Effects of Sampling Strategy, Detection Probability, and Independence of Counts on the Use of Point Counts[1]

Grey W. Pendleton[2]

Abstract: Many factors affect the use of point counts for monitoring bird populations, including sampling strategies, variation in detection rates, and independence of sample points. The most commonly used sampling plans are stratified sampling, cluster sampling, and systematic sampling. Each of these might be most useful for different objectives or field situations. Variation in detection probabilities and lack of independence among sample points can bias estimates and measures of precision. All of these factors should be considered when using point count methods.

Sampling strategies, variable detection probabilities, and independence among counts are aspects of point count methodology often overlooked but require emphasis when planning point count projects or analyzing point count data. At least as much planning should be focused on sampling schemes and potential sources of bias as on survey logistics.

The first consideration in planning a project with point counts is to explicitly state the objective. The three major objectives of point counts are to: (1) monitor trends, (2) assess habitat relationships, and (3) map bird distributions. Different sampling strategies best address each of these objectives; it might be impossible to design a single sampling plan that will provide data to do all three effectively. The effects of variable detection rates (e.g., birds are more easily detected at some time or place than at other times or places) and spatial correlation (counts at points close together are more similar than counts far apart even if in the same habitat) also differ among these objectives.

Sampling Strategies

Two aspects of sampling, the sample universe or frame and the sampling scheme, affect selection of the points to be sampled. The sample universe determines the area where samples may be located; this is also the area to which estimates or conclusions apply. The sampling scheme determines how the sample points will be chosen within the sample universe. Selection of the sample universe and the sampling scheme varies depending on the objective of the survey.

The sample universe for monitoring trends or mapping should be all areas where a species of interest is found within the overall study area. Studies assessing habitat relationships often sample exclusively in habitat blocks large enough to reduce the effect of neighboring habitats; mosaics of small habitat patches and edges are often avoided. Other studies might focus on bird abundance only in specific habitat types.

These examples illustrate the importance of selecting an appropriate sampling universe for the desired objective and the difficulty in trying to achieve combinations of objectives. Sampling only in specific habitats or avoiding habitat mosaics results in estimates and conclusions that apply only to the habitats or even habitat blocks of the size actually sampled. This sample would meet the objective of some habitat studies where only habitat-specific estimates are required. However, if the unsampled area is at all sizable, maps of bird abundance produced from these data could be misleading (e.g., when a species is abundant in an unsampled habitat). Also, if a species of interest occurs in the unsampled areas, overall trend estimates will be biased if the species' trend in the unsampled habitat is different from the trend in the sampled area. Habitat- or sample-specific estimates of trend could be produced, but these would be difficult to interpret.

Once the sample universe has been determined, a sampling scheme appropriate for the study objective can be selected. Simple random, stratified, cluster, systematic, or purposeful sampling are each appropriate in some studies. Completely random sampling is rarely used in point count studies for both theoretical and logistical reasons, but obtaining unbiased estimates requires some form of randomization. Purposeful sampling rarely is appropriate, but is sometimes used in some mapping studies to ensure that samples include transitions between areas with differing bird abundance.

The three most commonly used sampling schemes are stratified, cluster, and systematic (with a random start).

A stratified sample is one in which the sample universe is divided into groups of sample units (strata) that have more homogeneous bird abundance than the sample universe as a whole; for example, habitat types are often used as strata. Sample points are then randomly located within strata. Stratified samples reduce the variance of estimates when counts within strata are more similar than counts between strata. Strata need not have an equal number of samples, but weighted estimates might be needed for some unequal sample allocations (Cochran 1977). Stratum weights need to be known or estimated to obtain appropriate estimates and variances.

A cluster sample is one in which larger, primary sampling units are chosen (usually at random), then samples (i.e., point counts) are allocated within the primary sampling units (Cochran 1977). This sampling strategy is most useful when travel time between sample points is long and, therefore, simple random sampling is inefficient. However, under commonly encountered situations, the variance of an estimate based on a cluster sample is substantially larger than one based

[1] An abbreviated version of this paper was presented at the Workshop on Monitoring Bird Population Trends by Point Counts, November 6-7, 1991, Beltsville, Maryland.

[2] Biological Statistician, Patuxent Wildlife Research Center, USDI National Biological Service, Laurel, MD 20708

on a simple random sample. Effective use of cluster sampling can reduce travel time, allowing an increased sample size that overcomes the increased variance. Unlike stratified sampling, cluster sampling works best when the within-cluster variation is large and the between-cluster variation is small. Stratified or cluster sampling requires more complex estimators of means and variances than simple random sampling (Cochran 1977). Both stratified and cluster sampling could be used for all objectives of point count projects.

Systematic samples are those where a random starting point is selected and subsequent samples are located at uniform intervals. This results in samples uniformly distributed over the area of interest (uniformly spaced samples that cover only a subset of the area of interest would not be expected to give reliable results). With systematic samples, no unbiased estimates of variance are possible (see Independence of Counts below) (Sukatme and others 1984). However, under some conditions, estimates from systematic samples will be more precise than comparable random samples (Kingsley and Smith 1981, Sukatme and others 1984). Systematic samples perform poorly when populations have periodic fluctuations or clumped distributions. But, with careful planning of spacing between sample points, systematic samples can be useful because larger samples can be obtained because of the relative ease of locating sample points in the field. Mapping bird population distribution, which often does not require independent points or variance estimates, is best accomplished with uniform spacing of sample points, possibly with higher density strata in areas of particular interest.

Choice of sample universe and sampling scheme are important factors in designing a study using point counts. Equally important, once the universe and scheme are chosen, the appropriate estimation procedures should be used (i.e., based on stratified or cluster sampling).

Variation in Detection Probabilities

Detection probability is the probability of recording a bird's presence if the bird is at the point when the count is made. It has long been recognized as a. problem in wildlife surveys, including point counts (Lancia and others, 1994). Its potential effects should be seriously considered (Barker and Sauer 1992b). Two strategies have been used to reduce the effects of variable detection probabilities. The first is standardization of survey methods and the conditions when the survey is conducted, and the second is estimation of the detection probability, which is used to adjust the counts to get a population estimate (Lancia and others 1994). Standardized methods eliminate the effects of variable detection probabilities if they result in a constant fraction of the population in the count area being counted (e.g., exactly 57 percent of the animals are seen in all counts). This is not the same as having a constant detection probability (e.g., each animal has a 45 percent chance of being recorded on any given survey). The distinction between counting a constant fraction and having a constant detection probability is important because the second adds an additional source of variability (Barker and Sauer 1992a, 1992b). This additional variability occurs because, even though the detection probability is constant, the actual proportion detected is not exactly the same each time, in the same way that tossing a coin does not always result in exactly 50 percent heads and 50 percent tails.

Standardization of methods is encouraged, but there are weaknesses in data from standardized counts. It is unrealistic to think that all factors affecting detection probabilities can be controlled. One hopes that remaining variation in detection probabilities, after controlling as many factors as is feasible, is small relative to actual differences in abundance. However, because detection rates are not estimated, this cannot be evaluated.

The second approach estimates detection rates and adjusts counts for them (Lancia and others, 1994). These methods, including capture-recapture and variable circular plots, provide better information but are much more expensive than procedures that rely on standardized counts. For extensive surveys, these methods are usually not practical because of logistical restraints.

Both geographically and temporally variable detection probabilities can affect all uses of point count data. Estimation of population trends is more sensitive to detection probabilities changing over time (both within and among years). Also, interactions between geographically variable detection probabilities and actual population changes could mimic temporally changing detection rates. Analysis of habitat relationships and maps of bird distribution are sensitive to geographic changes in detection rates. If detection rates are unequal, an abundance map might show areas of high and low abundance that are actually areas of high and low detectability. But, combining data from different time periods (e.g., data from some habitat types in the spring and other habitat types in the summer) when there is a temporal trend in detection rates could bias resulting conclusions. For example, if detection rates were lower in the summer, a conclusion might be reached that a species was more abundant in the habitat sampled in the spring, when actually the species was equally abundant in both habitats.

Many factors affect detection probabilities, including differences among observers, annual variation in phenology, and weather. One major source of variation in detection probabilities with particular importance to many point count studies is differences in detectability among habitats. For example, clearcuts would likely have much different detection rates than neighboring forests. However, most uses of point counts involve some combination of data from different habitats. Mapping bird distributions with data from several habitats with habitat-specific detection probabilities will cause distortion of the distribution maps (Sauer and others, in these Proceedings). Bird habitat associations derived from point counts can also be biased by habitat-specific detections. Rather than ranking habitats by bird abundance, they would be ranked by the product of abundance and detectability, which could produce an entirely different pattern. Route-regression type trend estimators (Geissler and Sauer 1990) combine trends from individual routes that are often in more than one habitat. Trends of individual routes are weighted on the basis of the size of the counts from that route, which could also be distorted by habitat-specific variation of detection rates.

Independence of Counts

Independence among point counts is one factor often considered when proposing standardized point count methodology. However, there are two types of independence between counts that are important-conditional and unconditional independence. The number of birds counted at a point can be considered to be a function of the actual number of birds at a count location plus random error [$y_i = f(x_i) + e_i$, where y_i is the observed count, x_i is the actual number of birds, and e_i is the random variation].Conditional independence relates to whether a count being above or below the average value for that point is affected by whether neighboring points are above or below their averages [$\mathrm{cov}\{e_i, e_j | f(x_i), f(x_j)\}$]. This is a small-scale type of dependence that would likely include factors such as counting the same birds at successive points or having the calling rate of birds at a point affected by calling birds at a previously counted point. This is the type of dependence that the frequently suggested spacing between point counts (e.g., 100 m, 250 m, 500 m) is intended to reduce, although no empirical data are available to support selection of an appropriate distance.

The second type of independence, unconditional independence, is less often considered and relates to whether points close together have actual abundances more similar than points farther apart [$\mathrm{cov}\{f(x_i) f(x_j)\}$]. Unconditional dependence is probably related to the size of the area of interest, but would likely occur over larger geographic scales than are important for conditional dependence. For example, a sample point having a Scarlet Tanager (*Piranga olivacea*) might reveal nothing about the probability of a second point 5 miles away having a Scarlet Tanager if the area of interest is an eastern National Park. If the area of interest is the continental United States, however, then the tanager at the first point might provide substantial information about the second point's probability of also having a Scarlet Tanager.

If locations of point counts are randomly selected and measurement error is small relative to sampling variation, point counts will be independent regardless of any underlying spatial relationships (de Gruijter and ter Braak 1990). This is a large advantage for some form of random sampling over nonrandom sampling strategies. If random sampling is not used and points are close enough together so that there is unconditional dependence, variance estimates will be too small and power associated with statistical tests is artificially inflated (Sukatme and others 1984, Whysong and Miller 1987). Spatial dependence could affect statistical comparisons of abundance between areas or habitats and significance tests associated with trends. The distance between points to achieve unconditional independence would have to be estimated separately for each area of interest using methods such as variograms (Isaaks and Srivastava 1989). Many mapping procedures are unaffected by spatial dependence and some actually use this information. Some geostatistical methods (i.e., variograms) are useful for detecting spatial dependence and estimating the distance needed between sample points to obtain independence.

Conclusion

All of these factors, choice of sample universe and sampling scheme, variable detection probabilities, and independence among counts, can have substantial effects on estimates and conclusions based on point count data. Different objectives require different choices of sampling procedure and are affected differently by these factors. Adjustments cannot be made for most biases introduced by these factors when analyzing point count data. Point counts, however, are the only practical way to obtain data for many species. These problems and factors should be carefully considered when planning a project that uses point counts. Efforts should be made to investigate and reduce aspects of sampling that lead to biased estimates.

Acknowledgments

Helpful comments on this manuscript were provided by C. M. Bunck, Deanna K. Dawson, B. G. Peterjohn, John R. Sauer, and D. Smith.