias of 5–10% causes little change in the power to detect a 50% decline. See Bart et al. (2004) for a more complete discussion of this issue and quantification of how bias affects power.

PLOT SELECTION AND SURVEYS

Each subregion was partitioned into plots, most of which covered 0.12–0.16 km^2 and were designed to be surveyed in 1.0–1.5 hours. Plots were substantially larger in some early years of the study. We defined "wetland," "moist," and "upland" habitats in each subregion (see region-specific chapters, this volume, for definitions) and calculated the proportion of each plot covered by each habitat type. Plots with no habitat suitable for nesting birds were deleted. Plots with only small amounts of wetland, moist, or upland habitat that were primarily covered by water were unsatisfactory due to high travel costs (it did not make sense to spend hundreds or thousands of dollars to visit a very small plot). We therefore combined small plots with surrounding plots to reduce the variance in plot size.

Plots were assigned to wetland, moist, and upland habitat types. Habitat types were used (along with region) to define strata; thus each plot had to be assigned to a single habitat type. The rules used to make these assignments varied across the arctic because the extent of different habitats varied substantially. In many cases, plots were assigned to the type corresponding to the habitat that covered the largest fraction of the plot (e.g., if wetland habitat covered more than 50% of the plot, the plot was assigned to the wetland plot type). If wetland habitat was rare within a subregion, a different rule was used; for example, 'if wetlands habitat covers more than 20% of the plot, then assign the plot to the wetland plot type; otherwise, assign the plot to the type corresponding to the habitat that covers the largest fraction of the plot'. The rules used in each region are described in more detail in the regional reports (chapters 3–8, this volume).

The sampling plan for selecting plots to survey involved stratification using subregion and plot type (wetland, moist, upland), followed by selection of clusters of plots and then selection of plots. We selected plots to survey in groups to
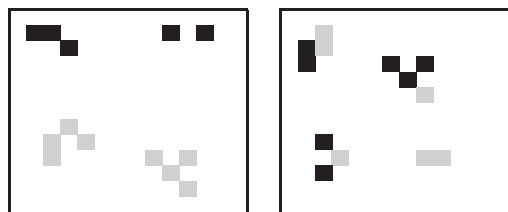

Figure 2.2. Standard cluster sampling (left) in which all plots in each cluster are in one stratum (indicated by shading) and our sampling plan, in which plots in a cluster were often in different strata (right). We termed the groups of plots (right) "zones." Plots of the same habitat type, within a zone, were a cluster.

reduce distances between plots being surveyed at the same time. These groups usually included plots in different habitats, and thus in different strata. Selection of plots was thus not independent in different strata (i.e., plots in different strata were close together much more often than if we had used independent selection). We referred to the groups of plots as "zones" (Fig. 2.2) to distinguish them from clusters, which, by definition, are plots in the same strata. Most zones covered 4–36 km^2 and comprised 25 to a few hundred plots. We acknowledged the lack of independence in selecting plots in different strata by modifying the standard formulas for cluster sampling (see below). Some reviewers have had difficulty grasping why we had to define zones, but had we ignored the lack of independence caused by selecting plots in different strata within zones, our variance estimates would have had substantial negative bias. Zones to be surveyed were selected systematically to ensure even coverage across the subregion. Simple random sampling was used to select plots within clusters.

In the early years of the study, we attempted to carry out the steps above by hand. With large subregions, this was not possible and we were forced to use short-cut methods which inevitably caused us problems later in the analysis. We therefore prepared a series of ArcGIS tools, collectively referred to as the Arctic PRISM ArcGIS extension (Table 2.1), to automate delineation of plots and assignment of plots to clusters, zones, and strata. This tool was essential for partitioning large regions into plots. It is available free from the senior author and may be useful to others who need to define a rigorous sampling frame for large, heterogeneous areas.

Random selection of locations for intensive surveys, which is part of the double sampling

TABLE 2.1
Description of the tools in the Arctic PRISM extension of Arcview.

Tool	Function
Create zones and plots	Generates sampling plots contained within larger zones (in polygon shapefile format).
Merge zones and plots	Merges plots and/or zones that contain less than a specified area of suitable habitat.
Plot habitat summary	Adds fields to the plot attribute table representing the proportion of each plot occupied by each habitat type.
Assign plot types	Codes a field in the plot attribute table using habitat types within each plot.
Cluster area	Produces a comma-delimited text file summary of the plots.
Plot selection	Creates a random sample of plots, stratified by plot type.

protocol, proved to be impossible. We found that many selected zones (groups of plots that we might have used for intensive surveys) lacked a suitable campsite close enough to the plots or the selected plots had too few birds to be able to compute detection ratios. These problems eventually forced us to select intensive camp locations non-randomly on the basis of logistics and expectations about shorebird density. Selection of plots around camps followed a similar method. Areas thought likely to have shorebirds were first identified and, when possible, a random selection was made of locations for the plots. When the habitat was variable, an effort was made to distribute plots across habitats. When initial surveys failed to reveal territorial shorebirds, the plots were moved, though this happened only occasionally. Because intensive plots were not selected randomly, results of rapid surveys on intensive plots were not combined with results from other rapid surveys. The results from rapid surveys and the estimated detection rates were thus independent, a fact that had implications for variance calculations, as discussed below. Although we were disappointed not to be able to select intensive plots in a fully random manner, analyses of detection rates showed little variation in either habitat or density, suggesting that the selection bias, if any, in our estimates of detection rates was small.

Field work was conducted during 1998–2006 at numerous locations widely distributed across the study area. Rapid surveys used time-constrained, area search methods. Total time per plot varied among regions but was usually about 1 hour in Alaska and 2 hours in Canada. Observers covered each plot thoroughly, walking transects 25 m apart when the habitat was uniform and following irregular paths when waterbodies or other obstacles were common. When habitats were variable, more time was spent in areas where birds were thought to be more common, but all parts of each plot were covered. Surveyors recorded "*indicated pairs*" (nests, probable nests, single birds by sex) on plot maps and summarized their observations in tabular form immediately after each survey (before continuing to the next plot). In the analyses, we assumed that each indicated pair represented two birds.

During most years and in most survey areas, 1–2 camps were established for intensive surveys (at previously selected sites for random surveys when feasible). At each camp, 4–6 plots were established. Two to four surveyors spent 3–5 weeks visiting these plots every 1–2 days, searching for nests and non-nesting birds. Search effort was usually more than 30 hours per plot and was greater than 50 hours per plot in a few areas where shorebird density was extremely high. The intensive searches began as birds returned to the plots and continued past hatch. Surveyors attempted to find all active shorebird nests and to identify all other territorial shorebirds.

The issue of how much effort, on intensive surveys, is required to find nearly all territorial birds was investigated in our study area by Smith et al. (2009). Surveyors searched for shorebird nests at four locations, widely distributed across the arctic. At each location, two or more teams, working independently, conducted surveys of the same plots. At the conclusion of the surveys,

the teams compared results to determine how many nests each had missed. Results were then used to model the likelihood of missing a nest as a function of number and timing of visits, species, and other factors. After five visits, the estimated proportion of nests found was 0.88 for the most difficult species to detect and 1.00 for the easiest species to detect. These results show that nearly all nests surviving for at least 7–10 days are found, but nests that survive only a few days may not be found. The study by Smith et al. (2009) did not assess the overall likelihood of detecting territorial birds because they only recorded nests found. During all Arctic PRISM surveys made through 2006, nests were not found for 22% of all birds recorded. This fact, in combination with the results obtained by Smith et al. (2009), suggests that the total detection rate of territorial birds was close to 1.0. As noted above, even if the intensive surveys were biased, this does not necessarily cause any bias in trend estimates (though bias in population size estimates does occur).

ESTIMATORS

For notational simplicity, we assume that population size is the same in each year during each survey period. Population declines, or other changes, during the survey period reduce precision but cause no bias. This can be seen by adding a subscript to indicate year and then showing that random selection of which areas to survey each year guarantees that the estimate, assuming it is unbiased within the year, also yields an unbiased estimate across years.

For a given species and period, the estimate of population density (d; birds per km²) was

$$d = \frac{\hat{\bar{X}}}{\hat{R}} \qquad (1)$$

where $\hat{\bar{X}}$ is an estimate of the mean density of birds that would have been recorded if an indefinitely large sample of rapid surveys had been conducted and $\hat{R}$ is an estimate of the detection ratio (birds recorded/birds present) on the rapid surveys. $\hat{\bar{X}}$ was obtained from the rapid surveys; $\hat{R}$ was obtained from the intensive surveys. From the standard equation for the estimated variance of a ratio of independent random variables (Cochran 1997),

$$\hat{V}(d) = d^2 \left(\frac{\hat{V}(\hat{\bar{X}})}{\hat{\bar{X}}^2} + \frac{\hat{V}(\hat{\bar{R}})}{\hat{\bar{R}}^2} \right) \qquad (2)$$

The estimated population size was

$$\hat{Y} = Ad \qquad (3)$$

where A is the size of the study area. The variance of $\hat{Y}$ was estimated as

$$\hat{V}(\hat{Y}) = A^2 \hat{V}(d). \qquad (4)$$

To derive estimators for the terms in expression (2), let

$\bar{z}_{uhi}$ = mean number of birds recorded per plot in the ith cluster of type h plots in region u

$\bar{b}_{uhi}$ = mean area covered per surveyed plots in the ith cluster of type h plots in region u

a_{uhi} = area covered by all plots in the ith cluster of type h plots in region u

n_{uh} = number of clusters of type h plots surveyed in region u

N_{uh} = number of clusters of type h plots in region u

$\hat{\bar{X}}$ was estimated using the "combined approach" (Cochran 1977) for ratios with stratification:

$$\hat{\bar{X}} = \frac{\hat{Z}}{a} = \frac{\sum_u^U \sum_h^H \frac{N_{uh}}{n_{uh}} \sum_i^{n_{uh}} a_{uhi}(\bar{z}_{uhi}/\bar{b}_{uhi})}{\sum_u^U \sum_h^H \frac{N_{uh}}{n_{uh}} \sum_i^{n_{uh}} a_{uhi}}$$

$$= \frac{\sum_u^U \sum_h^H \frac{N_{uh}}{n_{uh}} \sum_i^{n_{uh}} \hat{Z}_{uhi}}{\sum_u^U \sum_h^H \frac{N_{uh}}{n_{uh}} \sum_i^{n_{uh}} a_{uhi}} \qquad (5)$$

The quantity $\hat{Z}_{uhi} = a_{uhi} (\bar{z}_{uhi}/\bar{b}_{uhi})$ is the estimated number of birds that would be recorded if all the plots were surveyed in the ith cluster of type h plots in region u. The numerator in expression (5) is the mean of the $\hat{Z}_{uhi}$ times the number, N_{uh}, of clusters in the stratum. The numerator may thus be viewed as an estimate of the number of birds that would be recorded on rapid surveys if all plots in all regions were surveyed. The denominator is an estimate of the total area, based on the surveyed plots. The ratio is thus an estimate of density (uncorrected for the detection rate). The rationale for this estimator may be explained as follows. If the mean area of the surveyed zones is larger than the mean area of all zones, then the numerator will tend to exceed the true number of birds present, but the denominator will also tend to be greater than the

true area, so the ratio will tend to be closer to the actual density.

It will be convenient in deriving the variance estimator to express $\hat{\bar{X}}$ as

$$\hat{\bar{X}} = \frac{\hat{\bar{Z}}}{\bar{a}} = \frac{\sum\limits_{u}^{U}\sum\limits_{h}^{H} N_{uh}\hat{\bar{Z}}_{uh}}{\sum\limits_{u}^{U}\sum\limits_{h}^{H} N_{uh}\bar{a}_{uh}} \tag{6}$$

where $\hat{\bar{Z}}_{uh}$ and $\bar{a}_{uh}$ are the means of the $\hat{Z}_{uhi}$ and a_{uhi}, respectively, in stratum u-h (i.e., region u and type h). $\hat{V}(\hat{\bar{X}})$ was estimated using the standard formula for the estimated variance of a ratio of correlated random variables (Cochran 1997),

$$\hat{V}(\hat{\bar{X}}) = \left(\frac{\hat{Z}}{\bar{a}}\right)^2\left(\frac{\hat{V}(\hat{Z})}{\hat{Z}^2} + \frac{\hat{V}(a)}{a^2} - \frac{2C\hat{o}v(\hat{Z},a)}{\hat{Z}a}\right). \tag{7}$$

As noted above, sampling in different strata was not independent. To acknowledge this dependence, let the subscripts g and h indicate type, n_{ugh} = the number of zones in region u in which at least one type g plot and one type h plot were surveyed, and $\hat{\bar{Z}}_{ug|b}$ = the mean of the $\hat{Z}_{ugi}$ among the n_{ugh} plots. The "|b" (for "both") notation means the sum is restricted to zones in which both types g and h were surveyed. Let $\hat{\bar{Z}}_{uh|b}$, $\bar{a}_{ug|b}$ and $\bar{a}_{uh|b}$ be defined in a similar manner. With this notation, and using standard survey sampling methods, it may be shown that

$$\hat{V}(\hat{Z}) = \sum\limits_{u}^{U}\sum\limits_{g}^{H}\sum\limits_{h}^{H} N_{ug}N_{uh}\frac{n_{ugh}}{n_{ug}n_{uh}}\sum\limits_{i}^{n_{ugh}}(\hat{Z}_{ugi} - \hat{\bar{Z}}_{ug|b})$$
$$\times(\hat{Z}_{uhi} - \hat{\bar{Z}}_{uh|b})/(n_{ugh} - 1), \tag{8}$$

$$\hat{V}(a) = \sum\limits_{u}^{U}\sum\limits_{g}^{H}\sum\limits_{h}^{H} N_{ug}N_{uh}\frac{n_{ugh}}{n_{ug}n_{uh}}\sum\limits_{i}^{n_{ugh}}(a_{ugi} - \bar{a}_{ug|b})$$
$$\times(a_{uhi} - \bar{a}_{uh|b})/(n_{ugh} - 1), \tag{9}$$

$$C\hat{o}v(\hat{Z}, a) = \sum\limits_{u}^{U}\sum\limits_{g}^{H}\sum\limits_{h}^{H} N_{ug}N_{uh}\frac{n_{ugh}}{n_{ug}n_{uh}}\sum\limits_{i}^{n_{ugh}}(\hat{Z}_{ugi} - \hat{\bar{Z}}_{ug|b})$$
$$\times(a_{uhi} - \bar{a}_{uh|b})/(n_{ugh} - 1). \tag{10}$$

When $n_{ugh} = 1$, the corresponding term in expressions (8)–(10) cannot be evaluated even though the variance or covariance the term represents does exist. Models could be used to estimate the missing terms (e.g., using the mean of the terms for which $n_{ugh} > 1$), but we have not investigated this approach and therefore omitted terms in expressions (8)–(10) when n_{ugh} was 1.

The detection ratio, R, was estimated as

$$\hat{R} = \frac{\bar{x}}{\bar{y}} \tag{11}$$

where $\bar{x}$ is the mean number of birds recorded on rapid surveys of the intensive plots and $\bar{y}$ is the mean number of birds determined to be present on these plots through intensive surveys.

Let m = the number of camps, $\bar{x}_i$ = the mean number recorded per rapid survey at camp i (>1 rapid survey was made at intensive plots to increase precision), and $\bar{y}_i$ = the mean number actually present at all plots in camp i, then $\bar{x} = \Sigma\bar{x}_i/m$ and $\bar{y} = \Sigma\bar{y}_i/m$. The $\bar{x}_i$ were calculated as the simple means of the means/plot because sometimes plots at a camp were not all surveyed the same number of times by rapid surveyors. Camps were widely distributed across the study area, so they were treated as a simple random sample (rather than a stratified random sample). Under this assumption, the variance of $\hat{R}$ was estimated as

$$\hat{V}(\hat{R}) = \hat{R}^2\left(\frac{\hat{V}(\bar{x})}{\bar{x}^2} + \frac{\hat{V}(\bar{y})}{\bar{y}^2} - \frac{2C\hat{o}v(\bar{x}, \bar{y})}{\bar{x}\bar{y}}\right) \tag{12}$$

where

$$\hat{V}(\bar{x}) = \frac{1}{m}s^2(\bar{x}_i), \hat{V}(\bar{y}) = \frac{1}{m}s^2(\bar{y}_i), C\hat{o}v(\bar{x}, \bar{y}) = \frac{1}{m}cov(\bar{x}_i, \bar{y}_i)$$

and s^2 and $c\hat{o}v$ indicate the sample variances and covariances, respectively.

Estimates will often be needed within a stratum, habitat, or region. The point and interval estimators for such cases are easily derived from the expressions above. Within a single stratum we have simple random sampling. From (6), $\hat{\bar{X}}_{uh} = \hat{\bar{Z}}_{uh}/\bar{a}_{uh}$. The estimated variance has the same structure as expression (7), but $\hat{V}(\hat{\bar{Z}}_{uh}) = s^2(\hat{Z}_{uhi})/n_{uh}$, $\hat{V}(\bar{a}_{uh}) = s^2(a_{uhi})/n_{uh}$. For the mean within one habitat, across >1 region, expressions (6)–(10) apply, except that g and h are equal and constant, so expressions (8)–(10) simplify:

$$\hat{V}(\hat{Z}) = \sum\limits_{u}^{U} N_{uh}^2 s^2(\hat{Z}_{uhi})/n_{uh}, \tag{13}$$

$$\hat{V}(a) = \sum\limits_{u}^{U} N_{uh}^2 s^2(a_{uhi})/n_{uh}, \tag{14}$$

$$C\hat{o}v(\hat{Z}, a) = \sum\limits_{u}^{U} N_{uh}^2 cov(\hat{Z}_{uhi}, a_{uhi})/n_{uh}. \tag{15}$$

The estimate for >1 habitat within a single region is provided directly by expressions (6)–(10), with the sums in u having a single value. For all three

cases, the equations for the detection rate and its variance do not change except that the expressions may have fewer terms.

A comprehensive Windows-based program, DS, was written to carry out all of the calculations described above. DS produces estimated densities and population totals, along with standard measures of precision, by habitat, region, and species. It has many features to facilitate analysis. For example, any set of species can be used to estimate the detection ratios for any species. It is available free from the senior author along with a detailed user's manual.

DETECTION RATIOS

As noted immediately above, program DS allows the user to specify which species will be used in estimating the detection ratio for each species. It might seem that species-specific ratios should be estimated whenever feasible, but how do we define "feasible"? If intensive plots did not contain any pairs of a species, then we obviously cannot obtain a species-specific detection ratio. If the intensive plots only contained one or two pairs, then even though we could obtain a species-specific estimate, it would not be very useful because it would have a huge SE. Furthermore, species identity is only one factor that might affect detection ratios. When we compare densities across regions or habitats, we may be misled if detection ratios differ, but we use a combined rate. This problem is no different from those analysts face with any method for estimating detection ratios. For example, distance methods require 70 or more detections, yet, for many species (especially in specific habitats and regions), the number detected may be far smaller, so data must be combined across species (or regions and habitats) to obtain detection ratios. More generally, this is an example of a common statistical problem: how many "parameters" (in our case, detection ratios) to define. A common approach, and the one we largely follow in this monograph, is to perform a comprehensive test such as an analysis of variance (ANOVA) to determine whether we have good reasons for rejecting the null hypothesis that the rates are equal. Failing to reject the null hypothesis doesn't mean we conclude that the rates are equal; it just means we conclude that the data set is not large enough to justify calculation of separate rates. As a practical matter, if all of the confidence intervals

overlap, then it is very unlikely that the ANOVA and subsequent pairwise tests will support calculation of any species-specific rates. In subsequent chapters of this volume we often report whether any of the detection rates are significantly different and, if they are not (and especially if they are not even close to being significantly different), then we generally use the combined detection ratio. One exception, however, is that if we believe on the basis on biological information that the true detection ratios are quite different, then we often report the densities and population sizes using both the combined ratio and the species-specific ratio.

WHY DOESN'T POPULATION SIZE EQUAL THE SUM OF STRATUM-SPECIFIC ESTIMATES?

Readers will notice that estimated population sizes for two or more regions do not usually equal the sum of the region-specific estimates. Often the two figures are quite different. This is a standard problem in survey sampling that arises when a ratio estimator is used in each of several strata. The presence of a random variable in the denominator causes the estimate to be biased. As a simple example, suppose we were trying to estimate the inverse of mean plot size, as is the case in expression (6). To keep the example simple, suppose plots were actually of size 1 or size 9 (the units do not matter), and that the two sizes were equally common. The mean size is thus 5, so the inverse is 0.2. Now suppose we tried to estimate the inverse with a sample of size 1. We might get a plot of size 1, in which case our estimate of the inverse would be 1, or we might get a plot of size 9 in which case our estimate would be 0.11. Since the two sizes are equally common, the mean of the possible estimates (i.e., its expected value) would be $(1.0 + 0.11)/2 = 0.55$, which is substantially different from the true value, 0.2. If we made the plot sizes 1 and 2, then the difference between the expected value and true value (i.e., the bias) would be much smaller. If we used plot sizes of 1 and 9 but increased the sample size to 2, then the bias would also be smaller. More generally, the bias in a ratio estimate decreases with sample size and when the variance of the random variable in the denominator decreases.

With ratio estimates in stratified sampling, and either small sample sizes or large variance of the

random variable in the denominator, the bias in each stratum can be considerable. If an overall estimate is computed by adding up stratum-specific estimates, the bias also sums and can become quite large when estimates are summed over several strata. To avoid this problem, estimates are usually made by summing the numerators and dividing the result by the sum of the denominators. Thus, we might compute ("sum of birds")/("sum of areas"), known as the "combined estimate," rather than "sum of birds/areas," known as the separate estimate.

With our estimator (e.g., expression 6), the important random variable in the denominator is a_{uhi}, the area covered by all plots in the ith cluster of habitat h plots in region u. This quantity often varies substantially across clusters in a given habitat and region, and this variation would generate substantial bias if we used the separate approach in estimating population sizes. We therefore used the combined approach to reduce the bias, even though this means that the estimated population size for a region with several strata does not equal the sum of the stratum-specific estimates. The combined approach yields somewhat larger SEs than using the separate approach but much smaller bias. We consider the approach "conservative" in that sense.

CONCLUSION

When we began the study we expected to use distance, double-observer, or some other well-developed method for estimating detection rates and obtaining unbiased estimates of population size. We found, however, that some species responded to us strongly (which violates distance assumptions); in some plots the best habitat covered a small proportion of the area, so random selection did not seem feasible (which violates all methods that require random selection of points); and during parts of the season some birds had not arrived or some had already left (which meant that no single-visit survey could produce an accurate count). These problems forced us to adopt double sampling, a method that includes intensive efforts to find all the birds but requires few assumptions and thus is unaffected by the problems mentioned above. As noted in this chapter, the only assumptions are that the sampling plan is followed and that the intensive estimates are unbiased. The one remaining statistical problem was that standard cluster sampling was not very efficient; we needed to have plots in different habitats in close proximity. This necessitated the modification to typical cluster sampling described above in expressions (8)–(10). Use of double sampling also allowed us to avoid the use of index methods, which, in turn, means that the surveys can be modified as new analytic and field methods appear. We believe this is a major strength of the approach and will be appreciated by those who design the survey in the coming decades.

ONLINE CONTENT

Abstracts are available in French and Spanish from www.ucpress.edu/go/sab.
Une traduction du résumé est disponible en français.
Una traducción del resumen está disponible en español.