Experimental Design Considerations for Establishing an Off-Road, Habitat-Specific Bird Monitoring Program Using Point-Counts[1]

JoAnn M. Hanowski and Gerald J. Niemi[2]

Abstract: We established bird monitoring programs in two regions of Minnesota: the Chippewa National Forest and the Superior National Forest. The experimental design defined forest cover types as strata in which samples of forest stands were randomly selected. Subsamples (3 point counts) were placed in each stand to maximize field effort and to assess within-stand and between-stand variation for a variety of bird parameters. Data gathered in 1991 were used to evaluate several assumptions that were made in the experimental design and showed that variance of most bird parameters among strata were similar. This data indicated that a proportional stratified sample by forest cover type was reasonable. We also found that two subsamples per stand would be optimum when a variety of strata types and bird variables were considered. Analyses based on 120 stands in the Chippewa National Forest and 150 stands in the Superior National Forest indicated that a two-tailed *t*-test could detect a 25 percent change in bird numbers for common species. For most other species, we could detect less than a 50 percent annual change. Monitoring programs within regions should be habitat specific so that changes in bird numbers can be related to: (1) habitat changes that have occurred in the region; (2) natural fluctuations in bird numbers; or (3) other factors.

We recently suggested guidelines for determining the number of samples and size of study areas required for monitoring bird populations using line-transects (Hanowski and others 1990). We present a similar statistical approach to suggest an experimental design for establishing bird monitoring programs using point counts. An experimental design for determining monitoring programs in a region will be influenced by the objectives and resources available for each region. We assumed that the primary objectives of a monitoring program would be to: (1) monitor relative abundance of common bird species to assess annual changes, (2) define avian habitat relationships, (3) determine how forest management activities influence breeding bird abundance and distribution, and (4) provide a product that a regional wildlife biologist could use to plan forest management activities to accommodate a variety of bird species, especially those with specific habitat needs or declining populations in a region.

Our objectives here are to: (1) describe in detail the methods that we used to establish a habitat-specific bird monitoring program in two regions of northern Minnesota, (2) present results of statistical aspects of the experimental design in terms of sample stratification and allocation of samples and subsamples, and (3) describe an observer training and testing program that will provide quality assurance in the data collection.

[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] Research Fellow and Director, respectively, Center for Water and the Environment, Natural Resources Research Institute, University of Minnesota, Duluth, 5013 Miller Trunk Highway, Duluth, MN 55811

Study Areas and Methods

We established bird monitoring programs in two areas in Minnesota. The Chippewa National Forest is located in the north central portion of the State, and the Superior National Forest in the northeastern region. Major habitat types are similar within each Forest with the exception that upland spruce-fir (*Picea* sp. and *Abies balsamea*) forests are more common in the Superior National Forest.

Experimental Design

We designed our monitoring program so that it would integrate with each National Forest's method of describing vegetation cover types. Each unit or stand within each Forest is described by a forest cover type and age class code. With this method, bird census data can be directly linked to both the Forest Service inventory of cover types, total area of each cover type, and locations within the Forest. We used groups of cover types as our strata for sampling (*fig. 1*).

Our sample unit in the design was a forest stand that was ≥ 40 acres, the minimum size needed for three subsamples (point count). We subsampled stands for several reasons. First, we wanted to obtain measures of both within and between stand variation. This information would allow us to determine the optimal allocation of samples based on effort (see below). Second, subsamples would allow us to sum numbers from counts within a stand for species and thus would provide a better estimate of mean individuals per unit area. Subsequently, this sum would allow us to better meet assumptions for statistical tests (e.g., normality and homogeneity of variance) and the overall power of our statistical tests would be improved with larger means (Hanowski and others 1990).

This sample design was tailored to meet specific objectives of an individual Forest. With this particular method, small stands and habitats that are not managed for timber are not represented in the sample. The basic sample design can be modified to address specific questions, however, on a regional basis. For example, we have subsequently initiated a similar monitoring program in the Chequamegon National Forest in Wisconsin and have included upland and lowland shrub habitats as well as small wildlife openings in the sample.

Sample Stratification

The number of samples within each Forest was determined *a priori* by calculating the number required to detect about a 20 percent annual change in a common birds species abundance (Hanowski and others 1990). The first step in the sample stratification procedure was to determine the number of strata to sample. For our study, we used two different approaches, each to accommodate criteria established by each

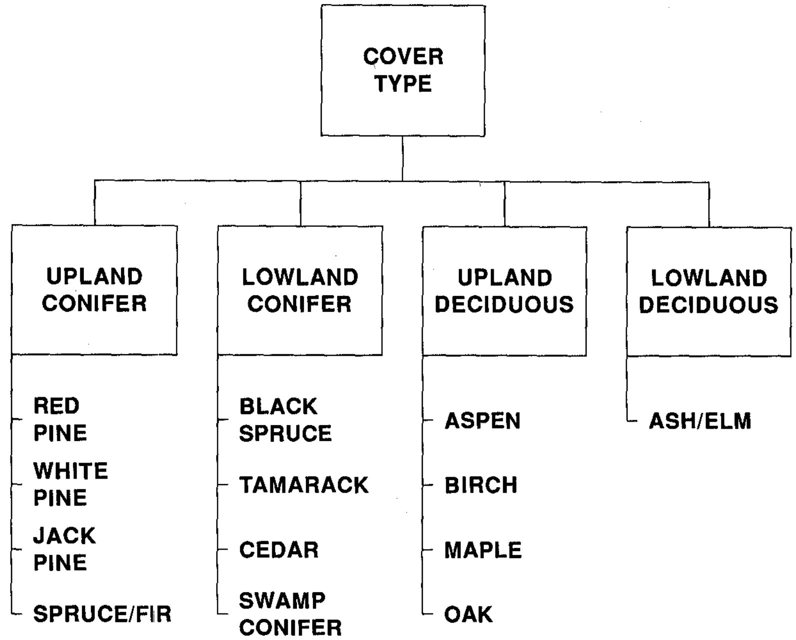

Figure 1--Four major strata (e.g., upland conifer) and cover types (e.g., red pine) sampled within the Chippewa and Superior National Forests in Minnesota.

Forest. We identified five habitat groups as strata in the Chippewa National Forest and four habitat groups or strata in the Superior National Forest. These included the types listed in *figure 1*, plus a regenerating type that included both upland deciduous and upland conifer types. Cover types selected represented areas where forest management activities are conducted. No lowland deciduous cover types were included in the Superior National Forest because no timber harvesting is done in these areas. Because the focus of the monitoring was forest birds, nonforested wetland habitats were excluded.

The next step was to determine; how many samples were required within each stratum. Because we had no estimate of the variance for point count data in Minnesota, we used data gathered previously in northern Wisconsin and Michigan. We found that the variances of several bird parameters in different habitat strata were similar (Hanowski and others 1990). Consequently, samples were allocated proportionally rather than optimally to strata.

We summed the number of acres and number of stands >40 acres and calculated the proportion of the total within each stratum. We could have stratified on the basis of total acres or by total number of stands. However, the proportion of samples that would be allocated to each stratum with either method was similar. We decided to use the number of stands in the Forest >40 acres to calculate the proportions of samples allocated within each stratum. This was done because the stand was considered the experimental unit for this study.

Another item that needs to be considered in the experimental design is the minimum number of samples required to provide biological information necessary for statistical analyses within each stratum. It was evident from the habitat breakdown in the Chippewa National Forest that there was a disproportionate amount of upland deciduous habitat in the Forest (approximately 45 percent). If we had stratified strictly on the proportion of habitat available, the majority of the samples would have been allocated to upland deciduous forest. The remaining habitats, however, would have been inadequately sampled. Therefore, we modified the stratification so that at least five samples were placed within each stratum.

Sample Unit and Selection

Forests are mapped in large management units (compartments). To select a stand, a compartment was first selected with a random number table, and stands within the compartment were chosen randomly. Because of travel time and, hence, cost to travel between stands and compartments, four or five stands were randomly selected within each compartment with the restriction that all could be sampled by one observer in one morning (between 0445 and 0930 c.d.t.). Stands selected within each compartment generally represented two or three different cover types. Other restrictions for stand selection were road access and physical barriers (e.g., large rivers and bodies of water could not be crossed). A total of 140 stands (420 subsamples) were selected within the Chippewa National Forest and 150 stands (450 subsamples) within the Superior National Forest. With this method, each stand is not randomly selected, and therefore the design is a cluster sample.

Bird Counts

All stands were located and count points were marked on compartment maps and aerial photos. Each stand was visited before counting to permanently mark locations and routes of travel between points. We conducted one bird count (10 minutes in duration) at each point during the breeding season (Reynolds and others 1980). Point counts are excellent for determining relative abundances of singing passerines but are inadequate for raptors, waterfowl, and other wide-ranging species. In addition, because only one count is conducted in the breeding season (June to early July in northern Minnesota), relative densities of early nesting species are probably underestimated (e.g., most permanent residents including woodpeckers and chickadees).

Six trained (see observer training section below) observers conducted the censuses, which were done from 0.5 hours to 4 hours after sunrise. Censuses were conducted only during good weather (i.e., wind <15 mph and no precipitation). Types of stands censused (forest cover type) were stratified by time of morning. For example, we avoided sampling all upland pine stands early or late in the morning. Forest cover types censused also were stratified by observer; each observer sampled essentially the same number of stands in each stratum.

We recorded weather (cloud cover, temperature, and windspeed) and time of day the census was conducted. All birds heard or seen from the center point were recorded in a circle with estimates of their distance from the center point (up to 100 m).

Observer Training

Four of the six observers in this study had conducted point counts previously but had not been specifically trained in the identification of northern Minnesota breeding birds or with the methodology used in this study. Observers were hired in April and were given a list of species that they were required to identify by sight and sound. Tapes of bird songs were provided as a learning tool for all observers. All observers were required to pass an identification test of 75 bird songs made by Cornell Laboratory of Ornithology Library of Natural Sounds. A standard for the number of correct responses was established by giving the test to experienced observers (>4 years) in the field identification of Minnesota's forest birds by sound. This was done to identify songs on the tape that were not good representations of songs heard in northern Minnesota. Based on results of trained observers, we set the standard for passing at 85 percent correct responses. Songs on the tape were grouped by habitat (e.g., upland deciduous) to simulate field cues that would aid in song identification.

Observer field training was done in late May. Observers were first instructed on the methods for recording data on the field sheets. Observers then conducted simultaneous counts (4 mornings; 40 points) and were allowed to ask questions about unknown birds after each count. Count information was compiled for each observer and their data were compared to data gathered by the experienced observers. Species lists and number of individuals recorded on the count by each observer were compared. Deviations from the average number of individuals observed or species missed were noted on the field sheets and returned to each observer.

In addition to training and testing, all observers were required to have a hearing test to ensure that their hearing was within normal ranges for all frequencies (125 to 8000 hertz). Normal ranges were standards established by audiologists.

Statistical Considerations

We made two assumptions in the sample allocation for monitoring. First, we assumed the variance of counts measured within each stratum was equal (based on data from Michigan and Wisconsin). Therefore, we stratified our sample strictly on proportion of stands within each stratum in each Forest. We examined data collected in 1991 to determine whether this assumption was valid. For these analyses we computed standard deviations of estimates for several bird community, bird guild, and species in strata where they occurred within each Forest. We were interested in determining whether the standard deviation in any one stratum for any variable was a factor of 2 higher than the standard deviation in any other strata. If standard deviations are within a factor of 2, a strict proportional stratification can be employed in a study (Kish 1965).

The second assumption was that three subsamples/sample would be optimal in terms of effort. We tested whether this assumption was valid by computing the components of variance among stands (samples) and among counts (subsamples) (SAS 1988, PROC VARCOMP). We then computed the optimum number of subsamples by assigning a cost to collecting a sample and subsample in the formula:

$$n_2 = \sqrt{\left(\frac{C_1 S_2^2}{C_2 S_1^2}\right)}$$

where n_2 = optimum number of subsamples; C_1 = cost of samples; C_2 = cost of subsample; S_1 = variance of sample; and S_2 = variance of subsample. For our calculations we assigned C_1 = 1.0 and C_2 = 0.4 (Snedecor and Cochran 1967: 532). Cost of sample values were estimated on the basis of our experience in the first year of the census.

We calculated an optimum number of subsamples for the total number of individuals, long-distance Neotropical migrants, and Ovenbirds (*Seiurus aurocapillus*) for all strata within the Chippewa and Superior (each Forest separately). We combined all strata for these analyses because we wanted to provide an overall recommendation for the number of subsamples/sample for a regional monitoring program. We realize that each stratum may provide a unique optimum number. However, because strata are not always comparable among regions, an average value would be of more use overall. In addition, if a monitoring program is being set up, it would be unrealistic (primarily for statistical analyses considerations) to place different numbers of subsamples within samples. In presenting these results, we assume that a monitoring program that uses subsamples is more cost efficient and would increase the overall power of statistical tests (see Experimental Design above).

In addition to testing the assumptions that we made in establishing the experimental design, we calculated the power of statistical analyses using means and variances of data collected in 1991. We used the formula presented by Lehmann and D'Abrera (1975: 78) to calculate the power of a two-tailed *t*-test (α = 0.05) for detecting annual differences in species abundances for two levels: a 25 percent and a 50 percent change in the mean per 40 acres. This was calculated for species that occurred in densities within each Forest of two to eight individuals per 40 acres and included about 90 percent of all species detected.

Density Calculations

Although point counts generally are used to assess "relative abundance" of birds, we calculated relative density values per unit area for each species. This information was used to determine baseline relative populations for species in a region. We calculated the number of territorial males in 40 acres by summing numbers of individuals for each species in three point counts within each stand. We determined the area of each sample (point count) on the basis of a radius of 100 m for most species. This was the distance that we used in our data collection. Although some birds could be heard beyond 100 m, we did not count them (Howe and others, in these Proceedings). We did this primarily because we did not want to count the same individual on adjacent points (our points were 250 m apart). We used a smaller radius for Cape May Warbler (*Dendroica tigrina*), Golden-crowned Kinglet (*Regulus satrapa*), and Bay-breasted 'Warbler (*Dendroica castenea*), because we were not confident that we could detect these species beyond 75 m. A relative density value for each bird in the forest can be calculated by multiplying the density value of a species within each habitat by the total amount of that habitat in the Forest.

The relative density calculations should be used with caution. They are not meant to be an absolute density value for the Forest. Rather, they should be viewed as base values with which future monitoring data can be compared to determine whether species populations are going up, down, or remaining the same. More importantly, as these data become coupled with forest change, they will allow an approximate measure of the effects of these changes.

Table 1--***Standard deviations for long-distance migrants and Ovenbirds within forest cover type and age class. Ovenbirds did not occur in lowland conifer types. Calculations were done on values from stands (e.g., sum of three subsamples/stand)***

Forest cover types	Age class	Long-distance migrants	Ovenbirds
Upland deciduous	Regenerating	1.4	0.8
	Pole size	1.7	0.9
	Saw size	2.3	0.9
Upland conifer	Regenerating	2.1	0.8
	Pole size	1.8	0.8
	Saw size	1.7	0.6
Lowland conifer	Regenerating	1.5	-
	Pole size	1.3	-
	Saw size	2.1	-

Results

Allocation of Samples to Strata

An estimate of the variance for bird species within different cover types is required to calculate the optimal allocation of samples to stand (equation 1). Standard deviations for numbers of long-distance migrants (mean of Chippewa and Superior National Forests) within nine cover types ranged from 1.3 to 2.3 (*table 1*). The range of standard deviations for numbers of Ovenbirds within six cover types ranged from 0.6 to 0.9 (*table 1*). We did not include standard deviations for the lowland conifer type for the Ovenbird, because it occurs in these types only occasionally.

Allocation of Subsamples to Samples

Components of variance between counts (subsamples) and between stands (samples) for three bird parameters were similar between the two Forests (*table 2*). The optimum number of counts per stand calculated for three bird parameters (total number of individuals, number of long-distance migrants, and number of Ovenbirds) indicated that between 1.5 and 2.0 counts per stand would optimize the effort involved in the sampling.

Table 2--***Between-stands (samples) and between-counts (subsamples) components of variance, and optimum number of counts per stand for three bird parameters in the Chippewa and Superior National Forests. See Methods section for details of equation used to calculate optimum number of counts per stand***

Bird parameter	Forest	Between stands	Between counts	Optimum number counts/stand
Total individuals	Chippewa	5.44	5.89	1.6
	Superior	5.38	4.75	1.5
Long-distance migrants	Chippewa	3.03	3.81	2.0
	Superior	2.49	2.95	2.0
Ovenbirds	Chippewa	0.51	0.48	1.5
	Superior	0.29	0.55	2.0

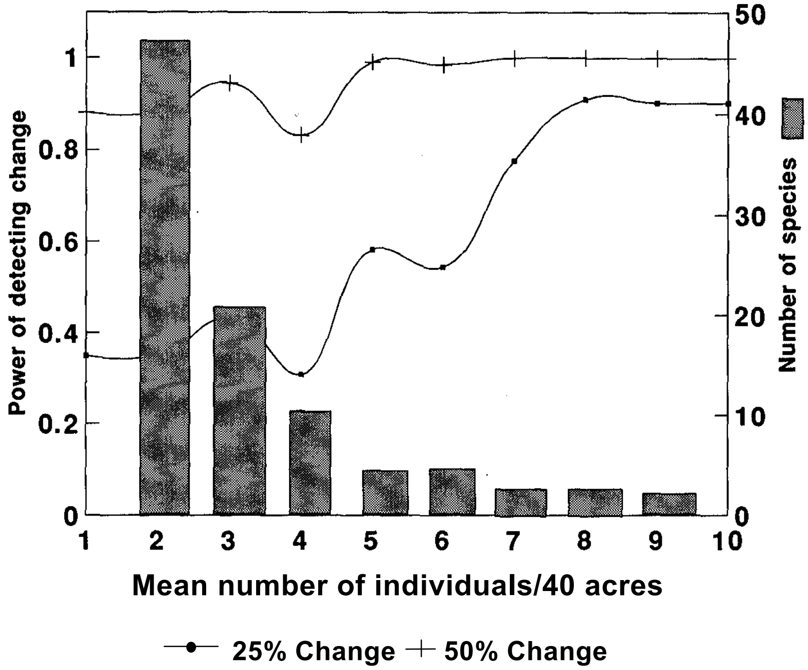

Figure 2--Power of a two-tailed *t*-test for detecting a 25 percent or 50 percent annual change in individual species that occurs within the Chippewa National Forest in densities from 2 to 10 per 40 acres. Number of species that occur at different densities is indicated on the Y2 axis.

Power of Statistical Analyses

Power analyses (two-tailed *t*-test) indicated that we would be able to detect a 25 percent annual change for those species that occurred with a relative density greater than seven individuals per 40 acres in the Chippewa National Forest (*fig. 2*). However, only two species, the Ovenbird and Red-eyed Vireo (*Vireo olivaceus*), occurred in the Forest at those densities (*fig. 2*). The power of detecting a 50 percent annual change, however, was greater than the standard (80 percent) used in most statistical tests for all species (*fig. 2*). This indicates that we would be able to detect a 50 percent or less annual change for many species that occur in the Forest. Results were slightly different for the Superior National Forest. In this region we should be able to detect a 25 percent annual change for six species that occur in densities greater than or equal to five pairs per 40 acres.

Observer Training

All new observers (four total) passed the song test on the first attempt (85 to 100 percent correct responses). Although we did not quantitatively examine results of the observer training sessions, three of the four observers recorded numbers of individuals and species that were similar to the experienced observers. One new observer at first tended to record fewer individuals than the other observers. This was brought to the attention of the observer and in later sessions the discrepancy was minimized. In addition, results of the hearing tests indicated that all observers had hearing within the normal ranges for all frequencies with one exception; one individual had a slightly lower threshold of detection at the highest frequency.

Discussion

Experimental Design

We have presented an approach for an experimental design that can be used to establish regional monitoring programs using point counts. It is critical that some aspects of the design be met in establishing a program, while others can be tailored to meet the objectives for the region. Number of strata sampled and allocation of samples to strata can be modified for each region. In our monitoring, we allocated 50 percent of our total sample to lowland and upland conifer in the Chippewa National Forest because the wildlife biologist wanted more information on bird indicator species that occur in conifer forests in the Chippewa. In the Superior National Forest, we concentrated the monitoring efforts in cover types that are managed for timber production. Some attention needs to be given to the variance within each stratum, although it is not a critical component of how allocation of samples to strata. Strata with higher variances will require relatively more samples (Kish 1965).

The number of subsamples within samples can also be modified to meet objectives and resources available within a region. Although we placed three subsamples within each sample, our *a posteriori* analyses indicated that about two subsamples would optimize the effort of field sampling. We stress that this is an approximate figure and would change depending upon the relative cost of collecting a sample and a subsample. It would also vary depending on the relative homogeneity (in terms of variance) of the stratum that is being sampled (*equation 1*).

There are at least two points that must be considered in all experimental designs. First is the manner of how samples and subsamples are treated in data analyses. The sample unit for our design is a stand and, therefore, analyses of counts within stands are not valid (e.g., the pseudoreplication of Hurlbert 1984). Second, it is critical that at some level the sampling (stands in this design) be random or most statistical tests will be invalid. The assumption of independence of errors is the only one in most statistical methods for which violation is both serious and impossible to cure after the data have been collected (Green 1979).

Bird Habitat Relationships

A major objective of our monitoring program was to relate bird numbers to forest cover types. Such information is required for any monitoring program that intends to relate annual variation of bird numbers to change in forests on the breeding grounds and to management practices of the region. For example, in our monitoring program we can calculate annual change in bird abundance for each species on the basis of data collected at all points. This information can then be linked to forest inventory data to estimate the number of birds in the forest. Therefore, annual changes in bird abundance can be attributed independently to either changes in abundance (based on points) or to changes in the amount of suitable habitat in the forest.

Another advantage of selecting samples based on forest cover strata in a region is that the samples can be linked to a regional data base and to a geographic information system (if present). This link can be a powerful tool for analyzing spatial patterns of bird distribution, identifying source and sink habitats for individuals species, and for determining on a gross scale the relative number of birds in the region. These monitoring goals are best designed for each region, and some comparisons (e.g., density) are not entirely comparable across regions because of differences in habitats and bird census methods. However, with our proposed methods, the count data could be compared across regions if similar methods are used.

Power of Statistical Analyses

The power of statistical analyses for detecting annual differences in numbers for individual bird species with this design and sample size was quite good in comparison to values we calculated in Wisconsin and Michigan (Hanowski and others 1990). This was primarily due to the larger sample that was gathered in the present study. Detecting annual changes of 25 percent to 50 percent in bird numbers on a species level in a region is reasonable with this design using a sample of 120 to 150 stands in each region. The power was somewhat higher in the Superior National Forest because the sample size was larger for that region.

More specific monitoring programs may be needed for rare species with very specific habitat requirements (Verner 1985). We are exploring how this monitoring program works for all forest species in a region and what modifications need to be made to deal with species that are not adequately covered with this method.

Observer Training

The quality of data collected in any monitoring program can be improved if observers are trained and then tested in their ability to identify regional bird songs. It is also helpful if data-recording training sessions are used to familiarize new observers with methods and regional dialects. Hearing tests document the levels (decibels) at which individuals can detect a variety of frequencies. However, although observer training and testing are necessary, they do not eliminate observer variation from the sample. Observers differ in their estimate of singing birds especially when a judgment needs to be made regarding multiple cues from the same species at a point. Because some sources of observer bias cannot be controlled, we recommend that observers be distributed over the strata that are being sampled. For example, to ensure that observer variation is distributed evenly across strata all observers should sample relatively the same proportion of strata in the sample.