

General Comments on *Monitoring Colorado's Birds*

23 December 2004

Dr. Paul F. Doherty, Jr.
Department of Fishery and Wildlife Biology
Colorado State University
Fort Collins, CO 80523-1474

Summary

Monitoring Colorado's Birds is an ambitious program, operating within a declining species framework, with the goal of detecting declining population trends. *Monitoring Colorado's Birds* addresses deficiencies in other bird monitoring programs by utilizing a two-stage sampling plan in which stratified random sampling (by habitat type) is the first stage and distance sampling is used in the second stage to incorporate detection probabilities into density estimates. *Monitoring Colorado's Birds* could possibly be improved by readdressing sample size calculations with the additional data collected in the past years, as well as utilizing more reliable software/methods. In some instances distance sampling is not being instituted correctly (i.e. river surveys). Other, novel, survey methods may also be considered for some species of interest (i.e. colonial nesters). I also suggest the underlying framework be moved nearer to a knowledge gaining framework in which the monitoring program is tied more closely to management actions. This will take additional prioritization and coordination by management agencies.

Comments

My comments will focus on three aspects. I will first focus on the underlying philosophical framework. I will then present suggestions for improving the probability-based sampling scheme as well as dealing with issues of detection. I also include as an appendix (Appendix 1) a report I recently wrote for the Colorado Division of Wildlife that provides in-depth details, and references, on these topics.

Underlying framework

The underlying philosophical framework to the monitoring plan is a declining species framework in which the goal is to detect declining population trends. In this framework, monitoring is a part of a four-step process:

- 1) Monitor trends in population abundance (or related index).
- 2) If the trend is negative or passes some 'trigger', then enter into a research program to determine the demographic parameter(s) causing the decreasing trend.
- 3) When the demographic parameter(s) is/are identified, further research the environmental factor(s) causing the change in the demographic parameter.
- 4) Ameliorate and mitigate the environmental factor(s).

Under this framework monitoring is used to detect declining trends. However, many years of effort can be spent with very little learned to better manage bird populations. Only if a declining trend is detected does actually research on possible management actions take place (steps 2 - 4). With the current effort, a linearly declining trend for many species will not be able to be detected for 15-30 years. By this time it may be too late to begin developing management goals and strategies. This is probably not the best use of possibly 30 years worth of effort.

An alternative framework, a knowledge-gaining framework, incorporates monitoring activities into the testing of management actions, and/or hypotheses to resolve biological uncertainty. Learning and the development of management actions/goals is incorporated with the monitoring throughout the time period, not only after a decline is detected. This is a more efficient use of resources.

This second framework suggests that the management agencies would need to better coordinate their activities and the setting of avian population goals. I put forth the questions: How has the current data collection been used by the management agencies? Is the data collection informing management? Could efforts be directed to better meet informational and management needs? If the monitoring plan can only reliably detect trends in the most common species, is it the best to focus efforts in this way? Does monitoring of 'common' species need to occur every year?

A proactive approach is generally better than a reactive one. Resources are always scarce and not all species can be monitored well or be the focus of management. Clear prioritization and goals need to be expressed, as well as possible management actions. This is not an easy task. However, if this is done, the monitoring plan will be more useful and defensible. These actions will also further the NABCII goals of unified management goals and actions across agencies and partners within bird conservation regions.

Additional discussion of philosophical frameworks can be found in the appendix.

Probability-based sampling

The sampling plan stratifies by habitat type. This seems very reasonable to me. However, given the guidelines for choosing transects, the sampling frame is strictly defined as areas within ~4 km of a road. Results can only be extrapolated, statistically, to these areas. This may need to be stated more explicitly in the annual reports. Any further extrapolation would need to open additional areas to sampling effort.

Other sampling schemes may be of use for some species of interest. These are detailed in the appendix. Of particular notice may be the dual-frame sampling scheme that has been proposed for bald eagle nests. This scheme could easily be altered for other raptors or colonies of birds. Clearly defining sampling frames for waterbirds of interest may also lead to improved sampling schemes.

By revisiting the same transects, the resampling scheme is focused on detecting changes in temporal trends only. If different transects were visited each year, changes in trend and/or distribution could be detected. However if temporal trend is the only interest, then varying transects would not be optimal. One possible resampling scheme that may be of interest is a rotating panel. This is a resampling scheme in which some of the same plots are visited each year, as well as some new ones (see appendix).

One problematic feature in the design of the plan was the reliance on Program Monitor for sample size calculations. This program is no longer being supported by the USGS Patuxent Wildlife Research Center due to limitations in the flexibility needed in such sample size calculations and corresponding unreliable results. With the additional data collected a better job of sample size calculations could be accomplished using more robust methods.

Detection probabilities

A mixture of point and line transects are generally used, with other survey methodologies (i.e. raw counts) being used for species for which these methods don't work well, such as colony-nesting species. It should be noted that line transect data is much more reliable than point transect data due to the inherent number of observations available near the zero line or point. In general point transects will need about 10-25% more data points as compared to line transects. Line transects should be relied upon as much as possible with the recognition that in multi-species surveys, or with surveys in rough terrain, point transects are often the option of choice.

The transects conducted on rivers are problematic for two reasons. First I would imagine that it must be very hard to maintain the 'zero' line in a canoe. Second, transects need to be placed randomly to avoid being aligned with an underlying gradient in bird distribution. By aligning transects with the river flow, the transect is probably perpendicular to the underlying gradient in bird density associated with a river. This is exactly the worst situation. Some sort of random placement, even if the transects need to be shortened, would be preferable.

None of the reports present figures of fitted models, or model selection results. It is impossible for the reader to judge how reasonable the fit is for differing detection functions. What detection functions are being used?

In addition to relying upon distance sampling estimators at points, if the time of each bird detection was recorded, then estimators based on auditory removal models, or upon presence/absence latent estimator models (Royle and Nichols 2003) could also be used. These would be valuable in corroborating the distance estimates and may be more appropriate for some species. I urge consideration of collecting these data. In essence the only additional data needing to be recorded would be a specific time for when each bird was first detected.

Monitoring Colorado's Birds has provided evidence that large scale monitoring programs can be improved upon with attention to utilizing probability-based sampling schemes and incorporating detection error. In both of these arenas there are continuing developments that may be of use as the program and methodology evolve, especially for some subsets of avian species. Many project ideas are available for further testing and refining survey methodology.

Reevaluating how the collected data are being used to inform management, as well as whether this is the best use of resources and the expertise at the Rocky Mountain Bird Observatory should also be revisited periodically. What are management goals? Does the data collection need to occur every year? Would devoting resources to more focused studies in intervening years be of value? These and like-minded questions, are the purview of the management agencies and cooperators and need to be explicitly addressed periodically.

The State of Colorado is fortunate to have an organization with the skills and expertise such as the Rocky Mountain Bird Observatory (RMBO). It is a credit to the organization that they have been able to improve upon the design of large-scale bird surveys when so many such surveys are found lacking. No program is static and I am sure RMBO will continue to expand and improve upon their monitoring activities.

**Conceptual and practical issues in monitoring animals: A review and
recommendations**

Paul F. Doherty, Jr.
Department of Fishery and Wildlife Biology
Colorado State University
Fort Collins, CO 80523-1474
5 August 2004
In fulfillment of RWO 5-30212

Table of Contents - Appendix

PHILOSOPHICAL APPROACHES TO MONITORING.....	7
DESIGN AND METHODOLOGICAL ISSUES	10
<i>PRIORITIZATION.....</i>	<i>11</i>
<i>GENERAL GUIDELINES FOR SURVEY DESIGN.....</i>	<i>12</i>
<i>DETECTION PROBABILITIES.....</i>	<i>13</i>
<i>Recent methodologies incorporating detection probabilities</i>	<i>14</i>
Distance sampling.....	14
Capture-recapture based methods	15
Capture-recapture of immobile objects	15
Double observer methods.....	15
Auditory removal methods.....	15
Estimating population growth (λ)	16
Robust design	16
Barker model	16
Presence-absence based methods	16
Occupancy probability.....	17
Extinction and colonization probabilities	18
Species co-occurrence	18
Estimating abundance from presence-absence data.....	18
Community-level parameters (species richness, extinction, turnover).....	19
<i>SAMPLING.....</i>	<i>20</i>
Simple random sampling.....	20
Stratified sampling.....	20
Cluster sampling.....	20
Systematic sampling.....	21
Adaptive sampling.....	21
Double sampling.....	21
Ratio estimation	22
Dual frame sampling.....	22
Multi-stage designs	22
Rotating panel.....	23
REVIEW OF SELECTED MONITORING PLANS	24
<i>CHRISTMAS BIRD COUNT (CBC)</i>	<i>24</i>
Description and underlying philosophical approach	24
Detection error	24
Sampling error	24
<i>NORTH AMERICAN BREEDING BIRD SURVEY (BBS)</i>	<i>25</i>
Description and underlying philosophical approach	25
Detection error	25
Sampling error	25
<i>MONITORING AVIAN PRODUCTIVITY AND SURVIVORSHIP (MAPS).....</i>	<i>25</i>
Description and underlying philosophical approach	25
Detection error	26
Sampling error	26
<i>NORTH AMERICAN WATERFOWL MANAGEMENT PLAN (NAWMP).....</i>	<i>26</i>
Description and underlying philosophical approach	26
Detection error	26

Sampling error	26
<i>USDA NORTHWEST FOREST PLAN</i>	26
Description and underlying philosophical approach	26
Detection error	27
Sampling error	27
<i>NORTH AMERICAN AMPHIBIAN MONITORING PROGRAM (NAAMP)</i>	27
Description and underlying philosophical approach	27
Detection error	27
Sampling error	27
<i>ENVIRONMENTAL MONITORING AND ASSESSMENT PROGRAM (EMAP)</i>	28
Description and underlying philosophical approach	28
Detection error	28
Sampling error	28
<i>NATIONAL PARK SERVICE VITAL SIGNS</i>	28
Description and underlying philosophical approach	28
Detection error	28
Sampling error	29
<i>US FOREST SERVICE NATIONAL PROTOCOL FOR MONITORING VERTEBRATES</i>	29
Description and underlying philosophical approach	29
Detection error	29
Sampling error	29
<i>SUMMARY</i>	29
CONCLUSIONS AND RECOMMENDATIONS	31
ACKNOWLEDGMENTS	33
LITERATURE CITED	33

INTRODUCTION

Interest in monitoring has increased in recent years with the realization that monitoring data can be a key component to understanding and managing biological systems. Societal pressure, starting in the 1960s, has caused monitoring activities to change from a main emphasis on game species and supporting the setting of harvest regulations, to focusing on the conservation of all flora and fauna. This enlarged purview is, in part, mandated by legislation such as the National Environmental Protection Act (1969), Endangered Species Act (ESA; 1973) and the Forest Management Act (1976). For the ESA in particular, monitoring data is instrumental in listing decisions as well as a required part of recovery and habitat conservation plans. The number of court actions related to this legislation has also increased in recent years, bringing additional attention to the integrity of monitoring activities and related management actions. This spotlight will continue to focus more narrowly on the scientific rigor and design supporting such data collection. Management decisions will need to be based on data that are defensible in court (Federal Judicial Center 2000) and additional emphasis needs to be placed on examining the role of monitoring plans, as well as their designs, in management.

However, even with the ubiquity and increased responsibilities of monitoring plans, some have viewed monitoring as “prostituting ... science to a trivial activity” (Krebs 1991), “displacement behavior” (Nichols 1999), “ecologically banal” (Krebs 1991), and as “datakleptomania” – the uncontrolled desire to collect more data (Hellawell 1991 cited in Vos *et al.* 2000). Many of these criticisms point to existing uncertainty concerning the role of monitoring, what to monitor, and what methods to employ, especially with limited resources. Below I will first explore two philosophical viewpoints on the role of monitoring. Second, I will concentrate on general methodological issues associated with monitoring. Third, I will review a selection of contemporary monitoring programs. And lastly, in summary, I will provide general recommendations to the Colorado Division of Wildlife. My discussion and examples will focus on terrestrial animal populations, but the conceptual and practical issues are applicable to other biota.

PHILOSOPHICAL APPROACHES TO MONITORING

Monitoring programs are generally grounded in two frameworks, 1) an early-warning (Vos *et al.* 2000) or retrospective (National Research Council 1995) framework and, 2) an early-control (Vos *et al.* 2000), knowledge-gaining (Nichols 1999), or predictive (National Research Council 1995) framework. The first is based on Caughley's (1994) declining species concept, which suggests much of conservation biology is concerned with identifying declining species and keeping species from declining. This early-warning framework is also grounded in a human health analogy with the view that “natural area managers are essentially family physicians for natural areas” (Davis 1993). Monitoring under this viewpoint is typified by the use of phrases such as ‘monitoring the pulse of the planet’ (Dougherty 2000) or monitoring ‘vital signs’ (National Park Service 2004) and can generally be described as part of a four-step program (Green 2002, Underhill and Gibbons 2002):

- 1) Monitor trends in population abundance (or related index).

- 2) If the trend is negative or passes some ‘trigger’, then enter into a research program to determine the demographic parameter(s) causing the decreasing trend.
- 3) When the demographic parameter(s) is/are identified, further research the environmental factor(s) causing the change in the demographic parameter.
- 4) Ameliorate and mitigate the environmental factor(s).

This framework is best typified by the case of raptors and DDT (dichloro-diphenyl-trichloroethane, Underhill and Gibbons 2002). In the 1960’s, ostensibly through monitoring, declines in birds of prey populations were noted (step 1). Through research the decline was found to be related to eggshell thinning and decreased reproductive success (step 2 - identify the demographic parameter) caused by metabolites of DDT (step 3 - identify the environmental cause). DDT was taken off of the market and raptor populations are recovering (step 4 - amelioration).

Monitoring programs, under this framework, are predicated on being ‘warning systems’ by methodologically collecting time series of data on a number of parameters (most notably abundance) and then searching for patterns within, and correlations among, these time series.

A number of criticisms have been levied against this early-warning framework (Krebs 1991, Suter 1993, National Research Council 1995, Wicklum and Davies 1995, Nichols 1999). First, the human medicine analogy is untenable (Suter 1993, Wicklum and Davies 1995). Mammalian health is a property of individual organisms and implies consistent structure and development, tight integration, homeostasis, (Suter 1993) and the existence of an optimum condition (Wicklum and Davies 1995). Populations, communities, or ecosystems are not organisms, do not have clear boundaries (such as skin), do not have consistent structure and development, and lack the control systems to maintain homeostasis (Suter 1993). Nor are there fundamental, optimal states for organizational scales above the individual. Such optima would invoke group selection theory (Wicklum and Davies 1995).

The analogy between human health and ecological systems is worth further exploration. First, a medical doctor does not only examine one measure and declare a patient “60% healthy” (Suter 1993). Rather, a large number of measures are examined (e.g. body temperature, body weight, blood pressure, blood chemistry, etc.) and specific diagnoses are made. In contrast ecological monitoring programs are often based solely on a single measure, such as an index to abundance in the North American Breeding Bird Survey (Sauer *et al.* 2003) and the ‘health’ state of the system is judged (Greenwood *et al.* 1995).

Second, due to homeostasis, variation in human health measures is low. For example the coefficient of variation associated with body temperature is approximately 0.04% (Mackowiak *et al.* 1992) whereas the coefficient of variation associated with animal abundance is frequently much greater than 10% (Link and Nichols 1994). Sampling error is much less in human health measures than in ecological measures and large sampling errors in the later can compromise the ability to recognize meaningful patterns (Nichols 1999). Finally experimentation in human health systems is under strict ethical and legal constraints. Because of these constraints, the gaining of knowledge is mostly limited to retrospective analyses and surrogate systems. Fortunately the gaining of knowledge is

not as severely limited in this way for ecological systems; however ecological scientists often do not take advantage of this ability.

Beyond the poor analogy, the concept of ‘health’ is a value-laden abstraction, often leading to imprecise thinking and goals (Davis and Slobodkin 2004). This concept is often invoked as being solely scientifically based without recognizing the role of societal (ethical, economical, and political) value judgments (Wicklum and Davies 1995).

For the early-warning framework other criticisms, besides those focused on the human health analogy, have been levelled at the ability to gain knowledge efficiently (Nichols 1999). This argument states, at best, monitoring data collected under an early-warning framework can only provide weak-inferential, retrospective results, with the inherent value being in hypothesis generation. Stronger inferential science (Platt 1964, Romesburg 1981) and research are reserved for steps 2 and 3. Under such a framework, possibly a career’s worth of data can be collected with very little knowledge gained (especially if no decline is detected). This is a small return on the money and time invested and leaves the strong-inferential, knowledge-gaining process, and informed management, until it is almost too late (i.e. after steep population declines are detected). The National Research Council (1995) also suggests that a retrospective approach, while applicable for some environmental threats, is inappropriate in cases focused on avoiding ‘biological extinction’ due to the poor ability to detect declines and an ‘infinite’ mitigation time lag.

In the second framework, called an “early-control” (Vos *et al.* 2000), knowledge-gaining (Nichols 1999), or predictive (National Research Council 1995) framework, monitoring is typified by being a part of a program that is attempting to gain knowledge about a biological system constantly, not just after a problem is suspected (Nichols 1999). Testing *a priori* hypotheses, especially through manipulative experiments, will provide the strongest inferential knowledge (Platt 1964). This gaining of knowledge may be of a basic scientific nature and/or focused on possible management actions. This framework contends that *a priori* hypotheses concerning driving factors or possible management actions exist, and monitoring programs should be designed to test predictions resulting from these hypotheses. The monitoring data collected can still be used in a retrospective time trend, hypothesis-generating exercise, as in the early-control framework; however this is not the *raison d’être*. Time trend analysis is the secondary product, rather than the main focus of monitoring. Under this framework useful knowledge is gained efficiently throughout the ‘career’s worth’ of data collection. If understanding the basic ecological processes of a system is the focus of a monitoring program, then much more will be known, especially if management actions are needed and must be developed. If management actions exist and are investigated throughout the time period, when a problem does occur, managers are more capable of addressing problems. This framework may influence data collection in small or large ways.

For example, a monitoring plan for the long-billed curlew (*Numenius Americanus*) has been proposed (Jones *et al.* 2003) with the goal of generating a time series of density estimates (an early-warning framework). However, this time series will do little to inform management for this species of interest. If, in fact possible management actions are deciding which lands to buy, or influence through agreements with landowners, (T. Remington – pers. comm.) a small adjustment, under a knowledge-gaining framework, would be to stratify the area by vegetation types or ownership to more efficiently test *a*

priori ideas about variation in curlew numbers and provide better information influencing likely management/conservation actions. Samples should be allocated differentially to those strata that would be most informative, or most likely to discriminate among the competing hypotheses about which habitats are most valuable to curlews. This approach would still produce an overall estimate of curlew numbers for a time trend analysis, but would provide more precise estimates of differences in curlew numbers associated with different habitat types or ownership. The use of this knowledge-gaining framework lends itself to tying a monitoring program to the gaining of needed knowledge, and possible management actions (and implied goals); an aspect that would seem especially important to a wildlife management agency such as a state division of wildlife.

The catch-phrase ‘adaptive management’ (Walters and Holling 1990) also falls under this second framework. Unfortunately this term means many things to many people. I reserve the use of the term in the spirit of a dual-control process in which the dual objectives of gaining knowledge and reaching management objectives are sought through an optimal decision/control analysis (Kendall 2001, Williams et al. 2002). These methods are often rooted in operations research, which provides a quantitative basis for decision making. One example of the use of this concept is waterfowl harvest management in North America (Williams et al. 1996, Kendall 2001). Explicit goals are stated (number and harvest of ducks), differing population models are constructed representing uncertainty in the population biology (e.g. presence of density dependence in reproduction, additive vs. compensatory harvest mortality) with relevance to possible management options (e.g. possible bag limits and season lengths). These biological models are assigned differing weights of belief. Through an optimality modeling exercise, using stochastic dynamic programming, predictions are made from each model, weighted, and an optimal management decision is determined and carried out. The predictions are tested and model weights updated iteratively using data from monitoring programs. Such a program dually meets management objectives, and gains knowledge leading to a reduction in the uncertainty in duck population biology, leading to better management. The monitoring data are used to parameterize the models, test predictions, and update the belief systems. They are also used to provide estimates of system state needed for state-dependent management decisions. This is a very different use of a time series of abundance estimates than in a traditional early-warning framework.

DESIGN AND METHODOLOGICAL ISSUES

No matter what philosophical bent is taken, of paramount importance are proper parameter estimation and experimental and/or sampling designs (Yoccoz et al. 2001). If attention is appropriately paid to these aspects, studies will have integrity, and ultimately, results will be defensible in court. Calls for attention to proper design of monitoring programs have been made a number of times (e.g. Davis 1993, Thompson et al. 1998, Gibbs et al. 1999, Nichols 1999, Block et al. 2001, Yoccoz et al. 2001, Sauer 2003, Diffendorfer and Doherty 2004), and below I will review the important issue of prioritization as well as general guidelines for survey design. Each specific situation requires a specific design and addressing every possible situation is beyond the scope of this paper. Therefore I will briefly discuss these issues in general. However, I will concentrate on two particular issues that have plagued wildlife survey designs – namely attention to recognizing the importance of detection probabilities and formal sampling

strategies. Sauer (2003), following Skalski (1994), recognize these two issues as a 2-stage sampling design, where choosing sample units, using a formal probability-based sampling scheme, is the first stage and estimation of animal density (incorporating detection probabilities) within units is the second stage.

PRIORITIZATION

What to monitor is one issue many struggle with when setting up a monitoring plan (Mace and Collar 2002, Noon 2003), especially with limited budgets and with concerns for many species. Without a clear focus and objectives, a monitoring plan has a high likelihood of failure. When many species are involved, the question essentially becomes a priority setting exercise in which parameters of interest are identified, rationalized, ranked and a program of action defined. The setting of priorities is based on available information, and because information is in a constant state of flux, priority-setting should be understood as a process, subject to modification and upgrading (Mace and Collar 2002). A monitoring program that evolves with gained information will fit more easily into a knowledge-gaining framework.

Delineating priorities is value-laden and not a trivial task. However, to have direction, an organization needs a hierarchy of priorities. For instance, some species will need to receive a high level of attention (possibly federally listed or candidate species) while other species may receive an intermediate level of attention and resources, and some species may receive little to no attention (i.e. widely distributed, common species). In addition, priorities should reflect the possibility of management goals and actions. A species for which no management goals are likely to exist and no management actions are possible is unworkable and should probably receive a low priority. One downfall of many monitoring plans is too many objectives, such that none is achieved. This is, again, a reason for careful prioritization.

Monitoring programs may focus on different organizational levels such as an individual species/population level, community-level, or larger ecosystem level. Most monitoring is driven, at least in part, by individual species (possibly as part of a larger framework such as a habitat conservation plan) because this is the scale at which legislation is enforced (e.g. ESA, hunting regulations, etc.). Thus I will mostly focus on individual species.

In deciding which species to focus upon, there are a number of considerations. Such considerations may include perceived severity of extinction risk, recent declines, rarity, restricted range, evolutionary uniqueness, protection status, susceptibility to anticipated land-use changes, and perceived role in an ecosystem (Mace and Collar 2002). In general, species considered threatened/endangered or of economic interest generally receive a higher prioritization. However, some of these species may already have recovery/management and monitoring plans in place and require few additional resources. In evaluating priorities, Noon (2003) suggests starting with a conceptual model of a system to help crystallize what species, drivers, and possible management actions to focus upon. Such a model can help identify uncertainties and guide monitoring and management actions. While science is only one influence on setting priorities, the setting of clear, and reachable, priorities and goals is imperative for a successful monitoring plan.

GENERAL GUIDELINES FOR SURVEY DESIGN

Sample surveys are most commonly invoked in monitoring plans, especially large-scale plans. A sample survey is defined as observing a population without experimental manipulation. Sample surveys are usually distinguished from the related field of experimental design, in which deliberate manipulation occurs in order to observe the effect of the manipulation with proper randomization, control and replication. A discussion of experimental designs is beyond the scope of this report but many references exist for the interested reader (e.g. Montgomery 2001). However, sampling may be an important aspect in experiments, especially large-scale experiments where not all areas can be sampled, or individuals observed. Sampling is also distinguished from observational studies in which little control exists over how observations are obtained. In sample surveys, a sample is deliberately selected to avoid factors making data unrepresentative, such as observations based on convenience, haphazard selection, judgment, or incidental sampling (Thompson 2002). Attention to sampling is important because the ‘integrity of the sample’ (Stuart 1984) is paramount for any larger inference to be made than only to the sampled individuals, and this integrity starts with data collection. Otherwise the problematic GIGO (garbage in-garbage out) principle will be in effect, which unfortunately, can be misconstrued as garbage in-gospel out (Scheaffer *et al.* 1996) by unsuspecting, or unknowing, users. In essence, when statistics are not based on strict adherence to underlying assumptions, they mislead instead of guide and, unfortunately, the human mind easily lets itself be taken in by the false appearance of exactitude which statistics retain in their mistaken use (Tocqueville 1889 cited in Scheaffer *et al.* 1996).

Thompson *et al.* (1998) provide general guidelines for designing sample surveys. First clear objectives need to be stated. This is an obvious point, but often ignored or not explicitly stated (Noon 2003). Data are too often collected for data’s sake without a notion of how the data will be used or in support of what objective. This relates back to the importance of priority setting.

Once an objective has been clearly defined, the study and sampling design can go forward. A target population and sampling frame need to be defined, as well as a plot design, enumeration method, sample size, and plot selection and reselection strategies (Thompson *et al.* 1998). The target population is the population of interest and the sampling frame should, ideally, include the entire target population. If the sampling frame and target population do not coincide, then adjustments to the inference must be made accordingly. In essence, all possible members of a target population must be open to being sampled with a probability > 0 . Plot designs (i.e. size and shape) are study specific, and have important influences on the estimation process. Factors such as the life history of the animal in question, available resources, and logistics will influence the plot size and shape. With preliminary variance estimates, optimal sample sizes can be computed (Thompson *et al.* 1998, Thompson 2002) and the distribution in space of the organism will affect the plot selection and reselection strategy. Although these guidelines need to be addressed in every survey, the details will be situation-specific. In designing a sample survey, the design team needs to consist of people with knowledge of the organism, knowledge of the logistics involved, and knowledge of statistical design and analysis. With the absence of any one of these knowledge areas, the possibility of failure increases.

I will further explore two areas where wildlife sample surveys have particularly failed in the past and which affect numerous aspects of the design and analysis; namely the issues of accounting for detection probabilities in estimation and the use of formal probability-based sampling schemes (Skalski and Robson 1992, Skalski 1994, Yoccoz et al. 2001, Pollock et al. 2002, Sauer 2003).

DETECTION PROBABILITIES

In enumerating wildlife, the issue of detection probabilities arises. This problem, while not unique to wildlife studies, is most common in wildlife studies. Introductory college statistics classes generally do not cover this subject matter because in most other fields (e.g. business, engineering) detection probabilities are not a major concern. Thus many wildlife biologists are not exposed to this issue in introductory college classes and are ill-prepared to deal with it. The problem arises because wildlife are naturally selected to avoid predators, are generally mobile, and thus often not easy to detect. Few survey methods permit the detection of all individuals in an area (i.e. a census). In essence, when wildlife are counted, the raw count is not the true abundance, but some unknown proportion of the true abundance (i.e. some individuals are not detected, Lancia et al. 1994). This relationship can be expressed mathematically by $C = Np$ where C represents the count statistic, N represents the true abundance, and p represents the probability of detection. Only if $p = 1$ does the count statistic equal true abundance (the parameter ecologists are generally most interested in). Detection probabilities can vary over time and space and generate patterns falsely ascribed to a biological cause, as well as hide patterns of biological interest. Detection probabilities are problematic and are often called ‘nuisance parameters’ because of their confounding effect on estimates of parameters of interest (e.g. abundance).

However, by rearranging the above equation such that $\hat{N} = \frac{C}{\hat{p}}$ an estimate of abundance can be obtained if an estimate of p is available. Many methodologies are available to estimate detection probabilities associated with count statistics (discussed below). If we rely on a count statistic without correcting for detection error, we are essentially using an index – an *ad-hoc* estimator of a parameter of interest. The estimators used should be based upon statistical theory, such as maximum likelihood theory, and not subjective intuition as are *ad-hoc* estimators. Too frequently indices with no theoretical underpinnings are relied upon (Anderson 2001, Yoccoz et al. 2001, Pollock et al. 2002, Sauer 2003, White *in press*) and uncertainty (partially due to detection error) in these estimators is ignored with no confidence intervals given. Often, the use of indices is justified with the claim that alternative methods are impractical, expensive, or have ‘too many assumptions’. Notably, in general, an index will have stricter assumptions, although often unstated, than a direct estimator. With recent advances in estimating detection probabilities, and the inherent drawbacks of *ad-hoc* estimators, focusing on indices as default-monitoring metrics is inadvisable. If indices are relied upon, underlying assumptions should be verified with calibration studies. Too often an index is used, a pattern identified in the collection of index values, and a statement made to the effect of ‘not trusting the index’ or ‘needing better estimation techniques to see if the pattern is real’ (e.g. Yen *et al.* 2004).

The misuse of indices is broader than in relation to detection probabilities. Such misuse is facilitated by often-used jargon and ill-defined concepts. Biologists frequently mislead themselves by using ‘habitat’, ‘umbrella’, ‘indicator’, ‘keystone’, or, more recently, ‘highly interactive’ species as indices to species richness, diversity, or some notion of ‘ecosystem health’ (Simberloff 1998, Andelman and Fagan 2000, Soule et al. 2003). These concepts are often employed in lieu of estimating the true parameter of interest (e.g. species richness) because estimating the parameter is thought to be too difficult or expensive. Sometimes monitoring plans have too many objectives, and indices are used as shortcuts. However, the cost associated with the uncertainty surrounding the index is unrecognized. The use of such short-cuts, without a careful calibration to the parameter of interest, is short-sighted and wasteful (Chase et al. 1998, Rubinoff 2001, Gu and Swihart 2003). Monitoring programs should avoid the use of indices as much as possible. When indices are used, assumptions should be carefully stated and, to the best of our ability, verified, and the index calibrated periodically.

When instituting a monitoring program, after setting objectives, comparative methodological studies should be performed to articulate the precision and bias of possible estimators as well as associated costs. Such information will lead to an optimal, defensible, sampling plan. If, however, little confidence can be had in available estimators and methodology, I contend either more should be invested in the situation or else effort should be allocated where it can be effective.

Recent methodologies incorporating detection probabilities

General references for estimating abundance and associated demographic parameters, as well as detection probabilities are Borchers et al. (2002) and Williams et al. (2002). Below I focus on recently developed methodologies and briefly describe distance sampling (Buckland *et al.* 2001), a number of methods utilizing a capture-recapture (resighting) framework (Nichols et al. 1986, Pollock et al. 1990, Nichols et al. 2000, Farnsworth et al. 2002, Nichols and Hines 2002), novel analyses focused on presence/absence data (MacKenzie et al. 2002, MacKenzie et al. 2003, Royle and Nichols 2003), as well as on community dynamics (Nichols *et al.* 1998a, b).

Distance sampling

Distance sampling (Burnham et al. 1980, Buckland et al. 2001), has received recent interest (Rosenstock *et al.* 2002) and is one methodology incorporating detection probabilities into estimates of abundance/density. This methodology is predicated on being able to detect all individuals on a transect line (or at a point) and being able to measure the perpendicular distance of each individual detected off of the transect line. Individuals located away from the transect line are, generally, harder to detect and a ‘detection curve’ or ‘probability density function’ can be fit to these data with the probability of detection decreasing from the transect line to areas away from the transect line. With a model of the detection function, corrected densities can be calculated. In using this methodology, the assumption that all individuals directly on the line are detected with $p = 1$, is critical. Either this assumption must be valid or a correction factor must be incorporated. Field trials and simulation are important first steps in employing this methodology. Another important design feature, but often overlooked, is the random placement of transect lines in the target area of interest. Results from transects along

roads can not be extrapolated to other areas due to the possibility of an underlying pattern associated with the road coinciding with the detection function. The freeware Program Distance (<http://www.ruwpa.st-and.ac.uk/distance/>) and periodic workshops are available to facilitate the use of this methodology.

Capture-recapture based methods

Many methodologies based on capture-recapture (or resight) approaches (Borchers *et al.* 2002, Williams *et al.* 2002) are available, and I assume readers are familiar with classical approaches to estimating abundance such as have been implemented in Program CAPTURE (Otis *et al.* 1978, White *et al.* 1982), or have access to the appropriate literature. I will focus on recent developments and new twists on historical methods I think may be unfamiliar to some and possibly used more broadly.

Capture-recapture of immobile objects

Capture-recapture methods have been mostly applied to mobile organisms. However, many immobile objects (such as bird nests) may be cryptic and difficult to detect. Nichols *et al.* (1986) used a capture-recapture approach to estimate the number of dove nests at a study site in Texas. I have not seen this methodology used often in this way, but I think the applicability may be broad. In their study, plots were delineated and a series of observers searched plots independently and marked nests by placing an observer-specific colored marker in each nest found. An encounter history for each nest was generated and analyzed using closed-capture models to estimate the number of nests as well as detection probabilities. Possible applications could be with nests, territories, or slow moving organisms, such as tortoises.

Double observer methods

The recently proposed double-observer approach (Nichols *et al.* 2000) has been used in a bird point-count setting with two people visiting each point. One observer (the secondary observer) tallies the birds detected by the primary observer, as well as additional birds detected by him/herself, but not detected by the primary observer. The primary and secondary observer switch roles periodically (e.g. every other point) and from these data (birds detected by one or both observers) detection probabilities can be calculated as well as abundance. This is an observation based approach, with no handling of animals needed. Keeping track of individual birds and avoiding double counting is crucial in this methodology.

Auditory removal methods

Farnsworth *et al.* (2002) modified a ‘removal’ protocol and estimator for use in an observational setting. This analysis is similar to the analysis used to estimate abundance in studies where fish are removed with a series of passes, or small mammals are removed using snap-traps. This methodology has utilized a single observer in a bird point-count situation, assumes the observer can identify and track individual animals and there is little movement of animals during the survey period. The observer notes when each individual is detected and ‘auditorily removes’ individuals from the population (i.e. each individual is first detected at a certain time and never double counted). The new twist is no animals are actually handled and removed. Classic capture-removal estimates can be used with

these data. Extending this methodology to an auditory ‘mark-recapture’ should be feasible.

Estimating population growth (λ)

Population rate of change (λ) is a parameter many monitoring plans focus upon. Population rate of change can be defined as the relative change in population size between two time periods such that $\lambda = N_{t+1} / N_t$. If λ is > 1 , the population is growing and if $\lambda < 1$ the population is declining. Abundance estimates derived from closed mark-recapture analyses or from distance sampling can be used in this calculation. Alternatively, estimates of survival, transition rates, and fecundity can be incorporated into a population matrix model and an asymptotic λ calculated (Caswell 2001). A third approach, (Pradel 1996, Nichols and Hines 2002) recently developed, estimates a realized λ , not projected into the future, from open mark-recapture data, such as are collected in a Jolly-Seber setting (Jolly 1965, Seber 1965) to estimate apparent survival. This is an exciting development in the mark-recapture field and one that will see much use in the future.

Robust design

The open mark-recapture setting can be combined with closed mark-recapture data (with a series of closed sampling periods [secondary periods] constituting a single open sampling period [primary period]) into a ‘robust design’ (Pollock et al. 1990, Kendall et al. 1997). A ‘robust design’ provides many sources of information to estimate population size, population rate of change as well as the ability to account for temporary emigration in which animals may not be available for capture in some years (i.e. birds not returning to the same area every year, tortoises aestivating under ground, etc.). In general, the robust design should be considered for any mark-recapture study because it provides the most information and flexibility in the analysis. The robust design has not yet seen wide use.

Barker model

Richard Barker (Barker 1997, 1999) developed methods to include incidental sightings into mark-recapture models. In most studies, there is an explicit season (assumed closed) in which data are collected. However, auxiliary information may be collected between seasons. For example, a banding study of eagles may take place in which eagles are banded (possibly with wing tags) and resighted every spring. The rest of the year, wildlife managers or the public at large may report sighting of the eagles. Previous to this model, these ancillary data were not used. With Barker’s modification these data can now increase precision of estimates and this model could have greater utility.

Presence-absence based methods

Due to difficulties in enumerating animal species, some monitoring plans focus on presence/absence data with an occupancy probability being the parameter of interest. In addition to monitoring efforts, estimation of occupancy rates and associated dynamics (i.e. extinction and colonization) from presence-absence data is fundamental to many

habitat models (Cabeza *et al.* 2004), and metapopulation studies (Hanski and Gilpin 1997). Such studies are interested in whether a defined ‘patch’ or ‘site’ is occupied by one or more individuals of a population or species. In general data are collected by visiting a number of sites and recording whether any individuals of interest are present (recorded as a 1) or not (recorded as a 0). Occupancy probability is often calculated as the proportion of sites that are occupied; extinction probability as the proportion of occupied sites at time t not occupied at time $t+1$; and colonization probability as the proportion of sites not occupied at time t occupied at time $t+1$. However such presence-absence data, and resulting estimates, can be confounded by detection error; namely that a recorded ‘absence’ may in fact be a ‘non-detection’ of available individuals, and not a true absence. Such data should more precisely be referred to as detection/non-detection data and not presence-absence data. Using such data with naïve estimators will most likely result in under-estimates of occupancy and over-estimates of extinction probabilities. If a detection probability can be calculated, then unbiased estimators of occupancy, extinction and colonization can be derived. Mackenzie *et al.* (2002, 2003) are the seminal papers describing such estimation methods. Practically, these methods share much in common with the classic closed population abundance estimator and the ‘robust design’ (Pollock *et al.* 1990). Below I will elaborate on the estimation of occupancy probabilities since these methods are among the more recent and I believe will have wide utility in monitoring programs.

Occupancy probability

Underlying assumptions in estimating occupancy in a single season or year are 1) surveyed sites are occupied by the species of interest throughout the duration of the study, with no sites becoming occupied or unoccupied during the survey period (i.e. the system is closed), 2) species are not falsely detected, but can remain undetected if present, 3) species detection at a site is assumed to be independent of species detection at other sites. For example a number of ponds may be visited to assess occupancy rates of salamanders. The surveys take place during a short period during the breeding season when ponds (sites) can be assumed to be occupied. Each pond is visited once a day for five days (or alternatively five locations within a pond could be sampled) and whether any salamanders are present at each pond is recorded. Data for a single pond (i) can take the form of an ‘encounter history’ such as:

01010

Where ‘0’ signifies a non-detection and a ‘1’ signifies a detection. Since salamanders were detected at least once, we assume the site was occupied across all five sampling occasions, but not detected on sampling occasions 1, 3 and 5. If we denote ψ as occupancy probability and p as detection probability for period i , we can designate the likelihood for the above encounter history as:

$$\psi(1-p_1)p_2(1-p_3)p_4(1-p_5).$$

Non -detection of a species does not imply absence. An encounter history of

00000

could easily occur. This is an important difference when compared to the classical abundance estimation models in which such encounter histories are never observed. In the occupancy arena, the site is the observation unit, not the individual animal. The probability statement for this encounter history would be

$$\psi(1 - p_1)(1 - p_2)(1 - p_3)(1 - p_4)(1 - p_5) + (1 - \psi)$$

read as the site was either occupied by salamanders but these salamanders were not detected over the five sampling occasions (the term to the left of the '+') or the site was not occupied (the term to the right of the '+').

The model can be extended to include covariates. Note, however that because occupancy does not change over time, appropriate covariates for occupancy will be time constant and site-specific while covariates for detection can be time varying. This is in contrast to classic mark-recapture models where time constant individual covariates can be used to model p and c using a Huggins model (Huggins 1989, 1991) but time-varying covariates can not be used (remember in the classic case individuals are not observed on all occasions and thus time-varying covariates would not be available).

Extinction and colonization probabilities

The occupancy model has been extended to estimate extinction and colonization over multiple years using a 'robust design' approach (MacKenzie *et al.* 2003). If our salamander pond was visited three years in a row, the encounter history may be

01010 00000 10111

with the three sets of five digits representing the three years. Occupancy rates for each year are calculated as well as extinction and colonization between years. In year two no salamanders were detected. This could represent an extinction event between year 1 and year 2, or it could represent a series of non-detections during year 2. The multinomial for this encounter history can be tediously written out. Interested readers should consult (MacKenzie *et al.* 2003).

Species co-occurrence

The occupancy model has also been extended to investigate whether the occupancy of one species influences the occupancy of a second species (i.e. co-occurrence). This development is among the most recent and is detailed in MacKenzie *et al.* (2004). This extension may be very useful in investigating the effect of one species on another (i.e. do prairie dogs influence the occupancy of grassland birds? Does an invasive species affect the occupancy of species x?). The models of MacKenzie *et al.* (2004) even deal with situations where the presence of one species influences the detection probability of the other.

Estimating abundance from presence-absence data

Royle and Nichols (2003) investigate the relationship between ‘patch’ or ‘site’ level detection probabilities and individual detection probabilities. If p_s represent the site-level detection probability and p_i represents an individual detection probability then the probability that a species was not detected at a site ($1 - p_s$) is equal to the probability that an individual was not detected raised to the number of individuals (n) at the site $(1 - p_i)^n$ or:

$$1 - p_s = (1 - p_i)^n .$$

If a site-level detection probability can be obtained and an individual detection probability can be modeled, a latent estimate of abundance can be obtained from detection/non-detection data.

Community-level parameters (species richness, extinction, turnover)

Managers may be interested in a community of animals (e.g. grassland birds) rather than any particular species and would be focused on community-level parameters (i.e. species richness, extinction and turnover). Methodologies have been developed to estimate species richness and associated metrics taking into account probabilities of detection (Boulinier et al. 1998, Nichols et al. 1998a, b). Above I outlined the importance of recognizing that not all individuals are detected when estimating population size. Similarly, when estimating the number of species in a community (i.e. species richness), not all species will be detected. Estimates of species richness need to incorporate detection probabilities. This methodology is analogous to estimating abundance except the focus is now on the number of species and not the number of individuals. A series of surveys, over a closed time period, are conducted within a plot. These surveys are analyzed similarly to a closed, mark-recapture data except the encounter histories are now of whether an individual species was detected on each survey or not, rather than whether an individual animal was detected. By combining series of surveys over years a ‘robust design’ result which allows the estimation of extinction and turnover rates (Nichols *et al.* 1998b). I believe focusing on community-level parameters may be useful in many situations, especially for groups of low-priority species. A community, or guild, can be arbitrarily defined and such a group may be ‘low priority’, grassland birds, for example. Further discussion of monitoring community-level parameters is given by Yoccoz et al. (2001).

The above discussion of enumeration methods is not meant to be comprehensive, but a quick review of recent developments that may be useful in a monitoring context and to illustrate the point that many methodologies have been recently developed. Another common reason for not estimating detection probabilities is that an individual data set or year is often felt to be inadequate. Beyond addressing these deficiencies in study design, modern software also allows modeling of parameters across data sets, such that each year’s data does not have to be completely adequate to estimate detection probability by itself (White *in press*). Many field researchers have complained about the lack of methodologies to estimate parameters of interest well. The statisticians have responded. The onus is now on the field researchers to incorporate these new methodologies into

field protocols and return to the quantitatively minded with advice on what worked and did not work so additional progress can be made.

SAMPLING

In some instances resources are available to conduct a total count, but a sample will provide enough information, and there is no need to spend additional resources conducting a total count. More often, resources are not available to conduct a census (even if possible) or visit all locations in an area of interest, and a sample must be relied upon to make inference to a larger population.

To make proper inference to a target population, the population needs to be sampled appropriately. The ‘integrity of the sample’ is of utmost importance (Stuart 1984). A sample must be collected under a probability-based scheme and proper estimators used. Essentially every individual in a population must be potentially sampled. If this is not the case, then there will be no statistical inference to a larger population. To be sampled, a target population needs to be separated into a list of sampling units. This list is often referred to as a sampling frame (Thompson 2002). Sampling units from this frame need to be chosen in a probabilistic manner. Below I will review standard sampling schemes as well as some newer developments I think may be useful. Providing estimators is beyond the scope of this report, but can be found in appropriate references (e.g. Cochran 1977, Thompson 2002).

Simple random sampling

Simple random sampling (SRS) is the most common starting point in a beginning statistics class and covers the situations in which every sampling unit has the same probability of being included in the sample. This sampling scheme is the most familiar and widely used. While data from this scheme are flexible in how they can be used, precision of estimates can be greatly improved with other schemes (discussed below). Overton and Stehman (1996) suggest SRS be used for long-term environmental monitoring designs due to the inevitable sampling frame update (e.g. associated with a range extension) and this equal-probability design permits the greatest adaptability and flexibility in terms of alteration.

Stratified sampling

Stratification can greatly increase precision. In a stratified sampling scheme, similar sampling units are grouped into strata. If a forest bird survey was being designed, possible strata may be associated with forest stand age. By grouping similar sampling units together, the variance within each stratum is lower than the variance of all units across all strata, and the combined variance from all strata can be lower than the overall variance from a SRS, resulting in an overall increased precision. Stratification also lends itself to testing hypotheses concerning groups of sampling units. There is very little risk (i.e. stratification generally will not result in a larger variance than SRS), and much to be gained by stratifying a particular survey at a point in time. However, if strata definitions change over time, the covariance structure of sampling points through time can present problems in analyses at levels lower than the entire target population.

Cluster sampling

Cluster sampling works best in situations of similar variances within clusters (vs. different variances within strata in stratification). Cluster sampling is often employed when not all sampling units can *a priori* be identified at the level of a sampling unit, or to economize the logistics of sampling. In essence a ‘cluster’ of sampling units is selected, and once a cluster is selected every unit within that cluster is sampled. A classic example of cluster sampling is the sampling of household adults. It may be difficult to obtain a list of every person in a city, but a list of street addresses, or city blocks, is relatively easy to obtain. For example a sample of city blocks (clusters) may be chosen, and within each block, every person (sampling unit) is surveyed. If city blocks do not differ markedly from each other, or can be stratified into similar groups, then cluster sampling can be very efficient. In the wildlife field, cluster sampling may be employed when logistics are challenging (especially in terms of travel to sampling locations). It can be useful to define clusters of sampling locations, travel to a cluster and sample all units within the defined cluster.

Systematic sampling

Systematic sampling is useful when a spatially balanced sample is desired (i.e. for a spatially explicit habitat model). A sample is chosen systematically from a list of all sampling units. For example, starting with a random location, every 5th unit could be sampled. Systematic sampling has the appeal of an ‘even’ spread across the sampling frame and may be especially good for spatial models. However, if there is an underlying oscillation or periodicity in the parameter of interest coinciding with the systematic sampling period, then biased estimates can result. Systematic sampling may have increased logistic costs if increased travel is needed.

Adaptive sampling

Adaptive sampling (Thompson and Seber 1996, Thompson 2002) is a new development focused on surveys for rare and clustered (e.g. colonial) organisms. If most survey locations contain zero organisms of interest, intuitively, when an organism is found, the organism is more likely to be in the surrounding area. Adaptive sampling is a framework for searching the surrounding area. Essentially, starting with a SRS (or other design), once an organism of interest is detected, the sampling units around the first unit are also searched until no more organisms are detected. Inclusion probabilities (the probability of a particular sampling unit being included in the sample) are used in calculations. Adaptive sampling has promise. However my experience in the field suggests increased logistical costs – especially with searching and locating surrounding sampling units once an organism has been found. This cost was, in part, due to there being no stopping rule – all units had to be searched until no more organisms were found. Recent developments (Thompson - pers. comm.) now incorporated stopping rules. However, unless the organism is highly clumped or aggregated, adaptive sampling will not improve efficiency.

Double sampling

Double sampling (not to be confused with double-observer) designs are when auxiliary information is taken on an initial sample and then a second sample is taken with both the auxiliary variable and variable of interest being observed (Thompson 2002).

Often the second sample is a subset of the first. Specifically for abundance, a large number of sampling units may be visited with a quick enumeration taking place (e.g. a point count for birds). A smaller number of the units are visited again and a more intensive enumeration method is used (e.g. estimating detection probability with one of the above methods). These more intensive plots are used to correct the estimates from the large sample of plots. Pollock *et al.* (2002) suggest this methodology receive more use in large-scale surveys. Examples of the use of double sampling in large scale surveys are the monitoring of duck breeding populations (Smith 1995) and monitoring breeding shorebirds in the Arctic (Bart and Earnst 2002).

Ratio estimation

Ratio estimators, have some similarity to double sampling, and use auxiliary data from all sampling units in the sample frame, regardless of whether or not the unit is actually sampled. The sample count from the chosen plots is related to an auxiliary variable (e.g. forest cover in a forest breeding bird survey) from the chosen plots using a ratio, difference, or regression estimator. This relationship is then extrapolated to all sampling units with the use of the auxiliary variable. Ratio estimation can markedly improve precision. Samples using ratio estimators are design and model-based and need to be chosen with care. Ratio estimation has not seen wide use in the wildlife field (for an example see Bowden *et al.* 2003), but I think in the future it will. With the advent of GIS technology, obtaining meaningful auxiliary information for an entire sampling frame is much easier than in the past; facilitating the use of ratio estimators.

Dual frame sampling

Dual frame sampling (Haines and Pollock 1998) is a useful design when a number of object locations are known and are somewhat permanent. For example, Haines and Pollock (1998) suggested this scheme for monitoring the number of bald eagle nests. In this example a number of bald eagle locations are known. However, some nest locations are not known. In essence this information can be divided into two sampling frames – a list frame of all sampling units with historically known nests, and an area frame of all other possible sampling units, containing an unknown number of nests. Sampling units can switch frames from year to year. If nests are found in sampling units in the area frame, the next year those sampling units can be added to the list frame. If an eagle nest was blown down, or destroyed, that particular sampling unit could switch to the area frame. Sampling effort can be divided between both frames. Equal effort does not have to be allocated to each frame; more effort may be spent surveying the list frame, but some effort needs to be devoted to the area frame. Information from both frames is incorporated to obtain an overall estimate for the target area. I think this sampling scheme could be useful in many situations, such as the monitoring of prairie dog colonies, the monitoring of colonial birds, as well as the monitoring of raptor nest locations.

Multi-stage designs

Multi-stage designs are more complicated designs and may combine a number of the different sampling schemes above. For instance, cluster sampling may be appropriate over subsets of a sampling frame, but not the entire sampling frame due to inherent

variance associated with habitat types. In this case the sampling frame may first be stratified by habitat type and then within each stratum a sample of clusters could be chosen in which each primary sampling unit was sampled. This would be a multi-stage sample with the first stage being the stratification and the second stage being the cluster sample. Multi-stage designs allow much flexibility and can be tailored to the biology of the animal as well as to available resources. I also note that the above categories are not mutually exclusive. For example many people view adaptive sampling as a 2-stage sampling design.

Rotating panel

When reselecting plots (as in a multiple-year survey), a rotating panel design (Manly 2001) may be useful. In plot reselection the same plots can be revisited, in which the focus would be on temporal variation through time, or different plots can be selected, focusing on temporal and spatial variation through time. A rotating panel design is somewhere in between these two reselection schemes, with some of the same plots and some new plots being visited in subsequent years. The below table (Table 1) illustrates the concept in which four plots are visited each year, consisting of three plots from the previous year and one new plot. This plot reselection design represents a trade-off in terms of investigating changes over time as well as in distribution.

Table 1. Rotating panel schematic.

Plot #	Year				
	1	2	3	4	5
1	x				
2	x	x			
3	x	x	x		
4	x	x	x	x	
5		x	x	x	x
6			x	x	x
7				x	x
8					x

In this section I briefly described issues related to the design and analysis of survey data. This section is not meant to give a full understanding of these methodologies and issues, but to identify and briefly explain some recent advances. Experts in the field of experimental and survey design and estimation should be consulted in incorporating these methodologies and helping with optimal sample selection and allocation. Few monitoring programs account for both detection and spatial sources of error, but I think to do so is the hallmark of a good monitoring program.

REVIEW OF SELECTED MONITORING PLANS

Olsen et al. (1999) and McDonald (2003) reviewed large-scale monitoring plans, however, most of the reviewed plans did not focus on animals. Below, I will review large-scale monitoring efforts focused on animals with respect to the philosophy taken, and attention to error associated with detection and sampling in space.

CHRISTMAS BIRD COUNT (CBC)

Description and underlying philosophical approach

The Christmas Bird Count (National Audubon Society 2003) started in 1900 and is the longest running avian monitoring program in North America. Each winter, generally on a single day between 15 December and 5 January, all birds within a 15 mile circle are counted by teams of observers. The number of CBC circles varies annually as does the number of observers. The CBC is used to track trends in winter bird populations and is conducted with an early-warning philosophical approach. Sauer (2003) reviewed the CBC and found it deficient in terms of incorporating detection probabilities and addressing sampling error.

Detection error

Incorporating detection error into estimates of abundance is not considered in the CBC design.

Sampling error

Sampling error is not considered in the CBC design. The count locations are defined by the observer and not chosen with a probabilistic sampling scheme. No sampling frame with a finite number of sampling units is considered.

NORTH AMERICAN BREEDING BIRD SURVEY (BBS)

Description and underlying philosophical approach

The North American Breeding Bird Survey (BBS, USGS 2001a) was initiated in 1966 as a long-term, large-scale avian monitoring program to track the status and trends of North American birds with the hope of detecting significant declines so the causes of the declines can be detected and ameliorated (USGS 2001a). Each year, on ~4100 survey routes throughout the United States and Canada, observers collect observation data on birds. Each survey route is located along a 24.5 mile stretch of road with survey stops located at 0.5 mile intervals. At each stop every bird seen or heard within a 0.25 mile radius, over three minutes is counted. The data are analyzed and used as an index to population abundance to track population trends at various scales (Sauer et al. 2003). The BBS was conceived under an early-warning framework. Sauer (2003) reviewed the BBS and found it deficient in terms of incorporating detection probabilities and addressing sampling error.

Detection error

The BBS is not designed to collect data to estimate detection probabilities. Although sophisticated models have been developed to address variation in detection probabilities associated with a single source of variation, observers (Link and Sauer 1998), much controversy exists over appropriate analysis methods (see Sauer et al. 2004 and Bart et al. 2004) and the meaning of this index. If the BBS design were changed to estimate detection probabilities directly, this controversy would be lessened. BBS data have been used to estimate community-level parameters incorporating detection probabilities (Nichols et al. 1998a, b, Doherty et al. 2003), however the BBS was not designed for this purpose and if these were metrics of interest, alterations to the design should be made.

Sampling error

Survey routes are located on secondary roads and picked with a sampling scheme based on 1-degree latitude-longitude blocks, and availability of qualified observers. There are many more routes in the east than the west (due to concentrations of observers) and rugged mountains and marshes are poorly surveyed (due to lack of roads, Bystrak 1981). Although the routes are widespread, statistical inference to any other areas besides the sampled secondary roads is unjustified. Habitat among roads is different, and changes differently than other habitats. The BBS could be improved by incorporating a better probabilistic sampling scheme and incorporating off-road survey routes.

MONITORING AVIAN PRODUCTIVITY AND SURVIVORSHIP (MAPS)

Description and underlying philosophical approach

The Monitoring Avian Productivity and Survivorship (MAPS) program (Desante et al. 1995, DeSante et al. 1999) attempts to collect demographic data to identify causal factors in population trends from this and other avian monitoring programs (i.e. BBS). This program is mostly run with the use of volunteers who operate banding stations consisting of 10 mist nets over an 8-ha area in which netting takes place once a day,

every 10 days for 10 - 12 periods. Species, age, and sex data are the primary data collected. These data are analyzed to provide an index of population size and change, as well as productivity indices and survival estimates. An early-warning philosophy underlies the program.

Detection error

Detection error for survival rate estimation is taken into account within mark-recapture analyses. Detection error is typically not dealt with for inferences about productivity (based on raw age ratios of netted birds).

Sampling error

One major problem with the MAPS program is the lack of a probabilistic sampling plan. Volunteers generally choose the banding sites and inferences to areas beyond the banding sites are unwarranted.

NORTH AMERICAN WATERFOWL MANAGEMENT PLAN (NAWMP)

Description and underlying philosophical approach

International treaties in North America mandate the protection of migratory waterfowl populations as well as providing hunting opportunities. Many waterfowl monitoring activities have developed into, arguably, the best large-scale monitoring program in the world (Nichols et al. 1995). These monitoring activities include banding and dead recovery of ducks, parts collections and flying surveys to count waterfowl and pond numbers. In the past, the analysis of these data have focused on retrospective analyses (Nichols et al. 1995), but more recently the waterfowl program has taken an active adaptive management (early-control) approach in which reducing uncertainty in the population biology of ducks, and maintaining management goals are dual objectives of the program (Johnson et al. 1993, Nichols et al. 1995, Kendall 2001, Williams and Nichols 2001).

Detection error

Detection error has been accounted for in the survival analysis with the development and use of band recovery models (Brownie et al. 1985) and in population counts through the use of double sampling.

Sampling error

Sampling error is taken into account with the use of a stratified sampling plan in the aerial population surveys.

USDA NORTHWEST FOREST PLAN

Description and underlying philosophical approach

The USDA Northwest Forest Plan was developed in 1994 to allow resource extraction (primarily timber) over the long term without degrading the health of the forest and environmental resources (Ringold et al. 2003) and focused on a large number of terrestrial, aquatic, animal, plant, and environmental objectives. Although adaptive management is recognized in the underlying philosophical approach, and conceptual models are employed to aid in identifying monitoring metrics, the underlying approach is

based on ‘early-warning’ with the focus on identifying trends in indicators, as well as triggers for further research and action (Lint et al. 1999, Madsen et al. 1999, Noon 2003). The two terrestrial vertebrates receiving the most attention have been the Northern spotted owl (*Strix occidentalis caurina*, Lint et al. 1999) and the marbled murrelet (*Brachyramphus marmoratus*, Madsen et al. 1999). I will focus on the spotted owl. Millions of dollars have been spent on monitoring spotted owls and periodic assessments of owl population and vital rate trends are made (i.e. Anthony et al. *in press*). Even with this immense effort, results are somewhat ambiguous (Anthony et al. *in press*). An alternative approach to primarily focusing on trends, is to have trend assessment secondary to investigating questions of biological and management interest. For the spotted owl, one important question may be whether the encroachment of barred owls (*Strix varia*) is detrimental to spotted owls and whether the management action of barred owl removal is warranted and effective. I contend this information may be more valuable to meeting owl management objectives than a, somewhat, ambiguous estimate of trend.

Detection error

For the spotted owl, detection error is accounted for in the design and analysis – primarily through a mark-recapture approach.

Sampling error

A probabilistic sampling strategy and sampling frame is not used in monitoring owls. However, a large portion of the owl’s range (~50%) is included in the sample.

NORTH AMERICAN AMPHIBIAN MONITORING PROGRAM (NAAMP)

Description and underlying philosophical approach

With the advent of concern over the status of amphibian populations (McCoy 1994), an amphibian monitoring program was initiated (USGS 2001b). This program was initiated by the same USGS group (Status and Trends group at the USGS Patuxent Wildlife Research Center) that runs the BBS. Accordingly, the North American Amphibian Monitoring Program (NAAMP) shares much in common with the BBS. NAAMP is a roadside survey focused on collecting data to estimate an index of population trends in many amphibian species. Survey routes consist of 10 stops located primarily near wetlands and are visited, for a 5-minute period, in the evening during the early, middle and late breeding season to listen for a series of amphibian calling species. Each calling species is categorized with a calling index. NAAMP was developed under an early-warning framework.

Detection error

Like the BBS, NAAMP was not designed to explicitly incorporate detection error into the estimation process. Although no formal analysis has been conducted on the data set as a whole, I anticipate similar problems as those encountered with BBS analyses, and the results being focused on a poorly-defined index.

Sampling error

Similar to the BBS, NAAMP routes are conditioned upon roads, although routes are picked with a stratified, random block design (USGS 2001b). Statistical inference to

amphibian populations others than those associated with roadside ditches and wetlands will be unwarranted.

ENVIRONMENTAL MONITORING AND ASSESSMENT PROGRAM (EMAP)

Description and underlying philosophical approach

EMAP originally was a program to monitor the status and trends in condition of the nation's ecological resources and to make information available to regulators, managers, and policy makers (National Research Council 1995). EMAP was designed to survey much of North America using a rotating panel design and a 12,600 hexagon sampling frame (Uruquhart et al. 1993). Some of the criticisms of the original EMAP were having an inadequate justification for what was to be monitored (i.e. no conceptual model, prioritization or clear objectives), being mostly a retrospective program (i.e. an early-warning philosophy), and not having the resources needed (National Research Council 1995). EMAP has evolved into a series of regional programs focused on "developing the tools necessary to monitor and assess the status and trends of national ecological resources" with a focus on aquatic systems and surrounding landscapes (Environmental Protection Agency 2004).

Detection error

Many of the current regional EMAP programs are not focused on animal populations where detection error is of primary concern.

Sampling error

The original sampling plan was a sophisticated probability-based sampling plan based on a 12,600 40-ha hexagonal sampling frame and a rotating panel (Manly 2001). Sampling plans associated with the regional EMAPs are generally probability-based.

NATIONAL PARK SERVICE VITAL SIGNS

Description and underlying philosophical approach

The Vital Signs program of the National Park Service is a recent development in which national parks and monuments are grouped into 32 networks of similar geography and natural resources. Within these networks monitoring plans are currently being developed with the sharing of resources and expertise within networks. The National Park Service has compiled an impressive web site of guidance and resources (<http://science.nature.nps.gov/im/monitor/index.htm>, National Park Service 2004) to aid the development of these programs. The underlying philosophy is an early-warning philosophy with an emphasis on a human health analogy. The 'vital signs' to be monitored are to be "ideal indicators ... representing the overall health or condition of park resources, known or hypothesized effects of stressors, or elements that have important human values." Although conceptual models are relied upon, the possible dependence on uncalibrated indices and a lack of attention to possible management actions could be problematic.

Detection error

A wide range of approaches is considered being used across the 32 networks, some accounting for detection probabilities, some not. The National Park Service web site advises detection error be considered in designs and analyses.

Sampling error

A wide range of approaches are being used across the 32 networks, some using probabilistic sampling plans. The National Park Service web site advises the use of probabilistic sampling plans.

US FOREST SERVICE NATIONAL PROTOCOL FOR MONITORING VERTEBRATES

Description and underlying philosophical approach

The Forest Service is developing a national protocol for monitoring vertebrates. This plan is currently in draft form (Manley et al. 2004) and is focused on monitoring trends in the percent area occupied (i.e. not abundance) by multiple species utilizing presence-absence data. The plan is currently structured utilizing an early-warning framework.

Detection error

The plan currently advises attention to detection probabilities in data collection (i.e. multiple visits) as well as in modeling the percent of the survey area occupied.

Sampling error

A probabilistic sampling strategy and sampling frame is advised and based upon the Forest Inventory and Analysis (FIA) grid system. The FIA grid is a systematic hexagonal grid with each grid containing about 2360 ha.

SUMMARY

Table 2 summarizes the selection of monitoring programs with respect to the underlying philosophy, and attention to detection and sampling error. Most monitoring plans are based on an early-warning framework with the focus on detecting trends in organisms or indices of interest. Only the North American Waterfowl Management Plan is specifically based upon an early-control ('adaptive management') framework. Many (6/9) of the monitoring programs I reviewed incorporate, or speak to the ability to incorporate, detection error into estimates of interest. Notably, plans currently being developed (i.e. by the National Park Service and Forest Service) speak strongly about the need to incorporate detection probabilities into estimation. Fewer plans allow extrapolation of results to populations of interest through the use of a probability-based sampling design. This is an area needing further emphasis. The only plan incorporating detection probabilities into estimation, using a probabilistic sampling plan, and relying upon an early-control framework is the North American Waterfowl Management Plan. This plan is also arguably the best large-scale monitoring plan (Nichols et al. 1995), with all other plans having been subjected to, sometimes severe, criticism. These criticisms are focused on underlying philosophies, such as the National Research Council (1995) suggesting plans such as these should not be based on a retrospective philosophy. Sauer (2003), in reviewing the CBC and BBS suggests incorporating detection probabilities into the design of these surveys and not further relying on the current index, as well as more

explicitly addressing and defining the area being sampled. Sauer (2003) also suggests survey evolve with improving methodology. Criticisms are also levied against the lack of objectives, or justification of objectives. Many of the reviewed plans are based on multiple objectives, or multiple species.

Table 2. Summary of selected national monitoring plans focused on vertebrates.

Monitoring Program	Underlying Philosophy	Detection Error	Sampling Error
Christmas Bird Count	Early warning	no	no
Breeding Bird Survey	Early warning	no	restricted
Monitoring Avian Productivity and Success	Early warning	yes	no
North American Waterfowl Management Plan	Early control	yes	yes
Northwest Forest Plan (spotted owl)	Early warning	yes	no
North American Amphibian Monitoring Program	Early warning	no	restricted
Environmental Monitoring and Assessment Program	Early warning	possibly	yes
National Park Service Vital Signs	Early warning	possibly	possibly
US Forest Service National Protocol	Early warning	yes	yes

I believe these multiple foci are to the detriment of many of the plans, because few, if any, of the objectives are met well. One size does not fit all, and a single study design does not work for all objectives. Clear and specific thinking, objectives, study plans, and execution are needed. Too often multiple objectives lead to unclear thoughts, objectives, and poor execution. I contend it is better to do one, or a few, things well than many things poorly. Finally criticism is levelled against poor study designs.

I believe monitoring plans more closely aligned with an early-control philosophy, with clear objectives, robust study plans (i.e. attention to error associated with detection and sampling in space), and with good execution will further our knowledge base and lead to better management, as well as be defensible. A management agency should benefit from such a situation.

CONCLUSIONS AND RECOMMENDATIONS

Monitoring programs are becoming more prevalent and important in wildlife management and decision making. With increased demand for, and use of monitoring data, monitoring programs will come under increased scrutiny in terms of underlying philosophical tenets, as well as in terms of design.

Monitoring is generally structured using one of two philosophical constructs. The first construct, an early-warning framework, views monitoring of ecological variables analogously to monitoring human health. Under this framework monitoring is viewed as the first step in a four step process, with the goal of detecting trends in parameters of interest. The second construct, an early-control framework, views monitoring as an integral part of a program to learn about biological systems of interest and possible management actions. Although the first construct is more prevalent in monitoring programs, the second construct will have more efficient returns for a management agency.

Many monitoring plans are predicated on meeting multiple objectives. Unfortunately, one size does not fit all, and I believe inattention to specifying a limited

number of accomplishable objectives with available resources is a weakness in many monitoring programs. This implies an agency needs to prioritize objectives. This prioritization process can be difficult, and is at least as strongly influenced by values as by science. Prioritization can include differing levels of investment for different priority levels. The same effort does not necessarily need to be given to common species, as given to species having more concern. Prioritization scheme should be, at least partly, informed by the use of conceptual models, consultation with managers, biologists, and statisticians, and with a focus on manageable systems. Prioritizations (and monitoring plans) should also change over time.

No matter what philosophical approach is adhered to, more attention to study design is needed. Of the monitoring programs reviewed, few began to approach the rigor needed to estimate quantities, and to make inference to populations, of interest. Two areas where wildlife studies and monitoring programs often lack are attention to error associated with detection and the use of probability-based sampling schemes. Many methodological advances have been made in these areas and these methods should be considered and incorporated into study designs. Statisticians have responded to field ecologists demands for better estimators of abundance and vital rates. It is now time for field ecologists to incorporate these methods and return to statisticians with their experiences so further developments can occur.

In summary:

- 1.) Monitoring should primarily be focused on informing management actions and secondarily on providing trend information. An early-control philosophy should be invoked more often. For a management agency, this focus on directly management actions seems especially warranted.
- 2.) Management agencies should have a list of priority species, metrics, or questions of management interest. This list of priorities should drive resource allocation. Although many activities can be done, all activities can not be done due to constraints.
- 3.) In designing a monitoring plan, the design team should consist of people with knowledge of the organism, knowledge of the logistics involved, and knowledge of statistical design and analysis. Partnerships and mutual respect within and between these knowledge bases should be enhanced.
- 4.) Monitoring programs should be designed with clear objectives and attention to sample integrity through the use of a probabilistic sampling scheme and attention to detection error.
- 5.) The use of indices, especially uncalibrated indices, should be avoided.
- 6.) Many new methodologies, including observation-based methods and methods focused on presence/absence data and community-level parameters, have been developed and such methods should be considered.
- 7.) Shortcomings of other monitoring plans should be avoided with a focus on aspects that have made plans successful. The most successful monitoring plan (e.g. NAWMP) is following an early-control framework, has clear objectives, management actions and addresses error associated with detection and sampling schemes in the design and analysis.
- 8.) Clear communication with contracting entities, specifying priorities and survey methods (if applicable) is needed.

9.) Monitoring plans should be expected to evolve as better methods are developed and as priorities change.

ACKNOWLEDGMENTS

I thank Ken Burnham, Bill Link, Jim Nichols, John Sauer, Barry Noon, Tom Remington, and Gary White, for discussions concerning monitoring activities. I thank the Colorado Division of Wildlife for funding this report.

LITERATURE CITED

- Andelman, S. J., and W. F. Fagan. 2000. Umbrellas and flagships: Efficient conservation surrogates or expensive mistakes? *Proceedings of the National Academy of Sciences of the United States of America* **97**:5954-5959.
- Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. *Wildlife Society Bulletin* **29**:1294-1297.
- Anthony, R., et al. *in press*. US Forest Service report on the status and trend of the northern spotted owl.
- Barker, R. J. 1997. Joint modeling of live-recapture, tag-resight, and tag-recovery data. *Biometrics* **53**:666-677.
- Barker, R. J. 1999. Joint analysis of mark-recapture, resighting and ring-recovery data with age-dependence and marking-effect. *Bird Study* **46**:82-91.
- Bart, J., B. Collins, and R. I. G. Morrison. 2004. Estimating trends with a linear model: reply to Sauer et al. *The Condor* **106**:440-443.
- Bart, J., and S. Earnst. 2002. Double sampling to estimate density and population trends in birds. *Auk* **119**:36-45.
- Block, W. M., A. B. Franklin, J. P. Ward, Jr., J. L. Ganey, and G. C. White. 2001. Design and implementation of monitoring studies to evaluate the success of ecological restoration on wildlife. *Restoration Ecology* **9**:293-303.
- Borchers, D. L., S. T. Buckland, and W. Zucchini. 2002. Estimating animal abundance - closed populations. Springer, New York.
- Boulinier, T., J. D. Nichols, J. R. Sauer, J. E. Hines, and K. H. Pollock. 1998. Estimating species richness: The importance of heterogeneity in species detectability. *Ecology* **79**:1018-1028.
- Bowden, D. C., G. C. White, A. B. Franklin, and J. L. Ganey. 2003. Estimating population size with correlated sampling unit estimates. *Journal of Wildlife Management* **67**:1-10.
- Brownie, C., D. R. Anderson, K. P. Burnham, and D. S. Robson. 1985. Statistical inference from band recovery data - a handbook. U.S. Fish and Wildlife Service, Resource Publication 156, Washington DC.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling - estimating abundance of biological populations. Oxford University Press, Oxford.
- Burnham, K. P., D. R. Anderson, and J. L. Laake. 1980. Estimation of Density from Line Transect Sampling of Biological Populations. *Wildlife Monographs* **72**:1-202.
- Bystrak, D. 1981. The North American Breeding Bird Survey. *Studies in Avian Biology* **6**:34-41.

- Cabeza, M., M. B. Araujo, R. J. Wilson, C. D. Thomas, M. J. R. Cowley, and A. Moilanen. 2004. Combining probabilities of occurrence with spatial reserve design. *Journal of Applied Ecology* **41**:252-262.
- Caswell, H. 2001. *Matrix population models - construction, analysis and interpretation*, second edition. Sinauer Associates, Inc., Sunderland, Massachusetts.
- Caughley, G. 1994. Directions in Conservation Biology. *Journal of Animal Ecology* **63**:215-244.
- Chase, M. K., J. T. Rotenberry, and M. D. Misenhelter. 1998. Is the California gnatcatcher an indicator of bird-species richness in coastal sage scrub? *Western Birds* **29**:468-474.
- Cochran, W. G. 1977. *Sampling Techniques*, third edition. John Wiley and Sons, New York.
- Davis, G. E. 1993. Design Elements of Monitoring Programs - the Necessary Ingredients for Success. *Environmental Monitoring and Assessment* **26**:99-105.
- Davis, M. A., and L. B. Slobodkin. 2004. Restoration ecology: the challenge of social values and expectations. *Frontiers in Ecology and the Environment* **2**:44-45.
- Desante, D. F., K. M. Burton, J. F. Saracco, and B. L. Walker. 1995. Productivity indices and survival rate estimates from MAPS, a continent-wide programme of constant-effort mist-netting in North America. *Journal of Applied Statistics* **22**:935-947.
- DeSante, D. F., D. R. O'Grady, and P. Pyle. 1999. Measures of productivity and survival derived from standardized mist-netting are consistent with observed population changes. *Bird Study* **46**:178-188.
- Diffendorfer, J. E., and P. F. Doherty, Jr. 2004. Lifting Cassandra's curse. *Conservation Biology* **18**:600.
- Doherty, P. F., Jr., T. Boulinier, and J. D. Nichols. 2003. Local extinction and turnover rates at the edge and interior of species' ranges. *Annals Zoologici Fennici* **40**:145-153.
- Dougherty, J. 2000. Biodiversity and human health. <http://www.wms.org/biod/summary/>.
- Environmental Protection Agency. 2004. *Environmental Monitoring and Assessment Program*. <http://www.epa.gov/emap/>.
- Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. *Auk* **119**:414-425.
- Federal Judicial Center. 2000. *Reference manual on scientific evidence*, 2nd edition. Federal Judicial Center, Washington DC.
- Gibbs, J. P., H. L. Snell, and C. E. Causton. 1999. Effective monitoring for adaptive wildlife management: Lessons from the Galapagos Islands. *Journal of Wildlife Management* **63**:1055-1065.
- Green, R. E. 2002. Diagnosing causes of population declines and selecting remedial actions. Pages 139-156 *in* K. Norris and D. J. Pain, editors. *Conserving Bird Biodiversity - General Principles and their Application*. Cambridge University Press, Cambridge.
- Greenwood, J. J. D., S. R. Baillie, R. D. Gregory, W. J. Peach, and R. J. Fuller. 1995. Some New Approaches to Conservation Monitoring of British Breeding Birds. *Ibis* **137**:S16-S28.

- Gu, W., and R. K. Swihart. 2003. Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. *Biological Conservation* **116**:195-203.
- Haines, D. W., and K. H. Pollock. 1998. Estimating the number of active and successful bald eagle nests: an application of the dual frame method. *Environmental and Ecological Statistics* **5**:245-256.
- Hanski, I. A., and M. E. Gilpin, editors. 1997. *Metapopulation biology - ecology, genetics and evolution*. Academic Press, San Diego.
- Hellawell, J. A. 1991. Developing a rational for monitoring. Pages 1 to 14 *in* B. Goldsmith, editor. *Monitoring for Conservation and Ecology*. Chapman and Hall, New York.
- Huggins, R. M. 1989. On the statistical analysis of capture-recapture experiments. *Biometrika* **76**:133-140.
- Huggins, R. M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* **47**.
- Johnson, F. A., B. K. Williams, J. D. Nichols, J. E. Hines, W. L. Kendall, G. W. Smith, and D. F. Caithamer. 1993. Developing an adaptive management strategy for harvesting waterfowl in North America. **58**:565-583.
- Jolly, G. M. 1965. Explicit Estimates from Capture-Recapture Data with Both Death and Immigration-Stochastic Model. *Biometrika* **52**:225-247.
- Jones, S. L., T. R. Stanley, S. K. Skagen, and R. L. Redmond. 2003. Long-Billed Curlew (*Numenius americanus*) Rangewide Survey and Monitoring Guidelines. USFWS.
- Kendall, W. L. 2001. Using models to facilitate complex decisions. Pages 147-170 *in* T. M. Shenk and A. B. Franklin, editors. *Modeling in natural resource management*. Island Press, Washington.
- Kendall, W. L., J. D. Nichols, and J. E. Hines. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. *Ecology* **78**:563-578.
- Krebs, C. J. 1991. The Experimental Paradigm and Long-Term Population Studies. *Ibis* **133**:3-8.
- Lancia, R. A., J. D. Nichols, and K. H. Pollock. 1994. Estimating the number of animals in wildlife populations. Pages 215-253 *in* T. A. Bookout, editor. *Research and Management Techniques for Wildlife and Habitats*. The Wildlife Society, Bethesda, MD.
- Link, W. A., and J. D. Nichols. 1994. On the Importance of Sampling Variance to Investigations of Temporal Variation in Animal Population-Size. *Oikos* **69**:539-544.
- Link, W. A., and J. R. Sauer. 1998. Estimating population change from count data: Application to the North American Breeding Bird Survey. *Ecological Applications* **8**:258-268.
- Lint, J., B. R. Noon, R. Anthony, E. Forsman, M. G. Raphael, M. W. Collopy, and E. Starkey. 1999. Northern spotted owl effectiveness monitoring plan for the Northwest Forest Plan. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station General Technical Report PNW-GTR-440, Portland.
- Mace, G. M., and N. J. Collar. 2002. Priority-setting in species conservation. Pages 61-73 *in* K. Norris and D. J. Pain, editors. *Conserving bird biodiversity: general principles and their application*. Cambridge University Press, Cambridge.

- MacKenzie, D. I., L. L. Bailey, and J. D. Nichols. 2004. Investigating species co-occurrence patterns when species are detected imperfectly. *J Anim Ecology* **73**:546-555.
- MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization and local extinction when a species is detected imperfectly. *Ecology* **84**:2200-2207.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83**:2248-2255.
- Mackowiak, P. A., S. S. Wasserman, and M. M. Levine. 1992. A critical appraisal of 98.6 degrees F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich. *Journal of the American Medical Association* **268**:1578-1580.
- Madsen, S., D. Evans, T. Hamer, P. Henson, S. Miller, S. K. Nelson, D. Roby, and M. Stapanian. 1999. Marbled murrelet effectiveness monitoring plan for the Northwest Forest Plan. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, General Technical Report PNW-GTR-439., Portland.
- Manley, P. N., B. Van Horne, and C. Hargis. 2004. Multiple-species inventory and monitoring national protocol technical guide - draft - http://www.fs.fed.us/research/monitoring_vertebrate.html.
- Manly, B. F. J. 2001. *Statistics for environmental science and management*. Chapman & Hall, New York.
- McCoy, E. D. 1994. Amphibian decline - a scientific dilemma in more ways than one. *Herpetologica* **50**:98-103.
- McDonald, T. L. 2003. Review of environmental monitoring methods: survey designs. *Environmental Monitoring and Assessment* **85**:277-292.
- Montgomery, D. C. 2001. *Design and analysis of experiments*, 5th edition. John Wiley, New York.
- National Audubon Society. 2003. Christmas Bird Count. <http://www.audubon.org/bird/cbc/>.
- National Park Service. 2004. Monitoring Natural Resources in Our National Parks. <http://science.nature.nps.gov/im/monitor/index.htm>.
- National Research Council, 1995. Review of EPA's Environmental Monitoring and Assessment Program - Overall Evaluation. National Academy Press, Washington DC.
- Nichols, J. D. 1999. Monitoring Is Not Enough: On the Need for a Model-based Approach to Migratory Bird Management. *in* R. Bonney, D. N. Pashley, R. J. Cooper, and L. Niles, editors. *Strategies for Bird Conservation: The Partners in Flight Planning Process*. Cornell Lab of Ornithology, Cape May, NJ.
- Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998a. Estimating rates of local species extinction, colonization, and turnover in animal communities. *Ecological Applications* **8**:1213-1225.
- Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998b. Inference methods for spatial variation in species richness and community composition when not all species are detected. *Conservation Biology* **12**:1390-1398.

- Nichols, J. D., and J. E. Hines. 2002. Approaches for the direct estimation of lambda, and demographic contributions to lambda, using capture-recapture data. *Journal of Applied Statistics* **29**:539-568.
- 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.
- Nichols, J. D., F. A. Johnson, and B. K. Williams. 1995. Managing North-American Waterfowl in the Face of Uncertainty. *Annual Review of Ecology and Systematics* **26**:177-199.
- Nichols, J. D., R. E. Tomlinson, and G. Waggenerman. 1986. Estimating Nest Detection Probabilities for White-Winged Dove Nest Transects in Tamaulipas, Mexico. *Auk* **103**:825-828.
- Noon, B. R. 2003. Conceptual issues in monitoring ecological resources. Pages 27 to 72 in D. E. Busch and J. C. Trexler, editors. *Monitoring ecosystems: interdisciplinary approaches for evaluating ecoregional initiatives*. Island Press, Washington D.C.
- Olsen, A. R., J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Young. 1999. Statistical issues for monitoring ecological and natural resources in the United States. *Environmental Monitoring and Assessment* **54**:1-45.
- Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical-Inference from Capture Data on Closed Animal Populations. *Wildlife Monographs*:7-135.
- Overton, W. S., and S. V. Stehman. 1996. Desirable design characteristics for long-term monitoring of ecological variables. *Environmental and Ecological Statistics* **3**:349-361.
- Platt, J. R. 1964. Strong inference - certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science* **146**:347-353.
- Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical-Inference for Capture-Recapture Experiments. *Wildlife Monographs*:1-97.
- Pollock, K. H., J. D. Nichols, T. R. Simons, G. L. Farnsworth, L. L. Bailey, and J. R. Sauer. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics* **13**:105-119.
- Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics* **52**:703-709.
- Ringold, P. L., B. Mulder, J. Alegria, R. L. Czaplewski, T. Tolle, and K. Burnett. 2003. Design of an ecological monitoring strategy for the forest plan of the Pacific Northwest. Pages 73-99 in D. E. Busch and J. C. Trexler, editors. *Monitoring Ecosystems*. Island Press, Washinton DC.
- Romesburg, H. C. 1981. Wildlife Science - Gaining Reliable Knowledge. *Journal of Wildlife Management* **45**:293-313.
- Rosenstock, S. S., D. R. Anderson, K. M. Giesen, T. Leukering, and M. F. Carter. 2002. Landbird counting techniques: Current practices and an alternative. *Auk* **119**:46-53.
- Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. *Ecology* **84**:777-790.

- Rubinoff, D. 2001. Evaluating the California gnatcatcher as an umbrella species for conservation of southern California coastal sage scrub. *Conservation Biology* **15**:1374-1383.
- Sauer, J. R. 2003. A critical look at national monitoring programs for birds and other wildlife species. Pages 119-126 in T. J. O'Shea and M. A. Bogan, editors. *Monitoring trends in bat populations of the United States and territories: problems and prospects*. USGS BRD, Information and Technology Report 2003-0003, Washington D.C.
- Sauer, J. R., J. E. Fallon, and R. Johnson. 2003. Use of North American Breeding Bird Survey data to estimate population change for bird conservation regions. *Journal of Wildlife Management* **67**:372-389.
- Sauer, J. R., W. A. Link, and J. A. Royle. 2004. Estimating population trends with a linear model: technical comments. *The Condor* **106**:435-440.
- Scheaffer, R. L., W. Mendenhall III, and R. L. Ott. 1996. *Elementary Survey Sampling*, 5th edition. Duxbury Press, Boston.
- Seber, G. A. F. 1965. A Note on Multiple-Recapture Census. *Biometrika* **52**:249-259.
- Simberloff, D. 1998. Flagships, umbrellas, and keystones: Is single-species management passe in the landscape era? *Biological Conservation* **83**:247-257.
- Skalski, J. R. 1994. Estimating wildlife populations based on incomplete area surveys. *Wildlife Society Bulletin* **22**:192-203.
- Skalski, J. R., and D. S. Robson. 1992. *Techniques for wildlife investigations: design and analysis of capture data*. Academic Press, San Diego; California, United States.
- Smith, G. A. 1995. *Critical Review of the Aerial and Ground Surveys of Breeding Waterfowl in North America*. U.S. Dept. of the Interior, National Biological Service, Washington, D.C.; District of Columbia, United States.
- Soule, M. E., J. A. Estes, J. Berger, and C. M. Del Rio. 2003. Ecological effectiveness: Conservation goals for interactive species. *Conservation Biology* **17**:1238-1250.
- Stuart, A. 1984. *The ideas of sampling*, 3rd edition. Oxford University Press, New York.
- Suter, G. W. 1993. A Critique of Ecosystem Health Concepts and Indexes. *Environmental Toxicology and Chemistry* **12**:1533-1539.
- Thompson, S. K. 2002. *Sampling*, second edition. John Wiley and Sons, New York.
- Thompson, S. K., and G. A. F. Seber. 1996. *Adaptive Sampling*, 1st edition. John Wiley & Sons, New York.
- Thompson, W. L., G. C. White, and C. Gowan. 1998. *Monitoring vertebrate populations*. Academic Press Inc., San Diego.
- Tocqueville, A. d. 1889. *Democracy in America*. Longmans, London.
- Underhill, L., and D. Gibbons. 2002. Mapping and monitoring bird populations: their conservation uses. Pages 34-60 in K. Norris and D. J. Pain, editors. *Conserving Bird Biodiversity - General Principles and their Application*. Cambridge University Press, Cambridge.
- Uruquhart, N. S., W. S. Overton, and D. S. Birkes. 1993. Comparing sampling designs for monitoring ecological status and trends: impact of temporal patterns. Pages 71-85 in V. Barnett and K. F. Turkman, editors. *Statistics for the Environment*. Academic Press, Chichester.

- USGS. 2001a. North American Breeding Bird Survey. USGS Patuxent Wildlife Research Center, <http://www.pwrc.usgs.gov/bbs/about/>. USGS Patuxent Wildlife Research Center <http://www.pwrc.usgs.gov/bbs/about/>.
- USGS. 2001b. North American Amphibian Monitoring Program. USGS Patuxent Wildlife Research Center, <http://www.pwrc.usgs.gov/naamp/>.
- Vos, P., E. Meelis, and W. J. Ter Keurs. 2000. A framework for the design of ecological monitoring programs as a tool for environmental and nature management. *Environmental Monitoring and Assessment* **61**:317-344.
- Walters, C. J., and C. S. Holling. 1990. Large-Scale Management Experiments and Learning by Doing. *Ecology* **71**:2060-2068.
- White, G. C. in press. Correcting counts: techniques to de-index. *Australian Wildlife Research*.
- White, G. C., D. J. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, Los Alamos.
- Wicklum, D., and R. W. Davies. 1995. Ecosystem Health and Integrity. *Canadian Journal of Botany* **73**:997-1000.
- Williams, B. K., F. A. Johnson, and K. Wilkins. 1996. Uncertainty and the adaptive management of waterfowl harvests. *Journal of Wildlife Management* **60**:223-232.
- Williams, B. K., and J. D. Nichols. 2001. Systems identification and the adaptive management of waterfowl in the United States. *Wildlife Biology* **7**:223-236.
- Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of vertebrate populations. Academic Press, San Diego.
- Yen, P. P. W., F. Huettmann, and F. Cooke. 2004. A large-scale model for the at-sea distribution and abundance of marbled murrelets (*Brachyramphus marmoratus*) during the breeding season in coastal British Columbia, Canada. *Ecological Modelling*.
- Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution* **16**:446-453.