A Sampling Design Framework for Monitoring Secretive Marshbirds

Authors: Johnson, Douglas H., Gibbs, James P., Herzog, Mark, Lor, Socheata, Niemuth, Neal D., et al.

Source: Waterbirds, 32(2) : 203-215

Published By: The Waterbird Society

URL: https://doi.org/10.1675/063.032.0201

WATERBIRDS

JOURNAL OF THE WATERBIRD SOCIETY

VOL. 32, NO. 2 2009 PAGES 203-362

A Sampling Design Framework for Monitoring Secretive Marshbirds

DOUGLAS H. JOHNSON[1*], JAMES P. GIBBS[2], MARK HERZOG[3], SOCHEATA LOR[4], NEAL D. NIEMUTH[5], CHRISTINE A. RIBIC[6], MARK SEAMANS[7], TERRY L. SHAFFER[8], W. GREGORY SHRIVER[9], STEPHEN V. STEHMAN[2] AND WILLIAM L. THOMPSON[10]

[1]U.S. Geological Survey, Northern Prairie Wildlife Research Center, Saint Paul, MN, 55108, USA

*Corresponding author; E-mail: Douglas_H_Johnson@usgs.gov

[2]State University of New York College of Environmental Science and Forestry, Syracuse, NY, 13210, USA

[3]PRBO Conservation Science, Petaluma, CA, 94954, USA

[4]U.S. Fish and Wildlife Service, Biological Monitoring Team, 2630 Fanta Reed Road, La Crosse, WI, 54603, USA

[5]U.S. Fish and Wildlife Service, Habitat and Population Evaluation Team, Bismarck, ND, 58501, USA

[6]U.S. Geological Survey, Wisconsin Cooperative Wildlife Research Unit, Department of Forest and Wildlife Ecology, University of Wisconsin, Madison, WI, 53706, USA

[7]U.S. Fish and Wildlife Service, Division of Migratory Birds, Laurel, MD, 20708, USA

[8]U.S. Geological Survey, Northern Prairie Wildlife Research Center, Jamestown, ND, 58401, USA

[9]Department of Entomology and Wildlife Ecology, University of Delaware, Newark, DE, 19717, USA

[10]National Park Service, Southwest Alaska Inventory and Monitoring Network, Anchorage, AK, 99501, USA

Abstract.—A framework for a sampling plan for monitoring marshbird populations in the contiguous 48 states is proposed here. The sampling universe is the breeding habitat (i.e. wetlands) potentially used by marshbirds. Selection protocols would be implemented within each of large geographical strata, such as Bird Conservation Regions. Site selection will be done using a two-stage cluster sample. Primary sampling units (PSUs) would be land areas, such as legal townships, and would be selected by a procedure such as systematic sampling. Secondary sampling units (SSUs) will be wetlands or portions of wetlands in the PSUs. SSUs will be selected by a randomized spatially balanced procedure. For analysis, the use of a variety of methods as a means of increasing confidence in conclusions that may be reached is encouraged. Additional effort will be required to work out details and implement the plan. *Received October 20 2008, accepted April 25 2009.*

Key words.—marshbirds, monitoring, population, sampling design, survey.

Waterbirds 32(2): 203-215, 2009

The design, implementation and reporting of population monitoring programs provide the foundation for much of our understanding of wildlife population trends, responses to management actions, and, more recently, phenological shifts associated with climate change (e.g., Stenseth and Mysterud 2002). In North America, the Breeding Bird Survey (BBS; Robbins *et al.* 1986) and the annual waterfowl surveys (Smith 1995) are the most comprehensive, long-term and large-scale avian monitoring programs in existence. Information from these programs has provided a basis for recommendations for harvest management, identified significant population changes, and heightened our awareness of the population declines in many migratory landbirds (e.g., Robbins *et al.* 1989). Despite these programs, many bird groups remain inadequately monitored. To that end, greater attention to comprehensive, large-scale monitoring programs has been directed towards other groups of birds, including marshbirds (Wheeler 2006).

The planning of a North American breeding marshbird monitoring program was formally initiated in 1998 (Ribic *et al.* 1999), when focal species - rails (*Rallus* spp. and *Porzana carolina*), American and Least Bitterns (*Botaurus lentiginosus* and *Ixobrychus exilis*), Wilson's Snipe (*Gallinago gallinago*), Pied-billed Grebe (*Podilymbus podiceps*), American Coot (*Fulica americana*), Common Moorhen (*Gallinula chloropus*) and Purple Gallinule (*Porphyrula martinica*) - were identified; a standardized field protocol was later developed (Conway 2008); and consideration was given to statistical design and sampling issues. It was also decided that an ideal monitoring program would be multi-scale so that both local and regional needs for information could be met and there would be a feasible way of "rolling up" results to make inferences about marshbird population status and trends at regional and national scales (Ribic *et al.* 1999).

Although reasons for monitoring marshbirds (or any group of birds) are numerous, here we focus exclusively on monitoring changes in population size over time. We do not consider other useful objectives, although actual implementation of a monitoring program might address some of them. We focus on recorded counts, as indices of abundance, rather than estimates of abundance for two major reasons. First, methods to account for detectability in surveys of diverse assemblages of birds over extensive areas are not available (Johnson 2008). The detection process is more complex than has generally been acknowledged (Alldredge *et al.* 2007b; Simons *et al.* 2007), and the assumptions required to appropriately adjust counts for incomplete detectability are difficult to meet (Johnson 2008). Second, problems associated with counts of birds are much greater for secretive marshbirds than they are for most landbirds. Secretive marshbirds are, well, secretive: many are inconspicuous, they often hide in heavy vegetation, some vocalize only infrequently, and their reactions to the presence of human observers or to broadcasts of calls played by observers are largely unknown. Further, the habitats occupied by marshbirds are not conducive to human travel to conduct surveys. Nonetheless, the proposed sampling design would be compatible with any protocols that might account for detectability.

To be successful, the sampling framework must be both statistically sound and logistically feasible. The sampling design must produce data that readily allow "rolling up" results in a hierarchical manner. The design also needs to accommodate certain realities, such as the fact that some areas (e.g., national wildlife refuges) are likely to be surveyed much more intensively than other areas, and that surveys in many areas necessarily would rely on volunteer observers. Species considered were those identified at the 1998 workshop (Ribic *et al.* 1999). To keep the issue manageable, we restrict the geographic scope to the contiguous 48 United States. In principle, much of the proposed framework is appropriate for other parts of North America, although implementation issues (e.g., types and sources of wetland data) may differ.

Herein we provide a framework for monitoring marshbirds on a spatially extensive basis. We consider the difficulties associated with sampling marshbirds, discuss analysis issues and consider implementation challenges.

SAMPLING DESIGN

Design Criteria

We developed a sample selection protocol to address several desirable design criteria: 1) probability sampling to provide a rigorous basis for inference, 2) hierarchical design structure to permit nesting of sub-regions within larger geographical entities (e.g., states nested within regions), 3) spatial balance to improve precision of estimates and to ensure the sample is spatially well-distributed, 4) spatial clustering of sample locations to reduce costs, 5) adaptable to accommodate the anticipated high non-response rate attributable to landowner resistance and heavy reliance on volunteers as observers, and 6) surveywide consistency to foster simplicity in describing methodology, analyzing the data, pooling and scaling up results, and reporting findings.

The Sampling Universe

The sampling universe is the totality of breeding habitat (i.e. wetlands) potentially used by the target species within the contiguous 48 states. Knowing the sampling universe is important when developing a sampling frame and analyzing survey data but often is difficult to specify completely.

To develop the sampling universe, spatial databases can be used in conjunction with geographic information systems to identify wetlands. Doing so will permit broad-scale evaluation of the geographic regions to be monitored, which will facilitate stratification, simplify the logistics of developing the sampling frame and allow spatial analysis of the data. However, several issues relate to identifying wetlands from spatial databases: 1) defining a wetland, 2) the varying nature of wetlands, 3) the amount of open water in the wetland, 4) accessibility, and 5) regional consistency.

We suggest using a broad, inclusive definition of wetlands. Because wetlands change between years due to varying water availability and processes such as ecological succession and wetland modification, a more narrow definition likely would exclude habitat in some years.

The nature of wetlands differs greatly both among and within regions. Areas such as the Prairie Pothole Region are dominated by small, discrete wetlands, whereas areas such as coastal marshes are dominated by large, extensive wetlands. Therefore, we recommend that the sampling universe be divided into one universe consisting of small (≤3 ha), discrete wetlands and another universe consisting of large (>3 ha), extensive wetlands. The 3-ha demarcation was based on a circle of radius 200 m, the distance at which many calling marshbirds can be heard (e.g., Allen *et al.* 2004; Conway and Nadeau 2006). Both types of wetlands can occur in close proximity, so the universes are not spatially distinct, and the design will select samples from both universes in a coordinated fashion. Sampling units would consist of entire wetlands in the universe of small, discrete wetlands and portions of wetlands in the universe of extensive wetlands. We propose no minimum size for wetland area because some target species (e.g., Sora [*Porzana carolina*] and Virginia Rail [*Rallus limicola*]) regularly occur on small wetlands (Brown and Dinsmore 1986), even on wetlands smaller than 0.1 ha (D. H. Johnson, unpublished data). The minimum size of wetlands in the sample universe most likely would be determined by the resolution of the available data used to construct the sampling frame.

Many wetlands have large areas of deep and open water that would rarely be used by most target species and would be difficult for observers to access. For those reasons, large open-water areas should be excluded from the sampling universe. Areas of emergent vegetation and open water within 200 m of the wetland edge should be retained.

Wetland accessibility may be very limited either because of landowners denying access or because of difficulties traveling to certain points. Accordingly, the sampling universe should be classified into accessible and inaccessible wetlands, which, somewhat analogously to the discrete and extensive universes, would be analyzed differently.

Decisions about which wetland types will be deemed available for surveying should be comparable within and among regions. Such decisions could vary regionally in part because of differences in the quality of wetland maps available. For example, National Wetland Inventory (NWI; Wilen and Bates 1995) maps may be used for most states, but National Land Cover Database (NLCD; Homer *et al.* 2007) maps may be needed for other areas. The comparability of these maps, and the potential bias introduced by drawing samples from different map types, need to be assessed.

Three spatial databases likely would be suitable as a start to determining the sampling universe. All are available across most or all of the contiguous 48 states; they are the

National Wetlands Inventory, the National Hydrography Dataset (NHD) and the National Land Cover Database. Each database has its imperfections. For example, all three databases were mapped or classified over multiple years, and varying precipitation levels over time can affect the number and area of wetlands mapped. However, even if wetlands were perfectly mapped, water levels of extant wetlands will vary among and within years, as will the extent and composition of emergent vegetation. Despite their shortcomings, the numerous advantages that these databases provide make their use preferable to alternatives such as subjective identification of wetlands to include in sampling. Whichever wetland database is used, the actual wetland universe will change over time as wetlands are created or are lost. These changes can be tracked over time and may require the collection of ancillary data to account for changes in water conditions and numbers of wetlands.

To ensure that the sampling frame uses the best available wetland data, the selection of which database to use and the determination of accessibility in a region should be done using local expertise and field evaluation by joint ventures or with federal or state agencies overseeing implementation. Wetland maps should be evaluated periodically when more information becomes available and revised as necessary; this will be particularly important for wetlands near urbanizing areas or elsewhere where the chance of destruction is likely highest. If the underlying wetland habitat is permanently destroyed, then the affected sampling points would need to be replaced; supplemental sampling (see below) is one way to add points.

Site Selection—Two-stage Cluster Sample

Primary Sampling Unit Selection Protocol. Selection protocols would be implemented within geographical strata, where the strata may be Bird Conservation Regions (BCRs), states, or Bird Conservation Subregions (BCSs, the intersections of Bird Conservation Regions and states; Fig. 1, [Bart 2006]). Within each stratum, two-stage cluster sampling would be employed (Cochran 1977). The cluster sampling design entails separate selection protocols for the two sizes of sampling unit, a first-stage selection of primary sampling units (PSUs) followed by a second-stage selection of secondary sampling units (SSUs) within each first-stage PSU. PSUs would be land areas, such as 40-km^2 or 648-km^2 Environmental Protection Agency hexagons (White 2007) or legal townships, and would be selected by a procedure such as systematic sampling. The PSU structure allows for obtaining improved frame information and possibly even refined maps of wetlands within the sampled PSUs; such detailed information would not be needed for the entirety of wetlands in the 48 states, thus saving significant costs associated with developing frame information needed to select the sample.

Figure 1. Strata could be formed by the intersections of state boundaries (dark lines) and Bird Conservation Regions (light lines; U.S. North American Bird Conservation Initiative Committee 2007), as suggested by Bart (2006).

Within each PSU, the sampling universe of marshbird habitat would be developed as described above. Wetlands would be characterized as either small (≤3 ha) and discrete wetland basins or large (>3 ha) and extensive wetlands. Areas of deep, open water more than 200 m from the wetland shore would be excluded. Each wetland or portion of wetlands would be categorized as accessible or inaccessible due either to difficulty of physical access (e.g., distant from roads) or to the lack of permission from owners. The inability to prospectively identify the accessibility of sites is an issue needing resolution. The areal extent of each wetland type, by accessibility, would be determined using a Geographical Information System (GIS). Model-based estimation would be used to extrapolate results to inaccessible sites.

Sample sizes of PSUs will affect the precision of estimators, will be limited by available survey effort, and likely will vary greatly among strata and regions. Although variable sampling effort among strata is not ideal for the purpose of monitoring overall population size, unbiased estimation would still be possible with such unequal effort. The flexibility to shift sampling effort among strata may be advantageous when the survey is implemented. For example, intensifying sampling effort in larger strata with greater populations of marshbirds at the expense of reducing PSU sample sizes in smaller strata with fewer marshbirds will lead to greater precision in the overall estimator of marshbird population changes.

Secondary Sampling Unit Selection Protocol.-Secondary sampling units (SSUs) will be wetlands or portions of wetlands in the PSUs. Secondary sampling units will be selected by a randomized spatially balanced procedure such as Generalized Random Tesselation Sampling (GRTS; Stevens and Olsen 1999, 2003, 2004) or the method of Lister and Scott (2009). Points within extensive wetlands should be at least 400 m apart (based on the field protocol of Conway 2008). For each SSU, on-the-ground observers will identify accessible locations well-suited for detecting marshbirds and record their positions with a Global Positioning System (GPS) unit, so that observers can revisit the same locations on subsequent occasions.

The SSU selection protocol depends on whether the PSU contains only small and discrete wetlands, only large extensive wetlands, or a mixture of both discrete and extensive wetlands. Wetlands may stretch across PSU boundaries, thus wetlands are defined as being within a PSU if the wetland's centroid, as determined from the frame information, falls within the PSU. All discrete (and accessible) wetlands within the PSU will be listed. If the PSU contains only discrete wetlands, the number of discrete wetlands sampled in each PSU will be a maximum of ten. If fewer than ten discrete wetlands are present in a PSU, all of them will be sampled. If more than ten discrete wetlands are present in a PSU, a GRTS or Lister-Scott protocol will be used to select ten SSUs from all wetlands present.

If the PSU contains only extensive wetlands, the GRTS or Lister-Scott protocol applied to a continuous spatial domain will be used to select a sample of point locations in accessible, extensive wetlands within the PSU. The number of points selected in a PSU depends on the area of extensive wetland within the PSU, as determined from the frame information. We tentatively propose the recommendations in Table 1, as used in a pilot study in Wisconsin in 2008. For that study, a grid of 40-km^2 hexagonal cells was used to delineate the PSUs. Sample size guidelines for SSUs for the Wisconsin pilot study were based on the number of SSUs that could potentially fit within the area of extensive wetland, while maintaining the recommended 400-m minimum spacing between SSUs (Conway 2008). Thus, there could be about one SSU for every 12.5 ha of extensive wetland.

For the Wisconsin pilot study, it was not possible to determine accessibility of SSUs prior to point selection. Instead, accessibility is being determined during ground-truthing of selected points prior to actual survey, and

Table 1. Proposed secondary sampling unit sample-size guidelines, based on the number of accessible discrete sampling sites (*k*) and the area of accessible extensive wetland within a primary sampling unit.

Available		In Sample	
Discrete (*k*)	Extensive (ha)	Discrete	Extensive
1-10	<1	All available	0
>10	<1	10	0
0	1-20	0	2
0	20-80	0	4
0	80-160	0	6
0	160-240	0	8
0	≥240	0	10
k (*k*>0)	1-20	min (8, *k*)	2
k (*k*>0)	20-80	min (6, *k*)	4
k (*k*>0)	80-160	min (4, *k*)	6
k (*k*>0)	160-240	min (2, *k*)	8
k (*k*>0)	≥240	min (2, *k*)	10-min (2, *k*)

adjustments due to accessibility (addition or deletion of SSUs) are made using an oversample generated during GRTS point selection (see "Supplemental Sampling" below).

If both discrete and extensive wetlands are present in a PSU, the sample will include both discrete wetlands and portions of extensive wetlands, the number combined not exceeding ten. Using the rule provided in Table 1 for determining the sample size for points in extensive wetlands, first establish the number of sample points to be located in the extensive wetlands. The number of accessible discrete wetlands to sample in that PSU would then be the number required to reach the maximum of ten sample locations in the PSU (Table 1). Clearly the number of discrete wetlands sampled will be limited by the number present in the PSU, and the extensive wetland sample points still must be located to maintain the recommended minimum distance of 400 m between points. Either the GRTS or Lister-Scott selection protocol will be implemented to select discrete wetlands for survey within the PSU.

The protocols for selecting the SSUs are such that discrete wetlands in different PSUs will not be sampled with the same inclusion probabilities, and points in extensive wetlands of different PSUs will not be sampled with the same inclusion densities. This unequal probability structure can be accounted for by the Horvitz-Thompson estimator (Horvitz and Thompson 1952) or its continuous-universe extension (Cordy 1993), which incorporates estimation weights based on the inverse of the inclusion probabilities for discrete wetlands or inclusion densities for extensive wetlands.

Changing Conditions at the Sampling Point

In some wetland systems that may undergo large annual fluctuations in size due to variable precipitation inputs (e.g., Prairie Pothole Region), initially suitable spots for observations might become unsuitable (e.g., temporarily dry, permanently destroyed). If the spot is in a small discrete wetland, the observer should move to another suitable observation point that would yield a census for that wetland. The assumption here is that the observer can effectively census the small wetland, so the spot where the observer stands can change from year to year. In extensive wetlands, the sampling location should not change if the observation location becomes unsuitable; a zero should be recorded for each species if the once-suitable habitat is no longer suitable. If the underlying wetland has been destroyed, the point still needs to be revisited until the wetland habitat data layer is updated and the point replaced. It will be important that notes about the condition of the point be preserved and used in subsequent analysis and revision of the survey design. In particular, a revised data layer should indicate that the destroyed wetland is no longer in the universe of wetlands from which samples can be drawn.

Supplemental Sampling

If some of the initial set of discrete sample wetlands or point locations in extensive wetlands cannot be surveyed, a list of replacement samples will be constructed as follows. (N. B. Replacement samples should be used *only* in situations where an existing wetland or point is not accessible or not practical to visit, or is not actually a wetland; replacements should not be used as a matter of convenience.) For discrete wetlands, the replacement sample will be the next wetland within the PSU specified by the GRTS or Lister-Scott method. The replacement samples must be used in the order specified by the sample-selection protocol. For example, it is not acceptable to skip over the next wetland in the replacement list to find a wetland closer to a wetland already included in the sample. For the extensive wetland sample replacements, the GRTS or Lister-Scott selection protocol will include extra sample locations to be visited. Again, it is important that replacement sample locations be selected in the order provided by the design to maintain as much as possible the probability sampling feature of the protocol.

Combining Results from the Two Sampling Universes

A challenging issue in designing an extensive marshbird survey involves the combination of results from the universe of small, discrete wetlands and from the universe of large, extensive wetlands. For the former, proper sampling units are wetlands themselves; for the latter, portions of wetlands are appropriate sampling units. If there were sampling units of only one type, analysis of results from the survey would be relatively straightforward.

In some situations, having two universes will pose little problem during analysis and interpretation. If, for example, counts of a species in both universes either increase or decrease at the same rate, it is clearly reasonable to conclude that the population in its entirety is increasing or decreasing at that same rate. Less restrictively, if estimates in both universes are either both increasing or both decreasing but at different rates, it is reasonable to assume that the entire population is following the same trend, although an estimate of the overall rate of increase or decrease is not immediately obvious. Only when estimates in the two universes trend in opposite directions is the overall pattern uncertain. In that case, it would appear that birds may be shifting from one universe to another. Such shifts can be expected when, for example, serious drought causes many small, discrete wetlands to dry up, and birds that remain in the area shift to larger, more extensive wetlands.

To resolve the problem and provide an overall estimate of the change in a population index, it is necessary to estimate the effective area surveyed for each sample wetland or point and to assume that the same fraction of birds is being detected in both discrete wetlands and extensive wetland areas. Effective area represents the area in which a bird would be detected if it were visible to the observer or if it emitted a call. Those values obviously would vary by species. For a small, discrete wetland, the effective area surveyed likely would be the entire wetland, at least for most species. For a survey at a large, extensive wetland, the effective area surveyed would be some fraction of the wetland size.

The effective area surveyed, especially on large, extensive wetlands, varies not only among species, but in relation to other characteristics, such as observer, size and shape of the wetland, presence and type of emergent vegetation, time of year, time of day, weather conditions, etc. For now, we recommend using empirically derived species-specific effective areas and then controlling for as many of those other characteristics as is feasible by following a designed study protocol. For example, training and testing of observers will reduce the effect of observer variability, and restricting surveys to certain calendar dates, time periods and weather conditions will reduce the influence of those features. Nonetheless, a "puddle" of unaccounted-for variation always will remain (Caughley and Goddard 1972: 136). At this time we recommend that such variation be acknowledged, but we do not propose a protocol to remedy it.

Initial general estimates of effective area surveyed in large, extensive wetlands can be obtained from available information or, if needed, expert opinion. A similar procedure was used by Partners in Flight to estimate bird densities from point count data collected on the North American Breeding Bird Survey (Rosenberg and Blancher 2005; Thogmartin *et al.* 2006). Refinement of initial estimates could come either from directed studies or possibly through internal evaluations during operational surveys. Should substantially improved estimates of effective area surveyed become available at some future time, it would be straightforward to retrospectively refine the estimated indices from previous years. We suspect that such refinement would have little influence on conclusions about population trends reached from the original counts, however.

ANALYSIS

Adherence to a rigorous sampling design such as the one described here will result in "good" data that can be analyzed in numerous ways that possibly involve different assumptions. We in fact encourage the use of various analysis methods as a means of increasing confidence in conclusions that may be reached. Somewhat different analyses may be appropriate for inventory versus monitoring purposes. For inventorying at a particular time *t*, standard sample survey analysis methods (e.g., Cochran 1977; Thompson 2002) should prove suitable for the accessible universes. For the inaccessible universes, model-based estimators (described below) will be required. For monitoring purposes, interest is in the difference in counts for each species at times, say, *t* and *t* + 1. In that situation, a more accurate estimate often may be obtained by using only sample points that were surveyed on both occasions.

Extrapolating count results from visited accessible sampling units to the totality of inaccessible area will most appropriately be based on predictive models that relate the response variable (counts by species) to relevant explanatory variables for which measurements at inaccessible sites can be obtained. Those explanatory variables likely would be habitat features and other variables that can either be assessed remotely (e.g., wetland class) or imputed from other data (e.g., wetland level based on precipitation patterns). Vital to this process will be model evaluation. Directed studies should be undertaken to evaluate these models, for example by collecting relevant data (explanatory variables and response variables) from a variety of sites to assess how well the observations conform to model predictions. Once again, if improved models become available in the future, they could be used to retrospectively re-analyze data from earlier years.

The hierarchical nature of the proposed design facilitates analysis at various levels. As one example, results from a number of wildlife refuges could readily be aggregated to evaluate patterns across the refuge system. As another example, findings from Bird Conservation Subregions could be combined to provide estimates either for a Bird Conservation Region or for a state. It also is possible to generate estimates on larger areas that do not conform to strata, typically called subpopulations or domains, but this often is more challenging (Cochran 1977).

The choice of analytic method will depend on the state of the art at the time enough data have been collected to estimate patterns. Accurate assessments of trends do not come quickly; many years will be required. For example, Urquhart *et al.* (1998) used a linear regression model to develop a recommendation of a minimum of 10-15 sample occasions (usually years) to detect even a moderate trend in water-quality variables. Obviously the length of time needed will depend upon the actual rate of change in the marshbird populations and the intensity of sampling each year (topics beyond the scope of this report). This said, a power analysis should be conducted once the overall analytical framework has been established and the requisite policy decisions have been made about desired accuracy of trend estimates, temporal windows for inference, and tolerable error rates for trend detection within the overall monitoring program. A power analysis is critical for elaborating the costs involved in launching a monitoring program that will deliver the information expected of it.

Possibilities for Abundance Estimation

The sampling design proposed here attempts to provide only indices of abundance. Although not founded on the assumption that estimation of density or absolute abundance will be feasible, the sampling design still would be appropriate for collecting data to be used to estimate density or abundance. A recent spate of efforts to account for incomplete detectability of surveyed birds includes distance sampling (Buckland *et al.* 1993), multiple-observer methods (Nichols *et al.* 2000; Alldredge *et al.* 2006), time-of-detection methods (Farnsworth *et al.* 2002; Alldredge *et al.* 2007a), double sampling (Thompson 2002; Bart and Earnst 2002), and methods based on multiple counts (Robson and Whitlock 1964; Royle 2004). Despite the mathematical rigor they bring to the issue, none of these methods has proven suitable for extensive surveys involving multiple species of landbirds (Johnson 2008). Moreover, their utility for secretive marshbirds is likely to be even less than for landbirds, because they indeed are secretive, and their behavior and detectability often are greatly influenced by the presence of an observer and especially the playing of calls.

Nonetheless, the sampling framework proposed here should be suitable for abundance estimation, should an accurate method for relating observed counts to actual numbers ever be devised. What would be required is the conversion of raw counts (indices) to abundance estimates at each sampled point. The aggregation of data from sample points up to strata and beyond would proceed just as with the count data.

USE OF CALL BROADCASTS

Most avian monitoring programs rely on birds to reveal themselves to observers by sight or through spontaneous vocalizations. Because secretive marshbirds often remain concealed in dense vegetation and vocalize only infrequently, many monitoring studies of these birds have used broadcasts of recorded calls to elicit responses (e.g., Glahn 1974; Marion *et al.* 1981; Johnson and Dinsmore 1986; Manci and Rusch 1988; Gibbs and Melvin 1993). The call broadcast survey method (also called tape-playback, playback or acoustic-lure survey methods) essentially exploits avian communication systems by mimicking (usually) a conspecific bird newly arrived at a site to stimulate an aggressive response from a resident bird, which then can be detected. The call broadcast method has been central to several proposals for a marshbird monitoring program in North America (Ribic *et al.* 1999; Bart 2006; Conway 2008).

The value of the call broadcast method relative to the costs, logistical issues and potential sampling biases associated with implementing it has received surprisingly little scrutiny. Notably, Conway and Gibbs (2005) identified many potential drawbacks to call broadcasts for monitoring marshbirds. These drawbacks include disturbance to the birds being surveyed by engaging them in unnecessary, energetically expensive vocal signaling (Kerlinger and Wiedner 1991), suppressing or eliciting the calling of one species by broadcasting the call of another and causing habituation among the surveyed population to the call's broadcast. Other issues are associated with deploying the audio equipment needed to broadcast calls, including the need to standardize equipment and its usage among users, across sites, and over the course of a monitoring program, as well as issues of "wear-and-tear" and subsequent variation in quality of broadcast calls associated with prolonged equipment use. Also, improvements in audio technology in the future may compromise the ability to maintain consistency in calls over long periods of time; that is, improved audio equipment may make future call broadcasts more effective. There are also significant economic costs associated with securing audio equipment required to broadcast calls. Which calls are most appropriate to broadcast in any given region is rarely obvious, nor is the optimal time for conducting surveys in different regions (Rehm and Baldassarre 2007). Last, although repeated exposure to broadcast calls can increase observers' probability of detecting birds, it also can interfere with observers' ability to hear birds calling. In the jargon associated with the detection issue (e.g., Johnson 2008), call broadcasts may increase availability but decrease perceptibility of calls of actual birds.

Ultimately, the value of call broadcast-derived data for monitoring marshbird populations must be assessed in the context of the quality of inferences such data can provide about changes in marshbird populations. Responsiveness of individual birds to call broadcasts is variable and is influenced by many features, including time of day, season, species and mating status (e.g., Legare *et al.* 1999). To provide an effective index of population change, call-broadcast methods should generate monitoring data that are both sound and precise; that is, detection probability should be high and have low temporal variation (Johnson 1995).

Several studies (e.g., Gibbs and Melvin 1993; Lor and Malecki 2002; Allen *et al.* 2004) have shown that call broadcasts increase, often dramatically, detection rates for a variety of species. Whether or not they improve count precision is less clear. A recent meta-analysis (Conway and Gibbs 2005) of simultaneously collected passive and call broadcast data provided insight on the contribution of call broadcast-derived data to both detectability and precision of marshbird counts. This synthesis of data from more than 16,000 point counts contributed by 15 cooperators using call-broadcast methods for twelve species revealed that broadcasting calls does indeed lead to greater detectability (increased the mean number of responses and the proportion of sites with a response) and generally increased (albeit modestly) precision by decreasing the variance in response rates. Although those authors endorsed the use of call broadcasts for monitoring marshbird populations, they did not recommend relying entirely on the method because calling activity by some species was depressed by the broadcast of other species' calls. For this reason, Conway and Gibbs (2005) recommended that the combination of passive and broadcast call survey methodologies previously employed in many marshbird survey efforts be continued. Such indeed has been the case in various, more recent formulations of a standardized protocol for monitoring marshbirds across the continent (e.g., Conway 2008). We still lack an analysis of the explicit costs and benefits of call broadcasts in terms of accuracy, precision, bias and costs relative to passive-only marshbird sampling programs but, issues of cost aside, call-broadcast surveys at present appear to have a useful role in facilitating large-scale monitoring of marshbird populations.

Including Other Species in Marshbird Surveys

It is always tempting, when conducting surveys of a selected suite of birds, to consider counting other species as well. If, for example, in a particular area some non-target species are of special interest, should observers also count those species?

The answer is not straightforward. It would seem that, because the observer already has made the (often considerable) effort to reach a survey point, it would be prudent to learn about other species at the same time. Conversely, time spent seeking, observing, and recording these non-target birds may detract from the quality of the counts of the target species. Such decisions should not be made lightly. We suggest these issues be considered: 1) the number of additional species, 2) separate surveys for target and non-target species, and 3) performance of surveys with and without non-target species included.

There should be a strict limit on the number of additional species to be recorded. Those species should be uncommon, with an identified need for specific information that can be met by the survey but not by feasible alternative sources. If species are added, the group of species should be kept the same over time. These restrictions will help avoid excessive effort being diverted from the target species.

If information on non-target species is needed, it may be prudent to conduct surveys for the target and non-target species in two consecutive time intervals. Observers could focus on the target species for the initial time interval and then turn their attention to the non-target species. It may be difficult (and inefficient), however, for observers to disregard detections of non-target species in the first interval and of target species in the second interval. Also, this approach will take longer, so fewer points could be surveyed in the same period of time.

If it is decided to include non-target species, some evaluation should be conducted to compare performance of the expanded survey with results obtained if only target species were counted. Such an assessment would usefully guide decisions about the value of expanding the survey to include non-target species.

REMAINING ISSUES

Although we provide a framework for a statistically sound marshbird monitoring plan, considerable additional effort is required to work out many of the details. One involves stratification criteria. We propose stratification based on regional boundaries, wetland type (small and discrete versus large and extensive) and accessibility. Implementation of such a stratified design should be coordinated nationally but will require considerable local expertise. The initial spatial stratification into bird conservation subregions developed by Bart and his cooperators (Bart 2006) should prove valuable. Stratifying by wetland type and by accessibility will involve spatial databases as described above, in combination with local expertise. Also to be determined is how to accommodate shifts of a wetland from accessible to inaccessible status, and vice versa, due to changing landowner attitudes or wetland condition.

Initial sample-size recommendations should be refined based on experience and more detailed evaluations in particular areas. The current proposal for a sample size of ten within a PSU is based on what is anticipated to be a reasonable workload for a single day or two. Variation from this sample size would not pose a problem as long as sample sites were visited in the order determined by the GRTS or Lister-Scott selection protocol.

The criteria for locating the observation points for the selected discrete wetlands need further resolution. This activity should be done in the field, rather than solely with remotely sensed information. Observation points should be selected 1) to be accessible by observers, 2) so that they offer good vantage points for observing the wetland site, and 3) so that these properties hold under a variety of conditions (e.g., during both wet periods and dry periods). The latter criterion is essential so that, for example, an observation point selected along the edge of a wetland during a dry period is not flooded and inaccessible during a wet period.

Determining the effective area covered at each sampling point will require considerable effort. Initial estimates could be developed based on available information and expert opinion, with refinements made based on directed studies and internal evaluations, as mentioned above.

Clearly a monitoring program such as outlined here should be phased in. Certain states and national wildlife refuges have begun already or will be ready to proceed very soon. The availability of suitable wetland maps may also influence the implementation schedule. Early experiences from these areas will offer useful guidance for a more complete implementation.

Although this proposal addresses only the contiguous 48 states, other states and nations need to be considered as well. Scientists in Canada, for example, have been extremely active in developing a marshbird monitoring program (e.g., Crewe *et al.* 2005).

While the focus here was on the sampling framework, it is recognized that sample design and survey protocols are intrinsically linked. The current protocol (Conway 2008) is largely the embodiment of proposals offered at the 1998 marshbird workshop (Ribic *et al.* 1999). Many of those proposals had been suggested based on personal observations and only limited information. It would seem very desirable to evaluate those protocols before they become too institutionalized. The extent to which they could be evaluated during the early phases of implementing the marshbird monitoring program should be fully explored.

The sampling framework recommended here meets the practical realities of a national marshbird monitoring and at the same time provides a statistically rigorous foundation for monitoring. The framework is predicated on implementing a probability sampling design to select the locations at which marshbird abundance will be observed. Probability sampling and associated design-based inferences are an accepted, commonly used basis for rigorous scientific inference from sample surveys (Särndal *et al.* 1992). Stratification by broad geographic regions (e.g., BCRs or states) is recommended to facilitate reporting of trends at meaningful geographic scales. However, this large-scale stratification does not preclude estimating trends in marshbird abundance for other spatial domains of interest. The design readily permits aggregating the data to produce results at different spatial scales, thus providing the needed feature of being able to "roll up" results for regional and national reporting. Two-stage cluster sampling is recommended because it provides a mechanism to reduce travel costs by concentrating the sample observations within a set of spatially well-distributed clusters. The two-stage design also readily accommodates sampling from two universes, discrete (small) wetlands and extensive wetlands, the two primary habitats for marshbirds. The GRTS or Lister-Scott selection protocols are recommended because they address the significant design problems associated with the high degree of non-response (e.g., denied access to private land) anticipated for such a monitoring protocol. These selection protocols provide a rigorous, statistically defensible way to replace sample locations that cannot be observed, while still maintaining the other benefits of the two-stage design (i.e. replacements are selected from the cluster in which the non-response occurs). The recommended design framework also allows the flexibility to increase the sample size within national wildlife refuges or other areas where resources and high interest justify more intensive sampling effort. Although the recommended probability sampling design is strongly suited to supporting design-based inference, it does not preclude model-based analyses of the data. In particular, we constructed the framework recognizing that model-based analyses will be necessary to provide estimates for the non-responding portion of the population.

ACKNOWLEDGMENTS

We are grateful to: P. Dresler, former manager of the USGS Status and Trends of Biological Resources program, for funding this report; the agencies and organizations that employ the authors; C. J. Conway, K. L. Drake, D. Granfors, H. Laskowski, S. J. Lewis, J. Wheeler and three referees for comments; and C. L. Amundson for assistance with the manuscript. The findings and recommendations presented in this article are those of the authors and do not necessarily represent the views of their employers.

LITERATURE CITED

Alldredge, M. W., K. H. Pollock and T. R. Simons. 2006. Estimating detection probabilities from multiple-observer counts. Auk 123: 1172-1182.

Alldredge, M. W., K. H. Pollock, T. R. Simons, J. A. Collazo and S. A. Shriner. 2007*a*. Time-of-detection method for estimating abundance from point-count surveys. Auk 124: 653-664.

Alldredge, M. W., T. R. Simons, and K. H. Pollock. 2007b. Factors affecting aural detections of songbirds. Ecological Applications 17: 948-955.

Allen, T., S. L. Finkbeiner and D. H. Johnson. 2004. Comparison of detection rates of breeding marsh birds in passive and playback surveys at Lacreek National Wildlife Refuge, South Dakota. Waterbirds 27: 277-281.

Bart, J. 2006. A sampling plan for the North American marsh bird monitoring program. U.S. Geological Survey Forest and Rangeland Ecosystem Science Center, Boise, Idaho. http://www.waterbirdconservation.org/pdfs/marshbirdmonitoring/a_sampling_plan_for_secretive_marshbirds_expanded.pdf, accessed 25 March 2009.

Bart, J. and S. Earnst. 2002. Double sampling to estimate density and population trends in birds. Auk 119: 36-45.

Brown, M. and J. J. Dinsmore. 1986. Implications of marsh size and isolation for marsh bird management. Journal of Wildlife Management 50: 392-397.

Buckland, S. T., D. R. Anderson, K. P. Burnham and J. L. Laake. 1993. Distance sampling: estimating abundance of biological populations. Chapman & Hall, London.

Caughley, G. and J. Goddard. 1972. Improving the estimates from inaccurate censuses. Journal of Wildlife Management 36: 135-140.

Cochran, W. G. 1977. Sampling techniques, 3rd ed. Wiley, New York, New York.

Conway, C. J. 2008. Standardized North American marsh bird monitoring protocols. Wildlife Research Report 2008-01. U.S. Geological Survey, Arizona Cooperative Fish and Wildlife Research Unit, Tucson, Arizona.

Conway, C. J. and J. P. Gibbs. 2005. Effectiveness of call-broadcast surveys for monitoring marsh birds. Auk 122: 26-35.

Conway, C. J. and C. P. Nadeau. 2006. Development and fieldtesting of survey methods for a continental marsh bird monitoring program in North America. Wildlife Research Report # 2005-11. USGS Arizona Cooperative Fish and Wildlife Research Unit, Tucson, Arizona.

Cordy, C. 1993. An extension of the Horvitz-Thompson theorem to point sampling from a continuous population. Probability and Statistics Letters 18: 353-362.

Crewe T. L., S. T. A. Timmermans and K. E. Jones. 2006. The Marsh Monitoring Program 1995 to 2004: A decade of marsh monitoring in the Great Lakes Region. Environment Canada, Port Rowan, Ontario. http://www.bsc-eoc.org/download/mmp10yrpt.pdf, accessed 25 March 2009.

Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines and J. R. Sauer. 2002. A removal method for estimating detection probabilities from point-count surveys. Auk 119: 414-425.

Gibbs, J. P. and S. M. Melvin. 1993. Call-response surveys for monitoring breeding waterbirds. Journal of Wildlife Management 7: 27-34.

Glahn, J. F. 1974. Study of breeding birds with recorded calls in north-central Colorado. Wilson Bulletin 86: 206-214.

Homer, C., J. Dewitz, J. Fry, M. Coan, N. Hossain, C. Larson, N. Herold, A. McKerrow, J. N. Van Driel and J. D.

Wickham. 2007. Completion of the 2001 National Land Cover Database for the conterminous United States. Photogrammetric Engineering & Remote Sensing 73: 337-341.

Horvitz, D. G. and D. J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47: 663-685.

Johnson, D. H. 1995. Point counts of birds: what are we estimating? Pages 118-123 *in* Monitoring bird populations by point counts (C. J. Ralph, J. R. Sauer and S. Droege, Eds.). General Technical Report PSW-GTR-149. U.S. Department of Agriculture, Forest Service, Washington, D.C.

Johnson, D. H. 2008. In defense of indices: the case of bird surveys. Journal of Wildlife Management 72: 857-868.

Johnson, R. R. and J. J. Dinsmore. 1986. The use of tape-recorded calls to count Virginia Rails and Soras. Wilson Bulletin 98: 303-306.

Kerlinger, P. and D. S. Wiedner. 1991. Vocal behavior and habitat use of Black Rails (*Laterallus jamaicensis*) in south Jersey. Records of New Jersey Birds 16: 58-62.

Legare, M. L., W. R. Eddleman, P. A. Buckley and C. Kelly. 1999. The effectiveness of tape playback in estimating Black Rail density. Journal of Wildlife Management 63: 116-125.

Lister, A. J. and C. T. Scott. 2009. Use of space-filling curves to select sample locations in natural resource monitoring studies. Environmental Monitoring and Assessment 149: 71-80.

Lor, S. and R. A. Malecki. 2002. Call-response surveys to monitor marsh bird population trends. Wildlife Society Bulletin 30: 1195-1201.

Manci, K. M. and D. H. Rusch. 1988. Indices to distribution and abundance of some inconspicuous waterbirds on Horicon Marsh. Journal of Field Ornithology 59: 67-75.

Marion, W. R., T. E. O'Meara and D. S. Maehr. 1981. Use of playback recordings in sampling elusive or secretive birds. Studies in Avian Biology 6: 81-85.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117: 393-408.

Rehm, E. M. and G. A. Baldassarre. 2007. Temporal variation in detection of marsh birds during broadcast of conspecific calls. Journal of Field Ornithology 78: 56-63.

Ribic, C. A., S. Lewis, S. Melvin, J. Bart and B. Peterjohn, Eds. 1999. Proceedings of the marsh bird monitoring workshop. U.S. Fish and Wildlife Service Region 3 Administrative Report, Fort Snelling, Minnesota.

Robbins, C. S., D. Bystrak and P. H. Geissler. 1986. The breeding bird survey: its first fifteen years, 1965-1979. U.S. Department of the Interior, Fish and Wildlife Service Resource Publication 157. Washington, D.C.

Robbins, C. S., J. R. Sauer, R. S. Greenberg and S. Droege. 1989. Population declines in North American birds that migrate to the neotropics. Proceedings of the National Academy of Sciences 86: 7658-7662.

Robson, D. S. and J. H. Whitlock. 1964. Estimation of a truncation point. Biometrika 51: 33-39.

Rosenberg, K. V. and P. J. Blancher. 2005. Setting numerical population objectives for priority landbird species. Pages 57-67 *in* Proceedings of the Third International Partners in Flight Conference (C. J. Ralph and T. D. Rich, Eds.). General Technical Report PSW-GTR-191. U.S. Department of Agriculture, Forest Service, Washington, D.C.

Royle, J. A. 2004. *N*-mixture models for estimating population size from spatially replicated counts. Biometrics 60: 108-115.

Särndal, C. E., B. Swensson and J. Wretman. 1992. Model-assisted survey sampling. Springer-Verlag, New York, New York.

Simons, T. R., M. W. Alldredge, K. H. Pollock and J. M. Wettroth. 2007. Experimental analysis of the auditory detection process on avian point counts. Auk 124: 986-999.

Smith, G. W. 1995. A critical review of aerial and ground surveys of breeding waterfowl in North America. U.S. Department of the Interior, National Biological Service, U.S. Biological Science Report 5, Washington, D.C.

Stenseth, N. C. and A. Mysterud. 2002. Climate, changing phenology, and other life history traits: nonlinearity and match-mismatch to the environment. Proceedings of the National Academy of Sciences 99: 13379-13381.

Stevens, D. L. and A. R. Olsen. 1999. Spatially restricted surveys over time for aquatic resources. Journal of Agricultural, Biological, and Environmental Statistics 4: 415-428.

Stevens, D. L. and A. R. Olsen. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics 14: 593-610.

Stevens, D. L. and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99: 262-278.

Thogmartin, W. E., F. P. Howe, F. C. James, D. H. Johnson, E. T. Reed, J. R. Sauer and F. R. Thompson III. 2006. A review of the population estimation approach of the North America Landbird Conservation Plan. Auk 123: 892-904.

Thompson, S. K. 2002. Sampling, 2nd ed. Wiley, New York, New York.

Urquhart, N. S., S. G. Paulsen and D. P. Larsen. 1998. Monitoring for policy-relevant regional trends over time. Ecological Applications 8: 246-257.

Wheeler, J. 2006. Moving ahead on a continental-scale marsh bird monitoring framework. Waterbird Conservation of the Americas, served by the U.S. Fish and Wildlife Service, Migratory Birds Division. http://www.fws.gov/birds/waterbirds/monitoring/marsh-monitoring.html, accessed 25 March 2009.

White, D. 2007. Spatial grid networks. U.S. Environmental Protection Agency, Western Ecology Division, Corvallis, Oregon. http://www.epa.gov/wed/pages/staff/white/getgrid.htm, accessed 25 March 2009.

Wilen, B. O. and M. K. Bates. 1995. The US Fish and Wildlife Service's National Wetlands Inventory project. Vegetatio 118: 153-169.