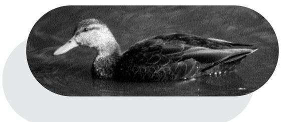
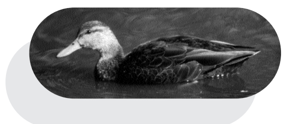
Statistical Analyses
Make the Christmas Bird Count

Relevant for Conservation





John R. Sauer[1], Daniel K. Niven[2], and William A. Link[1]
[1]USGS Patuxent Wildlife Research Center
12100 Beech Forest Drive
Laurel, MD 20708
[2]National Audubon Society
Illinois Natural History Survey
607 East Peabody Drive
Champaign, IL 61820

The North American Bird Conservation Initiative (NABCI) provides a broad vision for bird conservation based on "regionally based, biologically driven, landscape-oriented partnerships" (www.nabci-us.org/nabci.html). At the center of all discussions about implementing this initiative is an active (and sometimes heated) debate about the role of surveys such as the Christmas Bird Count (CBC) in bird conservation. Historically, surveys have been used to describe population change and to define which species are most in need of management. However, NABCI dictates a much more proactive role for surveys.

First, NABCI requires a greater understanding of the relationships between bird populations and features of the landscape that we can manage, such as amounts and dispersion of habitats. Much of our present understanding of these relationships is based on models developed from survey data, and NABCI has been the motivation for new modeling exercises from CBC and Breeding Bird Survey (BBS) data. Issues such as "how best to manage a national forest to enhance the populations of priority bird species," "how to predict the consequences of habitat management for birds on a refuge," and "how to achieve a bird population goal for a Bird Conservation Region" can only be addressed using these models.

Second, NABCI places a priority on use of surveys for direct assessment of the effects of management. The subtle shift away from "barometric" surveys, in which the general status of populations is assessed, to "management-oriented" surveys, in which the response of bird populations to specific management (or other environmental) changes is assessed, has enormous ramifications for surveys. Many surveys are poorly designed because their uses were poorly defined; new surveys are required to provide specific information, and hence require better designs focused on the new uses for the data. Management themes, such as adaptive management, place surveys in a central role in assessing the consequences of management and in improving our understanding of how management influences bird populations.

Third, NABCI reflects the understanding that many issues of geographic scale have become important in bird conservation. Managers need to have information for local areas to assist in local management, but also need to know the relative importance of their areas for regional populations of the species. Again, more emphasis is placed on the use of survey data to estimate population change and abundances in new contexts.

Finally, NABCI has caused introspection about surveys and their application. Uses of long-term databases such as the CBC and the BBS are now getting reevaluated in the context of this new focus. Suggestions for new analyses using data and changes in the design of the survey to enhance the quality of the information are a frequent topic of discussion in the bird conservation community.

For information about the authors, see page 10.

American Black Duck *(Anas rubripes)*. Photo/Glen Smart, U.S. Fish and Wildlife Service

Survey analysts and users of datasets such as the CBC are actively developing statistical methods that allow best use of survey data both for historical goals such as evaluating population change from surveys, and to advance the goals of NABCI. Audubon has been collaborating with the Patuxent Wildlife Research Center of the United States Geological Survey (USGS) in developing a hierarchical modeling approach to the analysis of CBC data that provides a coherent framework for evaluating population change and habitat associations at geographic scales relevant to NABCI.

❇

The CBC and Modern Bird Survey Analysis

The CBC has several features that make it relevant to bird conservation studies.

- CBC circles are located across most of North America. This huge spatial extent allows for "big picture" estimates over the wintering range of many species.
- Many sites of particular management interest (such as refuges, parks, important habitats) are covered by CBC circles.
- The CBC has been conducted since 1900, and long historical databases exist for some areas.

On the other hand, the CBC has some important constraints that need to be considered during analyses.

- It is difficult to model how the counts in circles actually reflect the population of birds in a circle. Factors such as amount of effort in counting are well-known to influence the proportion of birds that are counted, so any reasonable analysis of CBC data must at least evaluate whether effort influences counts.
- CBC circles are quite large (15 miles in diameter), and no information exists about within-circle distribution and abundance of birds. Unfortunately, many refuges and habitats are smaller than the circles.
- The CBC is not a survey in that circles were selected for convenience rather than as a statistical sample. Analyses of the data simply reflect that collection of sites, and any extrapolation to the entire region, requires some (probably invalid) assumptions.

While these limitations might seem onerous, it is useful to recognize that most bird surveys have constraints similar to those of the CBC. Only surveys specifically designed for waterfowl (the Spring Breeding Ground Survey, Smith 1995) can claim to have a design that incorporates appropriate detectability (i.e., estimation of how many birds are missed during counting) and spatial sampling (i.e., random site selection) as part of the design. For the CBC and other surveys, we attempt to accommodate through modeling the limitations during the analysis, and then try to be up front in stating the assumptions of the analysis.

A Better Mousetrap

Historically, CBC data have been analyzed using linear models, in which counts are predicted by a linear combination of effects such as circle, year, region, and amount of effort. Hierarchical models are an extension of this approach, in which the effects themselves are viewed as random variables (Gilks et al. 1996). Although this modeling (and most statistical formalism) may seem a bit arcane to people who do not work with it, the approach allows for an extremely flexible set of models. Many issues arising in estimation from CBC data are naturally handled using hierarchical models. We need to summarize information at multiple geographic scales and summarize results at the scale of circles, strata (states or physiographic regions), and over species ranges. We tend to view features such as effort effects or population change as varying through space, but reflecting a common underlying process. Hence among regions the features might reflect a common statistical distribution, but each region would be a random expression of that distribution. Also, hierarchical models that treat circle, year, and effort effects as random variables help us to accommodate one of the limitations that have bedeviled analysis of CBC (and other count) data: the variation in quality of information over circles and time. Information that is very difficult to estimate within many smaller regions can be estimated much more efficiently in a hierarchical model that treats it as a random effect. In the next sections, we define the essential hierarchical model, then describe some elaborations.

The Basic Model

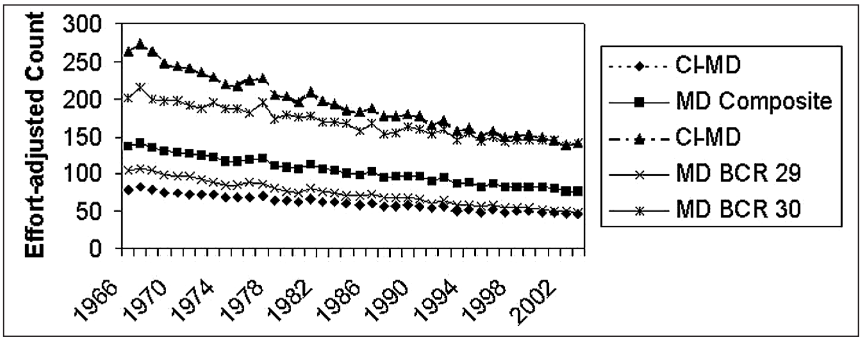
In this section, we descend into some statistical formalism to provide some details to readers who are interested in fitting these models to data. More information and programs for implementing these models for CBC data can be obtained from the authors.

The hierarchical model presently under development for the CBC is similar to the hierarchical model described for the BBS by Link and Sauer (2002), but is modified to include an effort adjustment (cf. Link and Sauer 1999). In the model, the count for a species at a circle in a year is assumed to follow an overdispersed Poisson distribution, the mean of which is governed by factors reflecting population change within strata (a slope coefficient and by year effects), effects of effort in counting, and regional variation in abundance for both the individual circle and for the stratum in which the circle occurs. These features are modeled as additive on the log scale. This model of population change is an extension of earlier CBC analyses (e.g., the loglinear model of Sauer and Link 2002a), with a few important innovations.

- First, it is a model for multiple strata, and permits estimation of change, effort effects, and relative abundance at several geographic scales.
- Second, the factors are hierarchically modeled, meaning that several levels of variation are explicitly modeled. Parameters (year effects, effort, and circle effects, and those describing overdispersion) are modeled as random variables, with distributions governed by higher level parameters (called hyperparameters). Choice of these distributions, and the

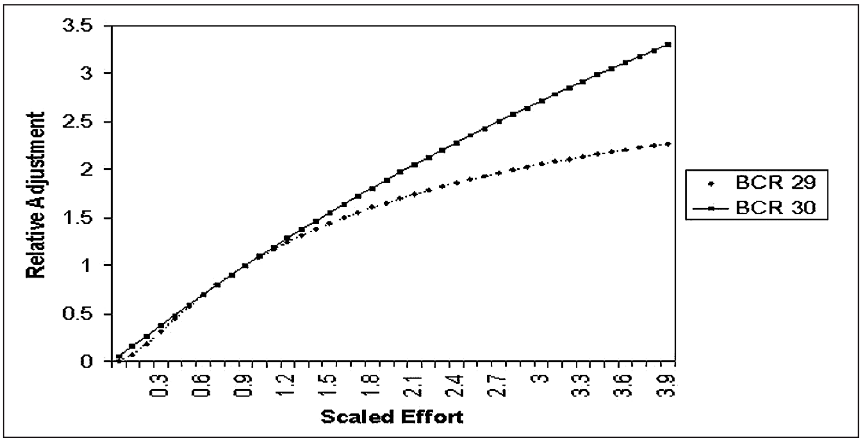
American Black Ducks in Maryland

Figure 1. Population change estimated using hierarchical models for American Black Ducks in Maryland Bird Conservation Regions 29 (Piedmont) and 30 (Coastal New England/Mid-Atlantic Coast). Composite results are presented for Maryland, along with 95 percent credible intervals (Bayesian confidence intervals).



Effort Adjustments

Figure 2. Proportional change in counts associated with differing changes in effort for Bird Conservation Regions 29 and 30 in Maryland. The curves show the relative adjustment of counts associated with varying levels of effort; effort data are scaled so that mean effort is given a value of 1.



extent to which they are allowed to vary among strata, form a critical part of the analysis. For a description of the distributional assumptions in hierarchical models we refer interested readers to Link and Sauer (2002). Not all effects are random; slopes and stratum effects are fixed, as are many of the hyperparameters.

Historically, fitting hierarchical models has been a daunting task. The recent development of Markov-chain Monte Carlo methods has facilitated fitting hierarchical models in a Bayesian framework. In these methods, a simulation approach is used to estimate the distributions of variables of interest (for example, trend), and means and variances of these variables are estimated from replicate outcomes of the simulations. Statistical software packages such as WinBUGS (Spiegelhalter et al. 1999) make these methods readily available for use by statisticians and ecologists.

Our basic model provides sufficient information to estimate composite yearly indices for regions. Indices are yearly measures of abundance, scaled to region and year using the slope, year effect, and stratum effect estimates from the hierarchical model (cf. Link and Sauer 2002). Population trend is estimated from these yearly indices to summarize change for a specific interval.

Examples: Fitting the Basic Model to CBC Data

Audubon and USGS staff are presently fitting this basic model to all species from CBC data, and analyses are presented in this volume for a group of species that breed in boreal North America (and hence are poorly surveyed by the BBS) but winter primarily in the CBC range (Niven et al., this volume). Annual indices of abundance and trend estimates for each species are estimated for Bird Conservation Regions (BCRs), which are physiographic strata designed for regional analyses in NABCI (see www.nabci-us.org/bcrs.html for details of BCRs), and for the entire species wintering ranges within the CBC survey area. After appropriate peer review, this analysis will be placed on a web site similar to that of the North American Breeding Bird Survey Analysis and Summary (www.mbr-pwrc.usgs.gov/bbs/bbs.html).

To demonstrate results from the method, we estimated population change for American Black Ducks for Mid-Atlantic Coast and Piedmont BCRs in Maryland over the interval 1966–2003 (Figure 1). Stratum-specific effort adjustments are shown in Figure 2. Overall, Black Ducks were declining in Maryland during the interval (yearly percent change as calculated from the first and last years of the time series: –1.51 percent/year, credible interval: –2.55, –0.54). We note that, when applied at the continental scale, care has to be used when considering what regions to include in analyses. Missing data can greatly decrease the precision of estimates.

Incorporating Explanatory Variables in the Analysis

As noted above, tying population change to environmental factors is a crucial step in making survey data relevant to management. The hierarchical modeling provides a means for doing this, and it can easily be extended to incorporate habitat data or any other relevant explanatory information that can be collected for individual circles or at regional scales. These covariates are included as additional factors in the model that can be additive on the log scale (as are the circle and strata effects, above) or can be parameterized to fit some hypothesized relationship to population size. One primary source of habitat data is the USGS National Land Cover Data (NLCD, http://landcover.usgs.gov/natllandcover.asp), which uses Landsat satellite data to provide a summary classification of land use for a 30-meter grid across the continental United States. Each "pixel" (individual grid cell) is grouped into one of 21 categories (e.g., Open Water, Row Crops, Deciduous Forest; see http://landcover.usgs.gov/classes.asp). Using a Geographic Information System, it is easy to find the pixels associated with each CBC circle, and to summarize the habitats within the circle. These data are also used to evaluate interdispersion of habitats and amount of edge, landscape characteristics thought to influence bird distribution.

For our American Black Duck example, we provide a summary of what the NLCD looks like for a single CBC circle, the Gibson Island Circle (Figure 3). We summarized the amounts of emergent wetlands for each CBC circle in Maryland, and included it as a covariate in the hierarchical model described above. As expected, there is a significant positive relationship between amount of wetland and Black Duck abundance (slope coefficient=0.55, credible interval=0.32,0.91). Transformed to the original units, this suggests that the effort-adjusted CBC counts increase by about 1.72 for each unit area (30 m^2) increase of wetland in the circle.

These models of associations between counts (and presumably population size) and wetlands have obvious relevance for managers, who can use this factor as an estimate of the relative change in abundance at a circle as a consequence of adding additional wetlands. Because the effect of wetlands is estimated in the hierarchical model, other factors influencing counts (such as population change and effort in counting) are controlled for. The continuing role of CBC is in defining these relationships between habitat, other environmental variables, and bird counts, and as an ongoing data source that can be used to test these models (all models need to be tested constantly). More importantly, the CBC can be used in conjunction with management following the adaptive management paradigm to improve our understanding of how birds relate to the habitats. By managing habitats, then using CBC data to evaluate how bird populations actually change, we are updating our understanding of the effects of management, and this knowledge can be used for more effective future management.

Perhaps in the future, the notion of using CBC data for evaluation of change for specific areas within a circle will lead to innovations both in how we sample within circles and in our methods for counting within circles. Data are often now collected for sectors (or observer "territories") within circles, but no formal means exists for summary and use of this local information. However, if within-circle data proved useful to managers, more formal arrangements could be made for storage and analysis of sector data. Similarly, counting approaches could be modified to incorporate population estimation if needed, and these estimates would better index the actual population than do the current counts.

Summaries Among Groups of Species

Managers often focus on management of habitats to collectively influence population change for groups of species such as grassland-wintering birds. A summary of population change for multiple-species groups can both assist in assessment of the management activities and be of more general ecological interest. Hierarchical models form a natural means to these summaries, as the question involves considering the underlying statistical parameters governing the distribution of a group of estimates of population change rather than considering the esti-

Table 1. Summary analyses of population change for 14 raptor species, 1955–1999. Trend estimates and their standard errors were estimated using loglinear models. A hierarchical model was then used to estimate the mean species rank and the probability that the estimated trend is within a "stable" interval (here defined as ±2 percent).

Species	Scientific Name	Trend	Standard Error (Trend)	# CBCs	Rank	Probability (Stable)
Swainson's Hawk	*Buteo swainsoni*	-7.09	0.954	235	1.0	0.00
Red-tailed Hawk	*Buteo jamaicensis*	0.62	0.291	2681	3.26	1.00
American Kestrel	*Falco sparverius*	0.66	0.091	1909	3.27	1.00
Red-shouldered Hawk	*Buteo lineatus*	0.93	0.257	1588	4.5	1.00
Northern Goshawk	*Accipiter gentilis*	0.79	1.1	1409	4.6	0.85
Rough-legged Hawk	*Buteo lagopus*	1.57	1.05	2054	6.3	0.66
Prairie Falcon	*Falco mexicanus*	1.67	0.44	488	6.7	0.77
Cooper's Hawk	*Accipiter cooperii*	1.82	0.213	1624	7.2	0.80
Golden Eagle	*Aquila chrysaetos*	2.88	0.735	1041	9.4	0.13
Peregrine Falcon	*Falco peregrinus*	3.10	0.61	666	9.9	0.04
Sharp-shinned Hawk	*Accipiter striatus*	3.73	0.387	2326	11.3	0.00
Bald Eagle	*Haliaeetus leucocephalus*	4.25	1.07	2073	11.7	0.03
Ferruginous Hawk	*Buteo regalis*	4.14	0.411	957	12.3	0.00
Merlin	*Falco columbarius*	4.66	0.226	1067	13.6	0.00

mate for a single species. Link and Sauer (1995) and Sauer and Link (2002b) formulated this summary in terms of a hierarchical model in which trends for a series of species are thought to be governed by a common mean trend and variance, with additional distributional assumptions for the variances. The analyses assist in ranking species by need for management action by sorting out species that have poor information from species that are truly extreme in population trend. Using the model, we can also define a convenient summary statistic (the proportion of species in the group with declining populations), and estimate probability that any species in the group is stable by calculating the probability that the trend is outside an interval defined by (-∂,∂) (∂ is a threshold value such as 2 percent/year).

In Table 1, we present an example of these analyses for a group of 14 raptor species (data are from a presentation by Sauer and Link [2000]). These analyses focus attention on characteristics of the collection of species, and provide a view of how to best judge which species show extreme population changes relative to other species in the group. For the raptor species, the hierarchical model analysis highlights the notion that several species are increasing in population, but one species is declining.

Maintaining a Dialogue Among Managers, Statisticians, and Biologists for Best Data Use

The CBC is being used for conservation, and part and parcel of this use is both controversy about the value of the data and an ongoing discussion of how best to use and analyze the data. The Audubon–USGS collaboration seeks to promote the appropriate use of the data by developing and applying statistical methods that accommodate some of the limitations of the data and make the analyses available to the conservation community. One reason that datasets are not used is due to technical limitations: Managers do not have access to analyses that they can use for conservation, statistical tools are misapplied or not made available, and biologists are not sensitive to the management and statistical needs for data collection. All of these partners must interact for successful use of the CBC. Hopefully, increased integration of the CBC with conservation planning will result in careful evaluations of how best to modify the data collection to provide better local information on bird populations, while still maintaining the public participation and educational aspects of the counts.

The hierarchical modeling tools we are developing are extremely flexible, and provide a neat framework for many analyses of CBC data. A variety of other summary analyses are possible in the hierarchical modeling framework. Aggregation of annual indexes among species to provide composite population trajectories have been suggested as an index to the health of bird populations, and we are evaluating approaches for estimation of these composite trajectories. By making these tools available, along with software, data, and expertise to apply the models, we hope to advance the use of this important database for conservation.

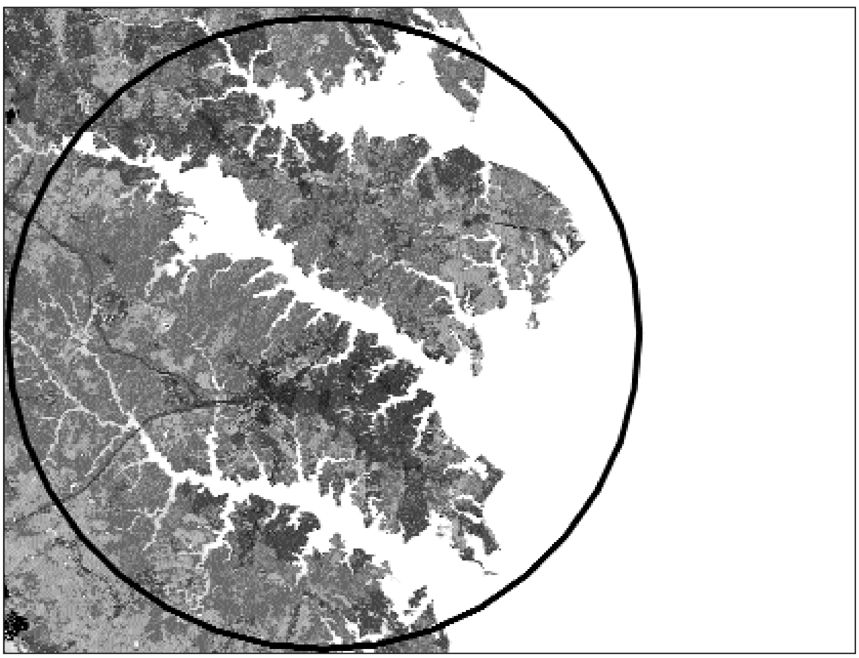


Figure 3. The shading of the land area on this USGS National Land Cover Data map represents different land cover types within the Gibson Island, MD, Christmas Bird Count, and reflects the interspersion of habitats such as emergent wetlands, residential and commercial developments, and hay and row crops.

Literature Cited

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996. Markov-chain Monte Carlo in Practice. Chapman and Hall, New York.

Link, W. A., and J. R. Sauer. 1995. Estimation of empirical mixing distributions in summary analyses. *Biometrics* 51:810–821.

Link, W. A., and J. R. Sauer. 1999. Controlling for varying effort in count surveys: An analysis of Christmas Bird Count Data. *Journal of Agricultural, Biological and Environmental Statistics* 4:116–125.

Link, W. A., and J. R. Sauer. 2002. A hierarchical model for population change with application to Cerulean Warblers. *Ecology* 83:2832-2840.

Sauer, J. R., and W. A. Link. 2000. Population change in raptors, 1955–1999, from the Christmas Bird Count. North American Raptor Monitoring Strategy Workshop. Annual Meeting, Raptor Research Foundation. Jonesboro, AR, September 2000.

Sauer, J. R., and W. A. Link. 2002a. Using Christmas Bird Count data in analysis of population change. *American Birds 56 (The 102nd Christmas Bird Count):* 10-14.

Sauer, J. R., and W. A. Link. 2002b. Hierarchical modeling of population stability and species group attributes using Markov-chain Monte Carlo methods. *Ecology* 83:1743-1751.

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

Spiegelhalter, D. J., A. Thomas, and N. G. Best. 1999. WinBUGS version 1.2 user manual, MRC Biostatistics Unit, University of Cambridge, Cambridge, UK.