

HIERARCHICAL MODELING OF POPULATION STABILITY AND SPECIES GROUP ATTRIBUTES FROM SURVEY DATA

JOHN R. SAUER¹ AND WILLIAM A. LINK

USGS Patuxent Wildlife Research Center, 11510 American Holly Drive, Laurel, Maryland 20708 USA

Abstract. Many ecological studies require analysis of collections of estimates. For example, population change is routinely estimated for many species from surveys such as the North American Breeding Bird Survey (BBS), and the species are grouped and used in comparative analyses. We developed a hierarchical model for estimation of group attributes from a collection of estimates of population trend. The model uses information from predefined groups of species to provide a context and to supplement data for individual species; summaries of group attributes are improved by statistical methods that simultaneously analyze collections of trend estimates. The model is Bayesian; trends are treated as random variables rather than fixed parameters. We use Markov Chain Monte Carlo (MCMC) methods to fit the model. Standard assessments of population stability cannot distinguish magnitude of trend and statistical significance of trend estimates, but the hierarchical model allows us to legitimately describe the probability that a trend is within given bounds. Thus we define population stability in terms of the probability that the magnitude of population change for a species is less than or equal to a predefined threshold. We applied the model to estimates of trend for 399 species from the BBS to estimate the proportion of species with increasing populations and to identify species with unstable populations. Analyses are presented for the collection of all species and for 12 species groups commonly used in BBS summaries. Overall, we estimated that 49% of species in the BBS have positive trends and 33 species have unstable populations. However, the proportion of species with increasing trends differs among habitat groups, with grassland birds having only 19% of species with positive trend estimates and wetland birds having 68% of species with positive trend estimates.

Key words: *birds; Breeding Bird Survey; conservation action; hierarchical models; Markov chain Monte Carlo methods; population stability; population trends; ranking estimates; species group attributes; surveys.*

INTRODUCTION

Biologists, managers, and the public often express questions about population attributes of birds and other taxa in terms of groups. Comparative analyses of population attributes such as change and abundance (e.g., Flather and Sauer 1996, James et al. 1996) often use species as replicates. Management directed at groups of species is increasingly being adopted as a strategy for avian conservation. Species using common regions or habitats for breeding (e.g., grassland) or wintering (e.g., short-distance migrants in the southeastern United States) are often identified as requiring particular conservation action (e.g., Peterjohn and Sauer 1999). Patterns of population trend in groups such as Neotropical migrant birds (e.g., Robbins et al. 1989), waterfowl, and other taxa have motivated conservation plans for the groups. Furthermore, monitoring information from individual species and groups of species is often used to measure population responses for research and management. Requisite summaries of group attributes include average population trend for the

groups, individual species estimates and ranks of species by population trend within groups, and determination of species with stable populations or changing populations.

Behind these questions and initiatives is the notion that the collection of species is meaningful, that some common characteristic of the species permits summary of population attributes among species, imposes a common need for conservation, and allows us to investigate patterns of variation among the species in the group. Wise implementation of management or comparative analyses based on species groups requires an acknowledgment of the dangers of artificial aggregation of species based on single attributes (Mannan et al. 1984, Harvey and Pagel 1991).

In this paper, we use bird survey data both to define the issues associated with analysis of collections of species and to provide examples of applications. Even when information on population attributes such as trend (interval-specific rate of population change; Link and Sauer 1998) are available for a collection of species from surveys such as the North American Breeding Bird Survey (BBS, Peterjohn and Sauer 1993), the quality of information may vary greatly among species. Appropriate handling of estimates of varying precision

Manuscript received 29 June 2000; revised 30 March 2001; accepted 6 April 2001; final version received 20 July 2001.

¹ E-mail: John_R_Sauer@usgs.gov

can be complicated; simple averages and summaries of estimates can be misleading. Link and Sauer (1995, 1998) suggested using empirical Bayes procedures to estimate proportions of species with positive trends and to rank species within groups by magnitude of trend; the empirical Bayes approaches take into account differences in precision among estimated trends.

Assessments of population status for groups of birds often include attempts at classifying populations as stable or declining. Here, too, there are difficulties associated with poor or imprecise information. Simple criteria based on the estimated magnitude of population trend seem reasonable, but an imprecise trend estimate may be quite large while still having a confidence interval large enough to include zero trend; the trend is not significantly different from zero. Of course, "statistical significance" (rejection of the null hypothesis of zero trend) alone is no basis for classifying populations as stable or declining: a very small rate of change may be identified as "statistically significant," but may be of no practical significance. The difficulty of separating notions of statistical significance from magnitude of trend has led to a variety of ad hoc characterizations of population stability.

Approaches to summary of group attributes

Summarizing population attributes for species groups requires a more complex statistical model than is needed for analysis of a single species. For analysis of a single species s , we generally assume that data Y_s are governed by a fixed, unknown parameter θ_s , which is estimated solely on the basis of the data Y_s , and without regard to other species. For analysis of species groups, the parameters θ_s are treated as random variables sampled from group-specific probability distributions, themselves governed by higher level parameters. Indeed, it is sometimes useful to view the higher level parameters also as random variables, sampled from probability distributions governed by yet higher level parameters. We may envision a hierarchy of probabilistic relations, in which parameters, like data, are realizations of random variables, the only distinction between data and parameters being that the data are observed, whereas the parameters can only be estimated. This conceptualization provides a framework for describing groups of species and for evaluating individual species in the context of the group. Multilevel models, in which distributions of data and parameters are described conditionally on realized values of parameters that are also random variables, are called hierarchical models in the statistical literature (Gelman et al. 1995).

Analysis of hierarchical models, most naturally handled by Bayesian methods, is carried out with two products in mind: probability statements about individual parameters θ_s , and probability statements about the group of parameters, i.e., statements about the distribution from which the values θ_s were drawn. Bayesian

analysis is model based, and thus allows for improved estimation of individual parameters by considering them in the context of the group, a characteristic commonly described as "borrowing strength from the ensemble" (Louis 1984, Rubin 1984).

A Bayesian analyst distinguishes several different probability distributions describing data and parameters. First, there is the sampling distribution of the data given the unknown parameter $f(Y_s | \theta_s)$; this is the distribution used in likelihood-based analyses, and that would typically be used in a single-species analysis. Next, there is the *prior distribution* of the parameters $\pi(\theta_s | \psi)$; this distribution is governed by a *prior parameter* (or *hyperparameter*) ψ . The prior and sampling distributions are combined using Bayes theorem to produce the *posterior distribution*, denoted $f(\theta_s | Y_s)$. This distribution combines the information available from the sampling distribution with knowledge based on modeled relations among values of θ_s . An attractive feature of Bayesian analysis is that the hierarchy of modeled relations among parameters can be extended; we may specify *hyperprior distributions* on ψ to express uncertainty regarding the prior parameters.

All Bayesian inference about θ_s is based on the posterior distribution. In particular, the shortest interval containing 95% of the posterior probability is the Bayesian confidence interval (sometimes referred to as a "credible interval"). Another descriptor of the posterior distribution is the "posterior mean," or "Bayes estimator" of θ_s . The Bayes estimator is typically found in an interval ranging from the prior mean to the estimate based on Y_s alone (the *raw estimate*). In fact, it is usually appropriate to regard the Bayes estimator as a precision-weighted average of the two end points, weighted by the variances of the sampling distribution and prior. Thus the weaker the information provided by Y_s alone, the closer the Bayes estimator will be to the prior mean; the Bayes estimator is said to "shrink" the raw estimate toward the prior mean.

Description of prior and hyperprior distributions can be a controversial aspect of Bayesian analysis, but as in any model-based application, the solution to such controversy is for the analyst to present clear statements of structural assumptions, allowing honest assessment of the sensitivity of conclusions drawn to the posited model. Various procedures exist for definition of prior and estimation of the posterior distribution. In our earlier analysis, we used empirical Bayes procedures to estimate the posterior distribution of the number of species with positive trend estimates. These procedures have been applied to analysis of BBS data (e.g., Sauer et al. 1997), but require use of empirical Bayes bootstrapping procedures and, hence, are not in common use among ecologists.

Here, we develop an alternative approach, estimating the proportion of species with increasing populations and the rankings of species using a hierarchical stochastic model. The hierarchical model formulation ex-

tends the Link and Sauer (1995) approach by accommodating uncertainty in estimates of variances of trend for individual species. It also provides a precise definition of “stability” of population trend for a species in the context of the group. This model provides a coherent structure for analysis of group attributes by specifying the prior distribution of parameters that are associated with the species groups. We then use Markov Chain Monte Carlo procedures (MCMC, Gilks et al. 1996) to calculate features of the posterior distributions. MCMC is a remarkable tool for estimation of parameters in these hierarchical models, allowing estimation of attributes of the posterior distributions (such as parameter estimates) based on specification of the underlying distributions of the parameters. We present the principles underlying the procedure, then apply it to a recent (1966–1999) analysis of BBS data for several groups of management interest.

METHODS

Hierarchical model for trend estimates

We specify a hierarchical model that describes the parameters of interest and how they vary among species. Some mathematical formalism is needed to adequately describe the model. First, we specify the sampling distributions of the data, conditional on the parameters. We assume that trend estimates $\hat{\beta}_s$ are available for species $s = 1, 2, \dots, n$; we write $[\hat{\beta}_s | \beta_s, \sigma_s^2] = N(\beta_s, \sigma_s^2)$ to denote that the sampling distribution for the estimator $\hat{\beta}_s$ is normal with mean given by parameter β_s and variance σ_s^2 . We also assume that variance estimates $\hat{\sigma}_s^2$ are available, and that

$$\left[\frac{v_s \hat{\sigma}_s^2}{\sigma_s^2} \middle| \sigma_s^2 \right] = \chi_{v_s}^2$$

(i.e., the variance estimate is distributed as a multiple of a chi-square with v_s degrees of freedom). By specifying distributions of the estimated variance, we account for uncertainty in the estimation of individual variances, which was not done in the Link and Sauer (1995) empirical Bayes approach.

Next, we define the models for prior distribution of parameters for species-specific trends as $[\beta_s | \mu, \tau^2] = N(\mu, \tau^2)$; thus we treat the trend parameters as being sampled from a distribution with mean μ and variance τ^2 . We used objective Bayes methods which specify extremely limited knowledge of the prior parameters. Thus, to complete the Bayesian specification, we specified diffuse normal priors for the mean $[\mu] = N(0, 1000^2)$, and flat inverse gamma distributions for variances τ^2 and σ_s^2 ; these are standard non-informative priors for Bayesian analysis (Spiegelhalter et al. 1995).

Based on these sampling distributions and prior distributions, we can calculate a set of distributions known as *full conditionals* for all of the parameters. The full conditional for β_s , for instance, is the distribution of β_s having fixed all other parameters at specified values; for

simplicity, we denote this full conditional by $[\beta_s | \cdot]$. As we will discuss, the full conditionals are a critical component of the procedure used to calculate the posterior distributions $[\beta_s | Y]$, $[\mu | Y]$, and $[\tau^2 | Y]$.

The Markov Chain Monte Carlo procedure

Hierarchical models are often extremely difficult to apply, as it is difficult to calculate the posterior distributions. MCMC is a convenient way to overcome these difficulties and permits estimation of parameters and hyperparameters for very complicated problems. The basic idea is to produce Markov chains of values, with the chains being defined in such a way that their stationary distributions are the posterior distributions of interest. This idea is very similar to familiar Monte Carlo methods, the only difference being that the sampled values are not independent, but have serial correlation like that of a Markov chain.

We used Gibbs sampling (Gilks et al. 1996) to produce the desired Markov chains. Gibbs sampling is cyclical sampling from full conditionals: a sample is drawn for β_s from the full conditional $[\beta_s | \cdot]$, for each $s = 1, 2, \dots, n$; then a new value for each variance is drawn from the appropriate full conditional, then a new value for μ from $[\mu | \cdot]$. At each iteration of the Gibbs samples, the full conditional distributions are calculated using the current values of the other parameters. This iterative sampling from the full conditional distributions produces a Markov chain of values which can be used to approximate the posterior distributions to any desired degree of accuracy. We present full conditional distributions for parameters of our model (Appendix A). See Gilks et al. (1996:10–11) for additional references and information on full conditional distributions and Gibbs sampling.

A computer program (BUGS; Spiegelhalter et al. 1995) is freely available that allows relatively simple model formulation and estimation using MCMC.² Program BUGS begins with model specification as previously described, and then produces Markov chains either by Gibbs sampling or by one of a variety of similar sampling algorithms. Alternatively, computer code for Gibbs sampling is easily written; in this analysis, we used the statistical programming package GAUSS (1994; Aptech Systems, Maple Valley, Washington, USA) to generate our Markov chains.

In conducting an MCMC analysis, an initial sequence of observations of each Markov chain is discarded (e.g., the first 5000 observations). This portion of the Markov chain (the “burn-in”) may not be representative of the stationary distribution (i.e., the posterior distribution from which inference will be made). The Markov chains are then observed for a sufficiently large number of cycles (e.g., 25 000 iterations). Several recent works provide useful introductions to MCMC (e.g., Gilks et al. 1996, Kass et al. 1998.)

² URL: <http://www.iph.cam.ac.uk/bugs/>

In our analysis, we use MCMC to calculate posterior distributions of β_s , μ , τ^2 , and σ_s^2 ; for point estimates, we used the means of the posterior distributions. The number of species from the group with positive population trend estimates (denoted as N_{inc}) and corresponding proportions p_{inc} are calculated as the mean number of estimates with positive estimates from the MCMC replicates. We derive 95% confidence intervals based on the observed percentiles from the MCMC replicates.

Definition of a stable population

Often, the need for management is phrased in terms of a magnitude of decline over a specific interval. For example, a 1% per year trend over a 33-yr period may be considered a threshold for management action. The notion of setting a threshold value of trend, beyond which the population is said to be unstable, is reasonable; unfortunately, estimates of population trend are often controversial as information for use in management. An estimate of population trend for a species may be extreme because it is imprecisely estimated, and many "extreme" estimates are imprecise (Link and Sauer 1997). Consequently, many managers who use population trend information do not rely solely on estimates of population trend, but instead assess significance of population trend, requiring the trend to be significantly different from 0 (based on results of a hypothesis test) before evaluating the magnitude of the trend. Unfortunately, this approach does not eliminate the consequences of imprecise estimates; "significant" estimates are not necessarily unstable with regard to a management threshold. When making relative comparisons of trends, additional information from the other species in the group provides a context of relative precision that allows improved estimates of magnitudes and rankings of trends (Link and Sauer 1996, 1997).

The hierarchical model allows us to accommodate the differences in precision among estimated trends, and to define stability in terms of magnitude of population trend. Hence, if δ is the maximum acceptable deviation from 0, the probability that a population is stable can be written as

$$\Pr(\beta_s \in (-\delta, \delta) | Y).$$

This value can be directly calculated from the posterior distribution; i.e., by calculating the frequency with which the Markov chain for β_s falls in the range $(-\delta, \delta)$. Given a value δ , and a value ρ (say, 0.90), we might choose to identify a species as having an unstable population if the probability β_s is not in $(-\delta, \delta)$ exceeds ρ ; this is equivalent to $\Pr(\beta_s \in (-\delta, \delta) | Y) \leq 1 - \rho$.

Application to North American Breeding Bird Survey

The BBS is a roadside survey of North American birds, conducted (primarily) in June along 50-stop roadside survey routes. At each stop, a 3-min point count is conducted, and the sum of the counts of birds

is used as an index to abundance for each species. The survey began in 1966, spanned the continental United States by 1968, and now covers the continental United States, Alaska, and southern Canada. See Robbins et al. (1986) for more details about the BBS. Interval-specific population trends were estimated for 399 species from the BBS using route regression analysis (Link and Sauer 1994). Details of the BBS estimation procedures are presented elsewhere (e.g., Link and Sauer 1994), and are based on estimating population trend route-by-route using a Poisson regression with log links (the estimating equations of Link and Sauer 1994), and then averaging the route trend to get estimates of regional trend. Variances are estimated by bootstrapping among routes, and differences in regional coverage and survey quality are accommodated by weighting route data in the regional averaging (Link and Sauer 1994). Although some controversy exists about appropriate methods for BBS analysis, the route regression method based on estimating equations that is applied here appears to provide results consistent with those from other procedures (Peterjohn et al. 1997). We estimated trend at the scale of the entire survey area, for the interval 1966–1999, noting that the survey did not begin until 1968 in some parts of the survey area.

Ecologists have been particularly interested in trends for groups of species, for example, Neotropical migrant birds, or birds that breed in grassland habitats. To meet this need, BBS results are often summarized for 12 groups of species, in four categories: breeding habitat (grassland, wetland/open water, successional/scrub, woodland, and urban habitats), nest type (cavity, open-cup), migration (permanent resident, short-distance, and Neotropical), and nest location (ground/low, mid-story/canopy). See Peterjohn and Sauer (1993) and Sauer et al. (1997) for summary lists of the species in these groups. A species can occur in only one group within each category, and we note that nest type and location categories are only defined for a subset of species (passerine birds and cuckoos).

Analysis

We used MCMC to fit the hierarchical model to BBS data for the 12 species groups and to the group of all species surveyed by the BBS. To ensure that the MCMC procedure had converged to a stationary distribution, we allowed a burn-in period of 5000 iterations of the Markov chain. We based our analysis of posterior distributions on the next 25 000 observations of each chain. Posterior distributions for parameters of interest are summarized by their means, standard deviations, and central 95% intervals. We denote posterior means with a superscript "B."

We defined two stability thresholds, $\delta = 2\%$ and 1% per year, and assessed the number of species in each group for which $\Pr(\beta_s \in (-2, 2) | Y) \leq 0.05$ or $\Pr(\beta_s \in (-1, 1) | Y) \leq 0.05$.

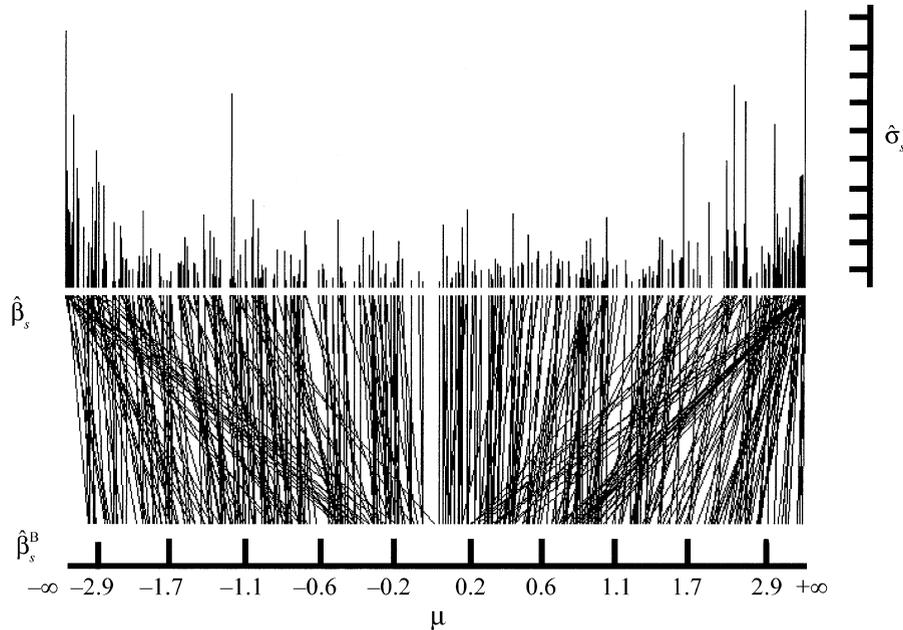


FIG. 1. Shrinkage plot for population trend estimates of 399 species of birds, estimated from North American Breeding bird survey data for the interval 1966–1999. Line segments connect raw estimates $\hat{\beta}_s$ and Bayes estimates $\hat{\beta}_s^B$ on two parallel scales, highlighting shrinkage toward the overall mean ($\mu = 0.06$); the more a segment departs from vertical, the greater the shrinkage. Estimated standard errors $\hat{\sigma}_s$ are plotted as vertical lines above corresponding estimates $\hat{\beta}_s$; the size of the lines represents the magnitude of the standard error. For clarity of presentation, the trend scale is uniform on the inverse logit scale.

RESULTS

Population trend was estimated for 399 species (Appendix B). Species varied greatly in quality of information, and some had extreme estimates of population trend. Generally, the Bayes estimates $\hat{\beta}_s^B$ based on all species groups tended to shrink the extreme estimates toward the estimated prior mean μ (0.06 ± 0.105 , mean ± 1 SE). A plot of ranked $\hat{\beta}_s^B$ and the $\hat{\beta}_s$ clearly shows the effect of the shrinkage toward the prior mean; many of the extreme values of $\hat{\beta}_s$ appear to be very imprecise (Fig. 1).

Estimates calculated among species groups show large intergroup differences in population trend. For the breeding-habitat groups (Table 1), we present the

Bayes estimates of μ , τ , N_{inc} , and p_{inc} . The 27 species of grassland breeding birds collectively appear to be declining, with $\mu^B = -1.13\%$ per year, $N_{inc}^B = 5.23$, so that $p_{inc} = 0.19$ with confidence intervals not including 0.50. Successional and urban breeding species also had negative estimates for μ (-0.54 and -0.63 , respectively), and estimated $p_{inc} < 0.50$. Wetland and woodland breeding birds had positive μ^B 's and p_{inc}^B 's, although the woodland p_{inc}^B had a confidence interval that overlapped 0.50. Note that the number of species with positive estimates can only take integer values; hence, the percentile confidence intervals of p_{inc}^B vary in confidence level from 0.944 to 0.973. Open-cup nesting species, Neotropical migrants, and ground-nesting spe-

TABLE 1. Summaries of posterior distributions for population trend parameters (1966–1999) of breeding bird habitat groups from the North American Breeding Bird Survey.

Breeding-habitat group	No. species [†]	Population trend measurements				CI for p_{inc}^B	Confidence level
		$\mu^B \ddagger$ (1 SD)	$\tau^B \S$ (1 SD)	$N_{incl}^B $ (1 SD)	$p_{inc}^B \P$		
Grassland	27	-1.13 (0.391)	1.67 (0.351)	5.23 (1.407)	0.19	0.11, 0.33	0.973
Wetland	80	1.39 (0.366)	2.67 (0.333)	54.70 (3.144)	0.68	0.61, 0.77	0.962
Successional	86	-0.54 (0.180)	1.37 (0.166)	29.54 (3.340)	0.34	0.27, 0.43	0.964
Woodland	118	0.16 (0.156)	1.45 (0.112)	65.23 (3.740)	0.55	0.49, 0.62	0.956
Urban	15	-0.63 (0.318)	0.93 (0.295)	3.12 (1.583)	0.21	0.07, 0.47	0.944

[†] Number of species in the species group.
[‡] Mean trend.
[§] Variance.
^{||} Number of species with positive trend estimates.
[¶] Proportion of species with positive trend estimates.

TABLE 2. Number of species, number of species with positive trend, and proportion of species with positive trend with confidence intervals for species groups and all species from the North American Breeding Bird Survey, 1966–1999.

Species group type	Total no. species	No. spp. increasing (N_{inc}^B)		Proportion of spp. increasing (p_{inc}^B)		
		Mean	(1 SD)	Mean	CI	Confidence level†
Nesting type groups						
Cavity nesting	50	30.16	(2.197)	0.60	0.52, 0.70	0.963
Open-cup nests	181	71.14	(4.639)	0.39	0.34, 0.45	0.959
Migration type						
Short-distance migrant	96	42.90	(2.898)	0.45	0.39, 0.51	0.963
Permanent resident	85	44.55	(3.404)	0.52	0.45, 0.61	0.961
Neotropical migrant	135	56.72	(4.098)	0.42	0.36, 0.49	0.961
Nest location						
Ground-nesting	110	36.97	(3.340)	0.34	0.27, 0.41	0.976
Midstory/canopy nest	124	66.03	(3.971)	0.53	0.47, 0.60	0.968
All species	399	195.88	(7.011)	0.49	0.46, 0.53	0.962

† Actual confidence level for the percentile confidence interval.

cies all have $p_{inc}^B < 0.50$ (Table 2). For all species, $p_{inc}^B = 0.49$, indicating that, in aggregate, ~50% of species have positive population trends (Table 2). Estimates based on the Link and Sauer (1995) procedure are very similar to the p_{inc}^B estimates, but generally have wider confidence intervals.

Based on all 399 species in the survey-wide analysis, 33 species were unstable with regard to the $\delta = 1\%$ per year criterion, with 14 increasing at a larger rate than would be expected based on the group characteristics and 19 declining at a greater rate than would be expected (Appendix B). Obviously, the 2% per year threshold characterizes fewer species as declining than does the 1% per year threshold. For example, Yellow-billed Cuckoo, with $\beta^B = -1.78\%$ per year, is not considered unstable ($p = 0.92$) with a criterion of 2% per year, because its estimate of trend is within the interval. However, it is unstable based on a criterion of 1% per year ($P < 0.001$).

We analyzed stability within breeding-habitat species groups (Table 3). As expected, groups with fewer estimated increasing species tend to have more species with unstable, declining populations. Grassland species have six of 27 species with unstable populations, and five of them are declining; successional species have eight of 86 species with unstable populations, and seven of them are declining. Woodland birds contain 17 species with unstable populations, with nine species showing declines.

DISCUSSION

Characterizing species as having increasing, declining, or stable populations is difficult. Few species are unequivocally declining or unstable in population, and most of these are well known. For the majority of species, however, we have to take imperfect information

and judge its use. Implicitly, most comparative analyses of population trend are conducted by first eliminating species with obviously inappropriate data, then evaluating the species information in the context of species that share some characteristic, from very general (all bird species) to specific (breeding or migration habitat). In these groups, we want to identify general attributes such as group tendencies for increase or decline, but still preserve the information of individual species in the group context. The hierarchical procedures presented here provide a general means for judging the quality of information for a species in the context of a collection of species.

The hierarchical procedures represent several advances for estimation of population trend. First, they provide a convenient way of estimating the proportion of species with increasing populations for groups of species, and provide a ranking of species by magnitude of population trend. Hierarchical models are a coherent framework for implementing the methods described in Link and Sauer (1995, 1996), but allow more reasonable underlying models (i.e., by modeling uncertainty in the estimates of sampling variances). The models also provide the basis for a reasonable definition of a stable population by defining stability in the context of the best information about the trend parameter, as summarized by the posterior distribution.

How much should group information inform individual species estimates?

Many biologists will have reservations about use of information from a collection of species to improve our understanding of estimates from individual species data. When a single species is the only interest in a study and no relevant group context exists, the individual species estimates are our best information for

TABLE 3. Species with unstable populations, when categorized within breeding habitat group, from survey-wide analysis of Breeding Bird Survey population trends, 1966–1999.

Group	Species increasing	Species declining
Grassland	Sedge Wren, 1	Horned Lark, 1 Sprague's Pipit, 1 Cassin's Sparrow, 1 Grasshopper Sparrow, 2 Eastern Meadowlark, 2
Wetland	Common Loon, 1 Great Blue Heron, 1 Canada Goose, 2 Wood Duck, 2 Gadwall, 2 Barrow's Goldeneye, 1 Osprey, 2 Sandhill Crane, 2 Boat-tailed Grackle, 1	
Successional	Lesser Nighthawk, 1	Northern Bobwhite, 2 Pinyon Jay, 1 Prairie Warbler, 1 Eastern Towhee, 1 Brewer's Sparrow, 1 Field Sparrow, 2 Lark Sparrow, 1
Woodland	Wild Turkey, 2 Barred Owl, 1 Ruby-throated Hummingbird, 1 Pileated Woodpecker, 1 Brown-crested Flycatcher, 1 Blue-headed Vireo, 2 White-breasted Nuthatch, 1 Winter Wren, 1	Yellow-billed Cuckoo, 1 Whip-poor-will, 1 Olive-sided Flycatcher, 2 Eastern Wood-Pewee, 1 Brown-headed Nuthatch, 1 Wood Thrush, 1 Cerulean Warbler, 1 Bachman's Sparrow, 1 Purple Finch, 1
Urban		Common Grackle, 1 House Sparrow, 2

Notes: Species are categorized by increasing or declining populations, and by two stability criteria, with a 1 indicating that the population trend is unstable with regard to the interval $(-1, 1)$ and a 2 indicating that the population is unstable with regard to the interval $(-2, 2)$.

management. However, in an omnibus survey such as the BBS, there is often uncertainty about the validity of information for particular species, as the survey is not optimized for any species and there are untestable assumptions that must be made in analysis of survey data (Link and Sauer 1998). Consequently, it is reasonable to assume that species likely to be influenced by similar environmental factors (such as local habitat change) would benefit from additional information associated with species experiencing common conditions.

Also, when attributes of the collection of species are of interest, there is a clear benefit to employing procedures that use information from the group of species. Efron and Morris (1977) provide a review of shrinkage estimators. They note that these estimators are superior to individual estimates in cases in which any attribute involving the group is of interest. These cases include deriving estimates with minimum mean-squared error, or in ranking estimates, or in correlation studies based on the collection of species. In particular, whenever a group of species is considered collectively, use of the estimates derived in this paper will be superior to individual estimates; hence, any multi-species analysis of stability or comparative population trend will benefit

from application of these methods. Judicious definitions of groups, by area as well as by species, could allow for more effective analysis of a variety of ecologically interesting hypotheses.

It is clear that careful consideration of species groups is prerequisite to any comparative analysis. Groups may be irrelevant or even misleading due to confounding of characteristics unrelated to the attribute shared by group members (Mannan et al. 1984). Harvey and Pagel (1991) discuss the possibility of associations among group members in comparative studies as a consequence of shared ancestry. They suggest some ameliorative steps to accommodate these associations in analysis. If concerns exist about taxonomic (or other) associations within the species groups, the hierarchical model approach described here can be modified to accommodate the associations by explicitly defining the subgroups as additional components in the model. Alternatively, these features can be considered in guild definitions (e.g., by restricting the analysis to groups of common ancestry). In conservation studies such as we describe in this paper, in which the primary interest is response related to a species group, defined at a single geographic scale in terms of an attribute (such

as habitat) that can be managed, the possibility of taxonomic associations may not be particularly important because the common response of all species is of interest. However, more complicated modeling exercises that incorporate spatial scale as part of the hierarchical modeling must pay particular attention to changes in species group constituency over space. As we noted earlier, in any model-based application the possibility of controversy about description of prior and hyperprior distributions always exists. Investigators must present clear statements of structural assumptions to allow honest assessment of the sensitivity of conclusions drawn from the posited model.

BBS analysis results

Any summary of BBS data can be criticized on the basis of methods and choice of temporal and spatial scales of analysis. For group analyses, group constituency is also a reasonable topic for criticism. We acknowledge these concerns, and note that investigators can apply these methods to any analysis that provides means and standard errors of estimated trend for a group of species over any time interval. The program BUGS can be obtained over the internet; source code for this particular problem is available from the authors, and we encourage its use.

Even though most species show temporal and spatial heterogeneity in population trend, the estimates of population trend and group attributes presented here provide reasonable summaries of large-scale patterns of population trend. Furthermore, the patterns are of management interest, as they are associated with habitats or other life history attributes. As in any correlative study, there are certainly other attributes of species within each group that may influence the patterns (e.g., taxonomic associations causing common life history strategies; Harvey and Pagel 1991), and these limitations should be considered when interpreting the correlations. Most grassland bird species are declining, and five of the species have declines that exceed our threshold of stability. Clearly, this group appears to share a common tendency for decline, although the causes of such declines cannot be divined from analysis of BBS data. This lack of association with causes is clearly apparent in other species groups. For example, the urban species also appear to be collectively declining. It is easy to speculate (although difficult to document) that their decline is not a population characteristic, but instead reflects an artifact of the counting method used in the survey. Observers may simply tire of counting these common and uninteresting birds, and over time their counts may decline.

Defining a stable population has proven difficult for managers, who often use a variety of ad hoc methods based on statistical significance for species with sufficient data for analysis. Our notion of stability analysis provides additional credibility to the result, as it is not simply a description of a significant population decline,

but instead explicitly incorporates the magnitude of population trend into the description of population stability. In this formulation, the definition of a stable population is set a priori, and we do not have the conceptual problem of a significant population trend of very small magnitude. Instead, the probability statement is made about whether the observed value falls within a range of interest. It also uses information from the collection to refine the individual species data, hence appropriately using among-species information (Efron and Morris 1977). Different management questions would probably require other levels of stability. We arbitrarily set our δ values at ± 1 and $\pm 2\%$ per year in this analysis, but, in practice, δ should be set in conjunction with management goals and objectives. For example, if a manager had an a priori expectation of expected magnitude of population trend for grassland birds due to change in habitat, the stability zone could be changed to allow detection of species that vary from the expectation.

Many of the species noted to be unstable in our analysis are well known to be declining, based on the individual species results (e.g., Peterjohn and Sauer 1999). However, it is useful to note that our perception of trend in quite a few species is shifted by the Bayes analysis. For example, Chimney Swifts, although well surveyed and precisely estimated, are not shown to have unstable populations in our analysis, even though they are significantly declining. This is, of course, due to our threshold values for a stable population, and reflects the notion that any critical level of stability must be clearly thought out and relevant to some management objective.

Finally, we note that we have followed the recent convention of bird conservation initiatives in defining population stability in terms of a long-term trend. However, the notion of stability as described here can be applied to any of a variety of estimates of population change for a collection of species. A population could be stable in terms of magnitude of year-to-year fluctuations, or in terms of multiyear population fluctuations, and often the time period associated with a "trend" estimated from a survey is not particularly relevant for analysis. Consideration of the appropriate descriptor for evaluating stability should be a component of any conservation activity.

ACKNOWLEDGMENTS

We thank D. H. Johnson, J. D. Nichols, J. A. Royle, and the referees for helpful comments and discussion. We also thank the participants in the North American Breeding Bird Survey project, who have contributed enormous amounts of time and skilled observation.

LITERATURE CITED

- Efron, B., and Morris. 1977. Stein's paradox in statistics. *Scientific American* **236**:119–127.
- Flather, C. H., and J. R. Sauer. 1996. Using landscape ecology to test hypotheses about large-scale abundance patterns in migratory songbirds. *Ecology* **77**:28–35.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. R. Rubin. 1995.

- Bayesian data analysis. Chapman and Hall, New York, New York, USA.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1996. Markov chain Monte Carlo in Practice. Chapman and Hall, New York, New York, USA.
- Harvey, P. H., and M. D. Pagel. 1991. The comparative method in evolutionary biology. Oxford University Press, New York, New York, USA.
- James, F. C., C. E. McCulloch, and D. A. Wiedenfeld. 1996. New approaches to the analysis of population trends in land birds. *Ecology* **77**:13–27.
- Kass, R. E., B. P. Carlin, A. Gelman, and R. M. Neal. 1998. Markov chain Monte Carlo in practice: a roundtable discussion. *American Statistician* **52**:93–100.
- Link, W. A., and J. R. Sauer. 1994. Estimating equations estimates of trends. *Bird Populations* **2**:23–32.
- 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. 1996. Extremes in ecology: avoiding the misleading effects of sampling variation in summary analyses. *Ecology* **77**:1633–1640.
- Link, W. A., and J. R. Sauer. 1997. New approaches to the analysis of population trends in land birds: a comment on statistical methods. *Ecology* **78**:2632–2634.
- 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.
- Louis, T. A. 1984. Estimating a population of parameter values using Bayes and empirical Bayes methods. *Journal of the American Statistical Association* **79**:393–398.
- Mannan, R. W., M. L. Morrison, and E. C. Meslow. 1984. Comment: the use of guilds in forest bird management. *Wildlife Society Bulletin* **12**:426–430.
- Peterjohn, B. G., and J. R. Sauer. 1993. North American Breeding Bird Survey annual summary 1990–1991. *Bird Populations* **1**:52–67.
- Peterjohn, B. G., and J. R. Sauer. 1999. Population status of North American grassland birds from the North American Breeding Bird Survey, 1966–1996. *Studies in Avian Biology* **19**:27–44.
- Peterjohn, B. G., J. R. Sauer, and W. A. Link. 1997. The 1994 and 1995 summary of the North American Breeding Bird Survey. *Bird Populations* **3**:48–66.
- Robbins, C. S., D. Bystrak, and P. H. Geissler. 1986. The Breeding Bird Survey: its first fifteen years, 1965–1979. U.S. Fish and Wildlife Service, Resource Publication **157**.
- Robbins, C. S., J. R. Sauer, R. S. Greenberg, and S. Droege. 1989. Population declines in North American birds that migrate to the neotropics. *Proceedings of the National Academy of Sciences (USA)* **86**:7658–7662.
- Rubin, D. B. 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Annals of Statistics* **12**:1151–1172.
- Sauer, J. R., J. E. Hines, G. Gough, I. Thomas, and B. G. Peterjohn. 1997. The North American Breeding Bird Survey results and analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, Maryland [Online, URL: <http://www.mbr-pwrc.usgs.gov/bbs/bbs.htm>].
- Spiegelhalter, D. J., A. Thomas, N. G. Best, and W. R. Gilks. 1995. BUGS: Bayesian inference using Gibbs sampling, Version 0.50. MRC Biostatistics Unit, Cambridge, UK.

APPENDIX A

Full conditional distributions of hyperparameters in the hierarchical model.

$$[\beta_i | \cdot] = N[w_i \hat{\beta}_i + (1 - w_i)\mu, w_i \sigma_i^2], \quad \text{where} \quad w_i = \frac{\tau^2}{\tau^2 + \sigma_i^2};$$

$$[\mu | \cdot] = N\left(\hat{\beta}, \frac{\tau^2}{n}\right), \quad \text{with} \quad \hat{\beta} = \frac{1}{n} \sum_{i=1}^n \beta_i;$$

$$\left[\frac{1}{\tau^2} | \cdot\right] = \chi_n^2 / \sum_{i=1}^n (\beta_i - \mu)^2, \quad \text{and}$$

$$\left[\frac{1}{\sigma_i^2} | \cdot\right] = \frac{\chi_{v_i}^2}{(v_i \hat{\sigma}_i^2)}.$$

APPENDIX B

A table presenting estimated population trend (1966–1999) and Bayes estimates for species of North American birds, based on the North American Breeding Bird Survey, is available in ESA's Electronic Data Archive: *Ecological Archives* E083-027-A1.