

Report

Cost-effective Sampling Design Applied to Large-scale Monitoring of Boreal Birds

[Matthew Carlson](#) and [Fiona Schmiegelow](#)¹

ABSTRACT. Despite their important roles in biodiversity conservation, large-scale ecological monitoring programs are scarce, in large part due to the difficulty of achieving an effective design under fiscal constraints. Using long-term avian monitoring in the boreal forest of Alberta, Canada as an example, we present a methodology that uses power analysis, statistical modeling, and partial derivatives to identify cost-effective sampling strategies for ecological monitoring programs. Empirical parameter estimates were used in simulations that estimated the power of sampling designs to detect trend in a variety of species' populations and community metrics. The ability to detect trend with increased sample effort depended on the monitoring target's variability and how effort was allocated to sampling parameters. Power estimates were used to develop nonlinear models of the relationship between sample effort and power. A cost model was also developed, and partial derivatives of the power and cost models were evaluated to identify two cost-effective avian sampling strategies. For decreasing sample error, sampling multiple plots at a site is preferable to multiple within-year visits to the site, and many sites should be sampled relatively infrequently rather than sampling few sites frequently, although the importance of frequent sampling increases for variable targets. We end by stressing the need for long-term, spatially extensive data for additional taxa, and by introducing optimal design as an alternative to power analysis for the evaluation of ecological monitoring program designs.

INTRODUCTION

Ten years have passed since ratification of the Convention on Biological Diversity, a landmark, international conservation commitment. However, the conservation of biodiversity continues to be constrained by our inability to reliably predict the effects of environmental perturbations on biota (Walters and Holling 1990). Given this uncertainty, an information feedback loop is required to assess the impact of human activities on biodiversity. Ecological monitoring, the repeated measurement of biotic response to disturbance (Hinds 1984), provides this feedback, thereby facilitating adaptive management (Halbert 1993) and the establishment of conservation and research priorities (Burbidge 1991, Stork and Samways 1995). Large-scale monitoring, where detection of long-term trends over broad spatial scales is the goal, improves our ability to distinguish human impacts from natural changes (Spellerberg 1991) and allows inference at scales relevant to policy (Urquhart et al. 1998).

Although large-scale ecological monitoring is required to conserve biodiversity (Johnson 1993, Minister of

Supply and Services 1995), examples of such programs remain rare, in large part due to the difficulty of achieving effective monitoring data under cost constraints (Committee on Environment and Natural Resources 1997). The magnitude of this challenge is illustrated by the Environmental Monitoring and Assessment Program (EMAP), designed to monitor the status and trends of ecological resources in the United States (U.S. Environmental Protection Agency 2000). Although the program had a large investment of effort and money, a review deemed it unlikely to succeed, due in part to insufficient sampling intensity (National Research Council 1995).

Sampling design is critical to the effectiveness of a monitoring program (Dixon et al. 1998) because it determines the quality of the resulting parameter estimates (Thompson et al. 1998). In many instances, the parameter of interest is a long-term trend in population size or community structure. The ability of a design to provide reliable trend estimates can be interpreted as the probability that it will detect a trend, should one exist. Statistical power is a suitable measure to gauge the influence of sample effort on

¹University of Alberta

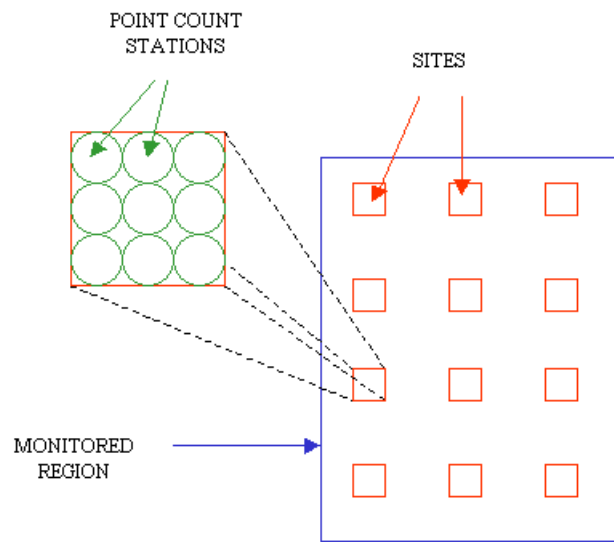
data quality (e.g., Peterman 1990, Gibbs et al. 1998). A low-power monitoring program will have little chance of detecting all but catastrophic trends, thereby providing an illusion of stability and preventing remedial action required to conserve biodiversity (Fairweather 1991).

The power to detect a trend of specified magnitude, with a given level of significance, is negatively related to variability and positively related to sample effort. Although increasing sample effort increases power to detect trend, excessive sampling wastes limited monitoring resources (Bernstein and Zalinski 1983). Extensive monitoring requirements that must be met under budgetary constraints have made cost-effective monitoring a research priority (Committee on Environment and Natural Resources 1996, Schneider et al. 1999). Funding limitations impose trade-offs between the allocations of effort to sample effort components, which compete for monitoring resources (Millard and Lettenmaier 1986). It follows that cost-effective monitoring design requires an understanding of the efficiency of allocating effort to sample effort components.

Here we investigate cost-effective sampling strategies for large-scale avian monitoring. As a group, birds are suited to monitoring because there are well-established (Bibby et al. 1992) and efficient (e.g., Croonquist and Brooks 1991) sampling protocols that achieve informative data due to the diverse responses of bird species to disturbance (Steele et al. 1984). We evaluate monitoring strategies in the context of the forested region of Alberta, Canada, where development of a large-scale forest biodiversity monitoring program is underway (Farr et al. 1999).

The design of a large-scale avian monitoring program can be generalized as follows (Fig. 1). Sampling occurs at multiple sites located randomly or systematically within the region of interest. Sampling spans multiple years, although monitoring may not occur every year. Within a sample year, sites are surveyed one or more times. At each site, birds are counted using one or more point-count stations, the preferred counting method for large-scale monitoring in forested regions (Verner 1988). We focus on the power and cost trade-offs implicit in allocating effort to the number of sites, the frequency with which these sites are sampled, and the number of stations and surveys used to sample the sites.

Fig. 1. Generalized large-scale avian monitoring program design assumed during efficiency analysis.



METHODS

To identify cost-effective sampling strategies, power and cost analyses were integrated. The power of sampling designs to detect trends in a selection of bird species and community metrics was estimated by simulating monitoring data, using mean and variance parameters estimated from empirical data. Based on the power estimates, a functional relationship between sample effort and power was derived for each monitoring target. A functional relationship between sample effort and cost was also identified. Partial derivatives of the functional relationships were taken to estimate rates of power and cost gain with sample effort. Comparison of these rates across sample effort parameters identified cost-effective sampling strategies.

Estimation of avian mean and variance parameters

Stochastic variability enters data collected via three components, according to Fig. 1. Data fluctuate at a site across years because of temporal variability in abundance and sample (methodological) error, which are captured by the within-site temporal variance (τ_1^2) component. Sample error, which includes both variation in the proportion of birds present that are detectable, e.g., singing (due to factors such as

weather) and variation in the proportion of birds that are detected due to observer ability, is expected to decrease with increased sample effort. This effect was evaluated by estimating τ_1^2 at multiple levels of within-site effort. In addition to temporal variability, data fluctuate spatially in two ways. The distribution of individuals across sites at the start of the monitoring period is not uniform, and is accounted for by between-site initial variance (τ_2^2). Trend magnitude also varies across sites about the regional trend mean, and is accounted for by between-site trend variance (τ_3^2). Because sample error is included in τ_1^2 , it was extracted from spatial variance estimates to the extent possible, to avoid double counting. Simulation of data therefore required estimation of τ_1^2 , τ_2^2 , τ_3^2 , as well as mean initial abundance (θ_0). Because we were interested in evaluating monitoring program design for the boreal forest of Alberta, data collected from this region were used for parameter estimation. τ_1^2 and θ_0 were estimated using data collected at three 100-ha control sites near Calling Lake, in north-central Alberta, over a 6-year period (1993–1998). The sites, part of a long-term experimental fragmentation study (Schmiegelow et al. 1997), occur in boreal mixed-wood forest, an ecotype representative of much of northern Alberta. At each site, 20 stations were sampled four times each year (surveys) during the breeding season (from late May to late June), at 10-d intervals, using 5-min point counts. Control sites (old forest devoid of forestry activity) were used to minimize the effect of non-natural (treatment) variation, which can exaggerate within-site temporal variation (Gibbs et al. 1998). τ_2^2 and τ_3^2 were estimated using Breeding Bird Survey (BBS) data (USGS Patuxent Wildlife Research Center 2000) collected from the closed boreal forest physiographic region of Alberta. The BBS is a volunteer-based long-term monitoring program spanning Canada and the United States. The program consists of routes that are sampled by volunteers each year during the breeding season (May or June). At each route, birds are sampled at 50 point count stations, each 3 min in duration, along 24.5-mile [39.4 km] routes (see Droege [1990] for additional BBS design information). All variance parameters were estimated as coefficients of variation (standard deviation divided by the mean) to make the estimates suitable for use in simulation studies. Details on parameter estimation are available in Appendix 1.

Species and community metric selection

To include species covering the spectrum of variance characteristics, we selected them using a plot of τ_1^2 and τ_2^2 estimates. As broad community measures, total species richness, species richness of the ground-nesting guild (see Appendix 1), and the Shannon-Weiner diversity index (Margalef 1958) were also included.

Power simulations

Power of 160 sampling designs (Table 1) was estimated for each species and community metric, using mean abundance and variance estimates in a Monte Carlo simulation approach (Appendix 2), following the methodology of Gibbs et al. (1998). Simulated trends were exponential and 20 years in length, with magnitudes of -3% and -1% per year for populations and community metrics, respectively.

Table 1. Sampling designs for a large-scale avian monitoring program evaluated in power simulations.

Effort parameter	Effort levels
Number of sites	10, 20, 30, 40, 50
Sampling frequency (sample years)	5, 6, 8, 11, 21
Number of point-count stations	4, 9, 16
Number of surveys	1, 2, 3, 4

Modeling the effect of sample effort on power and cost

Multiple regression was applied to develop models for each target that expressed power as a function of sample effort. A linear model was unsuitable due to the asymptotic nature of power as it approaches the 100% maximum. Instead, power estimates were fit to the following nonlinear model:

$$\text{power} = 1 - \exp[-1 * (a_0 + a_1s + a_2f + a_3p + a_4r + a_5sf + a_6sp + a_7sr + a_8fr + a_9fp + a_{10}pr)] \quad [1]$$

where a_{0-10} are model coefficients and s , f , p , and r are number of sites, sample frequency, point-count stations, and surveys, respectively. The model was fit to the power estimates by minimizing the residual sum of squares using procedure NLS in S-Plus (MathSoft 1999). Residual analysis indicated no violations of normality and homogeneity assumptions. A model relating sample effort to program cost is presented in Appendix 3.

Determining efficient sampling strategies

To evaluate the efficiency of allocating sample effort to each effort parameter, we estimated partial derivatives of the power and cost models for each effort parameter. Partial derivatives were interpreted as the increase in power or cost achieved by increasing the number of sites, sample frequency, point-count stations, or surveys by one, within the range of effort levels simulated (Table 1). Average effort levels of designs achieving between 80% and 90% power for each species/community metric were used to solve partial derivatives, in order to evaluate the efficiency of sampling strategies in close proximity to the 90% power goal. The significance of power partial derivatives, which were estimated from regression models, was tested using z tests. The standard errors of partial derivatives were estimated using the Krinsky-Robb method (Krinsky and Robb 1986). This required generation of 10,000 coefficient vectors by drawing random vectors from a multivariate normal distribution with mean and variance as estimated from nonlinear regression. Using the generated coefficient vectors, we calculated 10000 partial derivatives to approximate the distribution of the partial derivative, from which standard error was calculated. Normally distributed residuals from the regressions validated the required assumption of multivariate normality (Myers 1990).

Because power and cost were expressed in different units, it was necessary to convert partial derivatives to proportions for comparison. For example, partial derivatives of power related to each effort parameter were divided by the sum of power partial derivatives across effort parameters. We calculated 95% confidence intervals of power proportions by estimating 10,000 power proportions using the approximated power partial derivative distributions, and eliminating the top and bottom 2.5% of the distributions. For a given effort parameter, if the power proportion was greater than the cost proportion, the rate at which power increased by allocating effort

to that parameter, relative to the other parameters, was greater than the rate at which cost increased, indicating efficiency. If, on the other hand, the cost proportion was greater, the cost of allocating effort to that parameter, relative to the other parameters, was greater than the power gains achieved, indicating inefficiency.

Fig. 2. Temporal and spatial variability in bird species abundance and community metrics. Variance is expressed as coefficients of variation. Temporal variance refers to within-site temporal variance calculated using data collected near Calling Lake, Alberta, Canada from 1993 to 1998 using 20 point-count stations per site and four surveys per year. Spatial variance refers to between-site initial variation calculated using BBS (Breeding Bird Survey) data (from the closed boreal forest physiographic region of Alberta) collected from 1992 to 1998. The horizontal and vertical lines dissecting the plot represent mean between-site initial variation and median within-site temporal variation, respectively, across all species, excluding those with maximum within-site temporal variation (2.449). Median within-site temporal variance was used because of the skewed distribution of the parameter across species. An extreme and a moderate example were selected from each of the four variance categories delineated by the plot for inclusion in the efficiency analysis, for a total of eight species. The four-letter codes identify the eight species and three community metrics selected as targets for simulation studies: BCCH, Black-capped Chickadee *Parus atricapillus*; BGNW, Black-throated Green Warbler *Dendroica virens*; BHCO, Brown-headed Cowbird *Molothrus ater*; GRJA, Gray Jay *Perisoreus canadensis*; PIWO, Pileated Woodpecker *Dryocopus pileatus*; WBNU, White-breasted Nuthatch *Sitta carolinensis*; WTSP, White-throated Sparrow *Zonotichia albicollis*; YWAR, Yellow Warbler *Dendroica petechia*; SR, species richness; GN, richness of the ground-nesting guild; SW, Shannon-Weiner diversity index.

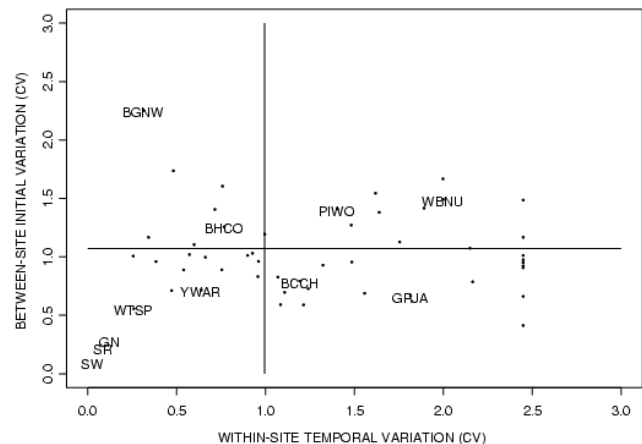


Table 2. Within-site temporal variance estimates (τ_1^2), expressed as coefficients of variation. Two estimates are provided for each species at each within-site effort level. The first estimate was generated using fixed timing of surveys, whereas the second estimate was generated using variable timing of surveys. Target codes are as in Fig. 2. Variance estimates are derived from data collected near Calling Lake, Alberta, Canada from 1993 to 1998.

Target species	No. surveys, four stations				No. surveys, nine stations				No. surveys, 16 stations			
	1	2	3	4	1	2	3	4	1	2	3	4
BCCH	2.389	2.274	2.020	1.903	2.185	1.922	1.732	1.570	2.035	1.740	1.455	1.264
	2.327	2.188	2.003	1.903	2.120	1.834	1.657	1.570	1.887	1.610	1.385	1.264
BGNW	1.219	0.806	0.734	0.615	0.773	0.544	0.470	0.407	0.605	0.442	0.367	0.348
	1.073	0.809	0.680	0.615	0.718	0.514	0.442	0.407	0.562	0.427	0.381	0.348
BHCO	1.978	1.747	1.689	1.629	1.659	1.464	1.448	1.229	1.422	1.399	1.331	0.925
	2.096	1.887	1.730	1.629	1.839	1.516	1.349	1.229	1.578	1.314	1.102	0.925
GRJA	1.849	1.849	1.849	2.326	1.849	1.849	1.849	2.256	1.549	1.549	1.549	1.848
	2.391	2.320	2.284	2.326	2.387	2.269	2.257	2.256	2.296	1.937	1.830	1.848
PIWO	2.449	2.449	2.449	2.064	2.449	2.449	2.449	1.857	2.449	2.449	2.449	1.314
	2.406	2.322	2.260	2.064	2.410	2.269	2.131	1.857	2.256	1.977	1.607	1.314
WBNU	2.449	2.449	2.499	1.999	2.449	2.449	1.999	2.224	2.449	2.449	1.549	1.999
	2.449	2.449	2.449	1.999	2.387	2.314	2.299	2.224	2.366	2.202	2.127	1.999
WTSP	0.671	0.468	0.406	0.357	0.573	0.383	0.353	0.313	0.484	0.350	0.327	0.276
	0.615	0.460	0.391	0.357	0.513	0.383	0.340	0.313	0.439	0.346	0.304	0.276
YWAR	1.801	1.358	1.158	0.991	1.462	1.087	1.036	0.795	1.218	0.964	0.990	0.688
	1.525	1.250	1.096	0.991	1.277	1.041	0.890	0.795	1.138	0.894	0.768	0.688
SR	0.206	0.153	0.143	0.122	0.141	0.100	0.099	0.105	0.117	0.089	0.092	0.086
	0.234	0.162	0.139	0.122	0.200	0.139	0.111	0.105	0.186	0.123	0.099	0.086
GN	0.329	0.231	0.230	0.185	0.247	0.168	0.183	0.160	0.247	0.158	0.164	0.124
	0.321	0.249	0.216	0.185	0.279	0.209	0.172	0.160	0.232	0.201	0.162	0.124
SW	0.100	0.069	0.057	0.046	0.062	0.044	0.035	0.033	0.047	0.029	0.023	0.026
	0.129	0.077	0.057	0.046	0.080	0.054	0.043	0.033	0.069	0.046	0.033	0.026

RESULTS

In addition to the three community metrics, eight

species were selected for the analysis (Fig. 2), two from each of the following categories: ubiquitous (low spatial and temporal variance), temporally fluctuating

(low spatial and high temporal variance), spatially patchy (high spatial and low temporal variance), and irruptive (high spatial and temporal variance). Averaged across selected species and community metrics, τ_2^2 was greatest in magnitude (1.143 for species and 0.203 for community metrics), followed by τ_1^2 (1.083 and 0.0787) and τ_3^2 (0.303 and 0.0507). Full presentation of parameter estimates, including the influence of within-site effort on τ_1^2 , is provided in

Tables 2–5. Relationships between variability and species abundance were explored by ranking the order of abundance and calculating a correlation coefficient for each variance parameter. There was a strong negative relationship between abundance and τ_1^2 ($r = -0.976$; $P = 0.000$). Abundance rank was also negatively correlated with the other two parameters, although the relationships were weaker ($r = -0.714$; $P = 0.047$ for τ_2^2 and $r = -0.476$; $P = 0.233$ for τ_3^2).

Table 3. Between-site initial variance (τ_2^2) and between-site trend variance (τ_3^2) estimates, expressed as coefficients of variation. Target codes are as in Fig. 2. Variance estimates are derived from BBS (Breeding Bird Survey) data collected from the closed boreal forest physiographic region of Alberta, Canada from 1992 to 1998 (τ_2^2) and 1989 to 1998 (τ_3^2).

Target	BCCH	BGNW	BHCO	GRJA	PIWO	WBNU	WTSP	YWAR	SR	GN	SW
τ_2^2	0.793	2.254	1.257	0.666	1.410	0.491	0.557	0.716	0.222	0.283	0.104
τ_3^2	0.487	0.167	0.404	0.301	0.290	0.449	0.139	0.188	0.038	0.050	0.064

In general, power increased with each sample effort parameter (Figs. 3–6). Power was highly sensitive to τ_1^2 , moderately sensitive to τ_2^2 , and only minimally sensitive to τ_3^2 (Table 6). Eq. 1 effectively modeled the relationship between power and sample effort, explaining much of the variability in power estimates ($R^2 = 0.94 - 0.98$). Models were not generated for species or community metrics for which all sampling designs achieved > 90% power. This excluded the Shannon-Weiner diversity index, species richness, and the White-throated Sparrow (*Zonotrichia albicollis*). Power to detect trend increased as the cost of monitoring programs increased, although the relationship was highly nonlinear (Fig. 7). Power partial derivatives were generally positive and significant (Table 7), as were cost partial derivatives except for point-count stations, reflecting the assumption that increasing the number of stations across the range simulated (4–16) incurred no additional cost (Appendix 3). This assumption, combined with positive rates of increase in power with point-count stations, implied efficiency. Thus point-count stations were excluded from further analysis. Table 8 presents results of the efficiency analysis, evaluated with reference to sampling designs achieving between 80% and 90% power. To evaluate the sensitivity of efficiencies to power level, the analysis was repeated using average effort levels from designs achieving: <60%, 60–70%, 70–80%, and

>90% power. Efficiencies were consistent across power intervals, with the exception of the efficiency of surveys for the Yellow Warbler (*Dendroica petechia*) and for species richness of the ground-nesting guild. Although surveys were marginally efficient for these targets when designs achieved > 80% power, they were marginally inefficient when designs achieved < 80% power.

DISCUSSION

Recent political support for biodiversity monitoring is providing new opportunities to obtain spatially and temporally extensive data sets. Maximizing these opportunities to achieve exciting new analyses requires efficient use of available funding. We identified cost-effective sampling strategies for forest birds, using existing data, power analysis, and statistical modeling.

Variability of bird species and community metrics

Because of the sensitivity of power to variability, and the wide range in variability across species, variance estimates were used to categorize and select targets for our analysis. Our four variance categories reflect species life-history traits. Species in the ubiquitous

category have broad niches that buffer them from both temporal and spatial environmental variation. In comparison, spatially patchy species have narrow niches that allow occupancy of only a subset of available habitat, although occupancy is consistent across years. Species in the temporally fluctuating and irruptive categories may be tied to temporally fluctuating habitat characteristics, such as food

availability. Other species may appear to be temporally variable due to sampling issues. For example, the Pileated Woodpecker (*Dryocopus pileatus*) is not an irruptive species, but its large home range relative to the scale of point-count stations resulted in variable counts, even in the presence of a potentially stable population.

Table 4. Mean abundance θ estimates for species and community metrics selected as targets for sample design evaluation. Two estimates are provided for each species at each within-site effort level. The first estimate was generated using fixed timing of surveys, whereas the second estimate was generated using variable timing of surveys. The number of surveys is indicated by r . Target codes are as in Fig. 2. Variance estimates are derived from data collected near Calling Lake, Alberta, Canada from 1993 to 1998.

Target	Four stations				Nine stations				16 stations			
	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 1$	$r = 2$	$r = 3$	$r = 4$
BCCH	0.299 0.291	0.163 0.214	0.163 0.163	0.153 0.153	0.380 0.377	0.259 0.341	0.258 0.287	0.281 0.281	0.500 0.572	0.347 0.525	0.417 0.457	0.458 0.458
BGNW	1.694 1.598	1.729 1.574	1.483 1.575	1.579 1.579	3.880 3.638	3.963 3.586	3.392 3.652	3.639 3.639	7.139 6.420	7.181 6.561	6.185 6.566	6.563 6.563
BHCO	0.293 0.294	0.228 0.205	0.181 0.183	0.172 0.172	0.463 0.461	0.407 0.394	0.333 0.374	0.365 0.365	0.806 0.718	0.667 0.637	0.556 0.601	0.604 0.604
GRJA	0.389 0.334	0.194 0.186	0.130 0.140	0.108 0.108	0.389 0.349	0.194 0.215	0.130 0.154	0.125 0.125	0.500 0.440	0.250 0.364	0.167 0.286	0.229 0.229
PIWO	0.250 0.208	0.125 0.118	0.083 0.084	0.077 0.077	0.375 0.252	0.188 0.158	0.125 0.125	0.117 0.117	0.417 0.336	0.208 0.206	0.139 0.194	0.188 0.188
WBNU	0.167 0.182	0.083 0.092	0.056 0.063	0.049 0.049	0.167 0.211	0.083 0.115	0.083 0.080	0.066 0.066	0.167 0.236	0.083 0.134	0.111 0.099	0.083 0.083
WTSP	4.426 5.108	4.632 5.041	4.972 5.092	5.076 5.076	9.926 11.27	10.44 11.35	11.16 11.37	11.38 11.38	17.33 20.09	18.22 19.89	19.42 19.90	19.91 19.91
YWAR	0.637 0.997	0.757 0.928	0.798 0.923	0.922 0.922	1.273 1.992	1.581 1.979	1.671 1.988	1.998 1.998	1.944 3.085	2.375 3.133	2.602 3.144	3.146 3.146
SR	11.81 11.19	15.29 15.63	17.24 18.43	20.43 20.43	16.73 15.84	20.57 21.05	23.14 24.44	26.77 26.77	20.39 19.34	24.58 25.63	27.44 29.16	31.64 31.64
GN	3.898 3.635	4.588 4.668	4.903 5.282	5.731 5.731	4.972 4.700	5.556 5.881	5.954 6.580	7.139 7.139	5.611 5.437	6.222 6.848	6.722 7.628	8.306 8.306
SW	3.179 3.062	3.417 3.409	3.486 3.562	3.640 3.640	3.558 3.459	3.689 3.705	3.742 3.798	3.855 3.855	3.733 3.641	3.814 3.838	3.860 3.918	3.956 3.956

Ubiquitous species will be easiest to monitor, due to the negative relationship between species abundance and the variance parameters. Their broad habitat requirements, however, imply less sensitivity to disturbance, reducing their value for monitoring. Rare species, although of greater monitoring interest, will require increased sampling intensity to mitigate high sample error (Link et al. 1994), as well as high natural variability, perhaps caused by specialized habitat requirements. If the expense associated with

monitoring such species is prohibitive, careful selection of target species that are sensitive to disturbance, yet exhibit low variability, is recommended. Temporal variation tends to influence power to a greater extent than does spatial variation. Species in the spatially variable category (i.e., the Black-throated Green Warbler, *Dendroica virens*), may be examples of efficient targets because of their habitat specificity and relative temporal stability.

Table 5. Percentage reductions in within-site variance (τ_1^2) due to increasing the number of surveys and stations and using fixed rather than randomly scheduled surveys. The response of τ_1^2 to the number of surveys was evaluated as the percentage reduction in τ_1^2 caused by increasing the number of surveys from one to four, averaged across the three station levels (assuming fixed survey schedules). The response of τ_1^2 to the number of stations was evaluated as the percentage reduction in τ_1^2 caused by increasing the number of stations from four to 16, averaged across the four survey levels (assuming fixed survey schedules). To evaluate the influence of timing of the surveys, the percentage reduction in τ_1^2 caused by using fixed rather than randomly scheduled surveys, averaged across all effort levels except when surveys equalled four, was calculated. Target codes are as in Fig. 2. Variance estimates are derived from data collected near Calling Lake, Alberta, Canada from 1993 to 1998.

Target	Reduction (%) in τ_1^2		
	Due to surveys	Due to stations	Due to survey timing
BCCH	28.79	25.04	-4.53
BGNW	46.46	47.24	-5.39
BHCO	26.17	28.11	0.44
GRJA	-22.37	17.31	21.01
PIWO	28.75	9.08	-13.88
WBNU	15.31	9.50	1.67
WTSP	45.05	23.81	-5.46
YWAR	44.70	26.62	-12.39
Species average	26.61	23.34	-0.02
SR	30.93	37.55	17.20
GN	42.93	29.55	4.07
SW	48.49	53.52	21.29
Metric average	40.78	40.21	14.19

Monitoring of community metrics provides an alternative. In this study, all community metrics exhibited less variance than even the most common species, which translated into low sample effort requirements. However, two issues must be resolved. First, the sensitivity of community metrics to disturbance must be determined. We used a smaller

magnitude change in community metric power analyses to account for a presumed reduction in sensitivity to disturbance, relative to populations of species. The magnitude of the reduction was subjective, however, and may not have adequately accounted for the relative stability of community metrics. Second, it is not clear which community

metric, or combination of metrics, provides meaningful information on the status of the bird community, and interpretation of community metrics is difficult. Nevertheless, the relative low cost of monitoring community metrics is promising and warrants increased research into their use.

Power estimates, and therefore our conclusions regarding cost-effective sampling strategies, were contingent on assumptions made when estimating variance parameters. A discussion of these assumptions is presented at the end of Appendix 1.

Table 6. Sensitivity of power to within-site variance (τ_1^2), between-site initial variance (τ_2^2), and between-site trend variance (τ_3^2). For each parameter, the power to detect -3% trends over 20 years was estimated at maximum and minimum values of the parameter as estimated across species and effort levels, while keeping the other parameters constant at their average value across species and effort levels and keeping the number of sites and the sampling frequency at minimum levels (10 sites sampled every five years). Sensitivity was then calculated by subtracting power at the maximum value from power at the minimum value of that variance parameter.

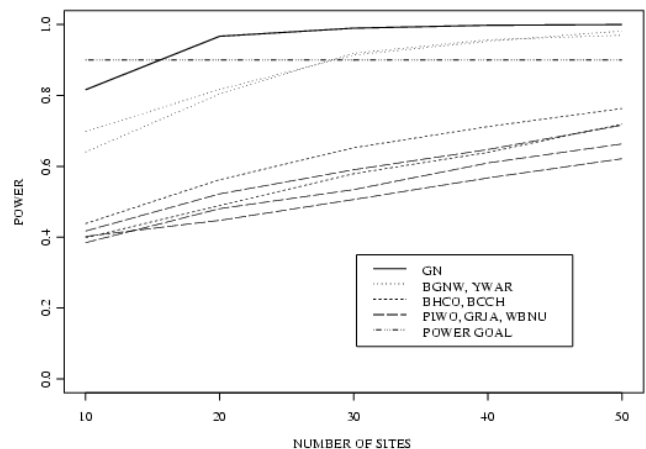
Variance parameter	τ_1^2	τ_2^2	τ_3^2
Sensitivity of power	0.551	0.116	0.055

Cost-effective sampling strategies

Although all effort parameters generally had a positive and significant influence on power, the magnitude of their influence and costs differed. Therefore, to establish efficient sampling design strategies, we compared rates of increase in power to rates of increase in cost across effort parameters. To focus on the response of power and cost to sample effort near the 90% power goal, we used average effort levels across designs achieving between 80% and 90% power. Efficiencies were consistent across power levels, however, indicating that the cost-effective sampling strategies reported here are not restricted to designs achieving between 80% and 90% power. Because optimal effort allocation depends on the relative magnitude of variance parameters (Bernstein

and Zalinski 1983), one would expect sampling efficiencies to vary across monitoring targets that exhibit different variance characteristics. The inclusion of multiple species and community metrics allowed comparison of sampling efficiencies across different targets, which identified consistent sampling efficiencies across targets, and efficiencies contingent on the variance characteristics of the monitoring target.

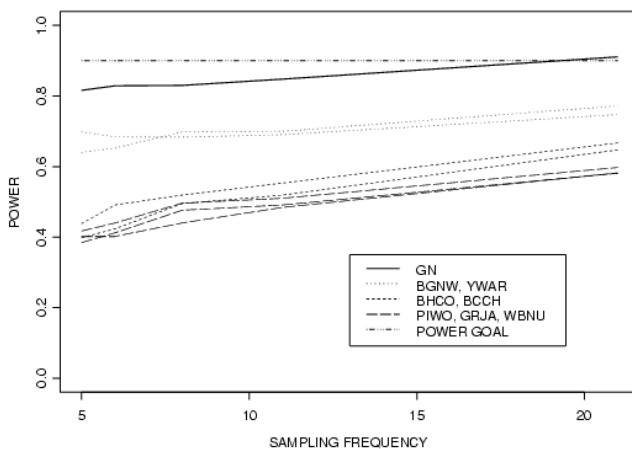
Fig. 3. Response of power to the number of sites as estimated using Monte Carlo simulations. In the figure, the number of sites is varied across the range, while the remaining sample effort parameters are kept fixed at minimum levels (sampling every five years, four point-count stations, and one survey). Species and community metric codes are as in Fig. 2. Species are listed from highest to lowest power within each group in the legend.



When designing a monitoring program, one must decide whether to monitor more sites infrequently, or fewer sites frequently. Allocation of effort to the number of sites was efficient for all monitoring targets, suggesting that increasing site density is a component of an efficient avian monitoring design (see also Link et al. 1994). Previous evaluation of the power of long-term avian sampling designs has assumed that sampling occurs every year (e.g., Gibbs and Melvin 1997, Gibbs et al. 1998, Lewis and Gould 2000). However, high sampling frequency is costly. We used simulations to estimate power for a range of sampling frequencies to evaluate the efficiency of allocating effort to sampling frequency. In general, the rate at which power increased with sample frequency was greater for temporally variable species. For temporally stable monitoring targets, however, the increase in data

quality achieved by increasing sampling frequency was small, and thus inefficient. These results suggest that more sites should be sampled relatively infrequently, although higher sampling frequency will be required for temporally variable targets. The number of sites rises in importance when sampling uncommon species, if sites with zero counts over the study period are eliminated from the analysis, as is the case with the BBS (Thomas 1997). This effectively reduces the sample size, so that the number of sites must be increased to account for the proportion expected to have zero counts. Similarly, more frequent monitoring will be necessary if detection of cyclical or irregular trends is required, rather than the constant trends simulated here (Scott 1998). If counts are expected to vary coherently across sites about a temporal trend (i.e., if they exhibit cyclic tendency), panel designs are an efficient strategy whereby some sites are sampled more frequently than others (Urquhart et al. 1998).

Fig. 4. Response of power to sample frequency as estimated using Monte Carlo simulations. In the figure, sample frequency is varied across its range, while keeping the remaining sample effort parameters fixed at minimum levels (10 sites, four point-count stations, and one survey). Species and community metric codes are as in Fig. 2. Species are listed from highest to lowest power within each group in the legend.



Two strategies for reducing sample error were evaluated: increasing the number of point-count stations and increasing the number of surveys at a station within a year. Sampling the maximum number (16) of point-count stations considered was always efficient because it increased power and incurred no

additional cost compared to sampling fewer point count stations. Although sampling of additional

Fig. 5. Response of power to the number of point-count stations as estimated using Monte Carlo simulations. In the figure, the number of point-count stations is varied across its range, while keeping the remaining sample effort parameters fixed at minimum levels (10 sites, sampling every five years, and one survey). Species and community metric codes are as in Fig. 2. Species are listed from highest to lowest power within each group in the legend.

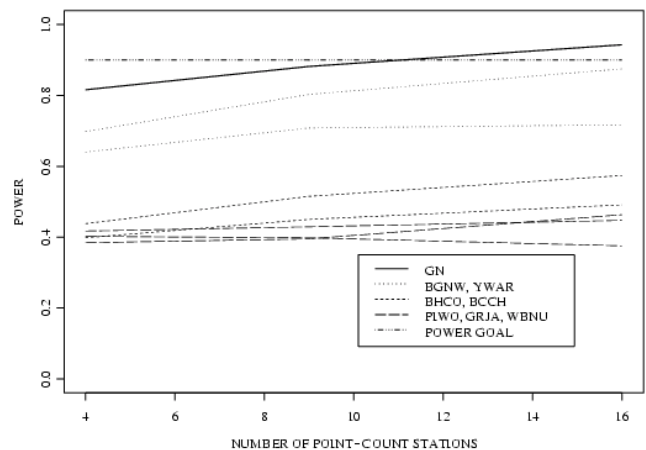
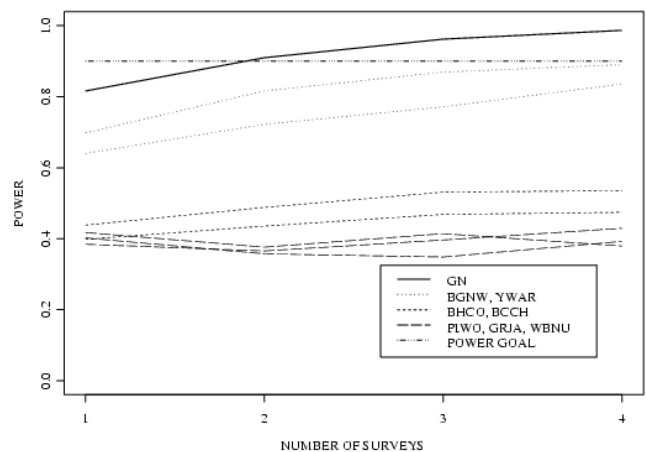


Fig. 6. Response of power to the number of surveys, as estimated using Monte Carlo simulations. In the figure, the number of surveys is varied across its range, while keeping the remaining sample effort parameters fixed at minimum levels (10 sites, sampling every five years, and four point-count stations). Species and community metric codes are as in Fig. 2. Species are listed from highest to lowest power within each group in the legend.



stations is not, in reality, free, the associated cost will be minimal compared to the cost of accessing sites, particularly where access is limited. In our case, one crew could visit only a single site on a given day due to logistical constraints, and could sample up to 16 point-count stations. As long as additional within-site sampling (i.e., point-count stations) reduces sample error, high within-site, within-survey sample effort should be a component of a large-scale monitoring program. Repeated surveys within a year, on the other hand, were generally inefficient. Although they are marginally efficient for some temporally stable species and metrics, multiple surveys will usually be unnecessary to detect trends in these targets. Surveys

were dramatically inefficient for species requiring high sample effort to detect trends, suggesting that repeated visits to a site within a year are not cost-effective (see also Link et al. 1994). Allocating effort to achieve consistent timing of surveys across years was also inefficient, based on the negligible effect of timing on temporal variance. This result is fortunate, as consistent timing of surveys is difficult to achieve in large-scale monitoring. However, if potential sample dates were to extend beyond the one-month period during which our baseline data were collected (i.e., the peak of breeding activity for many species), the importance of survey timing should increase.

Table 7. Rates of change (and *P* values) in power to detect -3% per year trends in species' abundance and -1% per year trends in community metrics over 20 years, for four sample effort parameters. Rates of change for each species/community metric were calculated by taking partial derivatives of the power model and solving using average effort levels of designs achieving between 80% and 90% power for each species/community metric. Species and community metric codes are as in Fig. 2.

Species	No. sites	No. sample years	No. point-count stations	No. surveys
BCCH	0.0058 (0.000)	0.0133 (0.000)	0.0095 (0.000)	0.0237 (0.000)
BGNW	0.0114 (0.000)	0.0016 (0.000)	0.0169 (0.000)	0.0643 (0.000)
BHCO	0.0062 (0.000)	0.0100 (0.000)	0.0104 (0.000)	0.0296 (0.000)
GRJA	0.0054 (0.000)	0.0145 (0.000)	0.0046 (0.000)	-0.0003 (0.891)
PIWO	0.0052 (0.000)	0.0142 (0.000)	0.0068 (0.000)	0.0224 (0.000)
WBNU	0.0058 (0.000)	0.0137 (0.000)	0.0016 (0.000)	0.0026 (0.147)
YWAR	0.0156 (0.000)	0.0043 (0.000)	0.0053 (0.000)	0.0470 (0.000)
GN	0.0265 (0.000)	0.0075 (0.000)	0.0125 (0.000)	0.1431 (0.000)
Average	0.0103	0.0098	0.0084	0.0416

Extending the analysis to additional taxa

Comprehensive monitoring of biodiversity must include diverse taxa. This ambitious goal might best be achieved through a single, broad-based biodiversity monitoring program, rather than a collection of specialized programs. To facilitate design of such a program, our analysis should be extended to multiple taxa. An extensive search for data sets to facilitate parameter estimation for our study area revealed severe limitations (S. Rangen and W. Schaffer, *unpublished report*). In shortest supply were data collected across multiple years (necessary to estimate τ_1^2 and τ_3^2); in addition to birds, such data were only available for components of terrestrial vegetation and aquatic macroinvertebrates. This stresses the need for

implementation of long-term monitoring programs.

Alternative sample design

We interpreted the problem of designing a cost-effective monitoring program as achieving high power to detect trend at low cost. An alternative framework is an optimal experimental design (Silvey 1980), in which sample design is viewed as a constrained optimization problem. The goal is to select the design that achieves maximum information under design constraints, such as cost. A model is assumed for the phenomenon of interest, and the analysis proceeds through an evaluation of the inverse of the error covariance matrix of the best linear unbiased

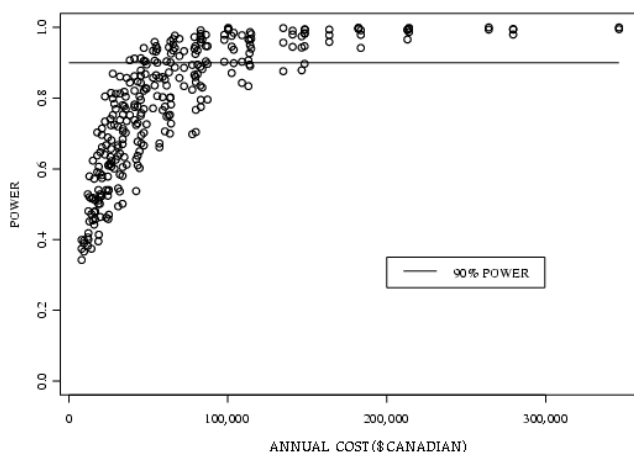
estimator. This matrix, known as the Fisher information matrix, contains information about measurement errors and parameter sensitivities,

allowing the quality of parameter estimates achieved by a design to be quantified (Versyck et al. 1999).

Table 8. Relative efficiencies (and 95% confidence intervals) of effort parameters for detecting 20-year trends in various targets. Elements in the table represent power proportions divided by cost proportions for each effort parameter, excluding stations. An efficiency value greater than one indicates that the effort parameter is efficient. Target codes are as in Fig. 2.

Target	Sites		Frequency		Surveys	
	Mean	CL	Mean	CL	Mean	CL
BCCH	2.315	(2.186, 2.457)	1.591	(1.476, 1.716)	0.742	(0.701, 0.779)
BHCO	2.138	(2.037, 2.252)	1.104	(1.014, 1.196)	0.875	(0.844, 0.904)
BGNW	1.159	(1.109, 1.211)	0.138	(0.112, 0.164)	1.146	(1.136, 1.156)
GRJA	5.121	(4.256, 6.413)	4.833	(4.043, 6.011)	-0.017	(-0.330, 0.192)
PIWO	2.117	(1.952, 2.307)	1.884	(1.722, 2.068)	0.704	(0.648, 0.753)
WBNU	4.668	(4.027, 5.545)	5.134	(4.446, 6.075)	0.141	(-0.057, 0.286)
YWAR	1.686	(1.622, 1.751)	0.356	(0.312, 0.399)	1.030	(1.013, 1.047)
GN	1.532	(1.409, 1.658)	0.374	(0.328, 0.420)	1.024	(1.006, 1.040)

Fig. 7. Power to detect -3% per year exponential trends in Black-capped Chickadee abundance over 20 years vs. annual program cost. Each circle represents one of the 160 sampling designs evaluated (see Table 1). Power is estimated using Monte Carlo simulations; cost is estimated using a simplified cost model (Appendix 3).



There is a fundamental difference in the monitoring philosophies implicit in power analysis and optimal design. Our power analysis assumed that the monitoring goal is regional detection of long-term trends using fixed sites randomly located within the region of interest. There is no consideration of site attributes, such as disturbance or habitat type. Such a program is not designed to answer a particular question, such as the effect of a disturbance type on biodiversity, but simply to detect broad-scale changes in biodiversity, regardless of the cause. This does not maximize information gain, nor can it link biodiversity trends with causal factors. Thus it restricts the translation of monitoring findings into policy change. Optimal experimental design takes a very different approach. Treatment levels at potential sites are explicitly considered in the design, in order to achieve the best estimate of the parameter of interest. Such a design has greater potential for providing estimates of the effect of specific management policies on biodiversity, thereby facilitating adaptive

management. However, the long-term utility of such a monitoring program may be limited. Data from a program optimally designed to answer today's questions are unlikely to address tomorrow's environmental dilemmas.

The two design strategies produce different styles of monitoring programs, each with its strengths and weaknesses. Instead of relying solely on either strategy, a hybrid approach would exploit the strengths of both. This would consist of optimal designs that track the effects of specific disturbances, superimposed on a large-scale design that achieves regional trend estimates with high power.

Responses to this article can be read online at:
<http://www.consecol.org/vol6/iss2/art11/responses/index.html>

Acknowledgments:

This article is the result of graduate research, and benefited greatly from the perspectives of supervisory committee members Brad Stelfox, Vic Adamowicz, and Peter Boxall. Steve Cumming provided statistical and computer programming guidance. Interaction with the Alberta Forest Biodiversity Monitoring Program Technical Committee, especially Dan Farr, was valuable. Without the hard work of those who collected the bird data, the research reported here would not have been possible. Many thanks to Breeding Bird Survey volunteers and to those who have collected data over the years at Calling Lake. Thank you also to those who provided computer code. This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Sustainable Forest Management Network, and the Alberta Forest Biodiversity Monitoring Program. The article benefited from the comments of two anonymous reviewers. Finally, thank you to family and friends for your support.

APPENDIX 1. Avian Mean and Variance Parameters

Mean and variance parameter estimation methods

To estimate θ_0 and τ_1^2 , the mean response and coefficient of variation at each Calling Lake (Alberta, Canada) site across the 6-year period was calculated and averaged across the three sites. The response at a site in a given year was total detections across all point-count stations averaged across surveys. Sites for which no individuals of target species were detected over the 6-year period were excluded to avoid undefined coefficients of variation. To evaluate the effect of within-site sample effort on τ_1^2 and θ_0 , subsamples of the data set were taken and the parameters were estimated using the reduced data set. Within-site effort levels of 4, 9, and 16 stations and 1, 2, 3, and 4 surveys were evaluated, resulting in 12 estimates of θ_0 and τ_1^2 for each monitoring target. Two strategies were used when subsampling surveys. To simulate a monitoring program in which the timing of site visits can be kept constant from year to year, survey 2 was always sampled, and surveys 3, 4, and 1 were added successively as the number of surveys increased from two to four. To simulate a monitoring program in which timing of site visits cannot be kept constant from year to year, the surveys were randomly selected each year, without replacement. This process was repeated 100 times to achieve 100 subsamples of randomly selected surveys across the six years. When subsampling to reduce the number of stations, multiple estimates of τ_1^2 and θ_0 were generated for each site to represent the various contiguous arrangements of the reduced number of stations. The mean of these estimates was taken as an overall estimate for each site.

We estimated τ_2^2 as the response variance between BBS routes in the closed boreal forest region of Alberta within a year. This estimate reflected not only variability between routes, but also variability due to sample error at each route (Link et al. 1994). Sample error, already accounted for in τ_1^2 , was extracted to avoid double counting. To isolate spatial variation in abundance, we multiplied variance estimates by (1 - the proportion of between-route variance due to sampling error). Estimates of the proportion of BBS between-route variance due to sampling error (α) were taken from Link et al. (1994), who estimated this quantity using repeated counts at BBS routes within seasons. For species not evaluated by Link et al. (1994), α was estimated based on their finding that α is negatively related to abundance according to a regression of $\text{logit}(\alpha)$ on the natural logarithm of abundance. Sampling error could not be extracted for the community metrics that we examined because Link et al. (1994) did not evaluate sampling error at the community level. After extracting sampling error, we calculated the coefficient of variation in response across routes. Estimates were made using data collected from 1992 to 1998 (11–16 routes per year), and the average of these coefficients of variation was used to estimate τ_2^2 . To avoid undefined

coefficients of variation, we excluded years with zero mean response across routes from the analysis.

We estimated τ_3^2 as the variation in trend across BBS routes in the closed boreal forest region of Alberta between 1989 and 1998. Although counts from 1989 to 1991 were not used to estimate τ_2^2 due to reduced numbers of routes sampled in this time frame, they were included here to achieve longer trajectories and to enhance τ_3^2 estimates. No counts prior to 1989 were used, because of the infrequency of route visits in the region. For species, only routes satisfying the U.S. National Biological Service data selection criteria (Thomas 1997) were used in the analysis. After applying these criteria, data from four to 15 routes were available, depending on the species and community metric. The trend at each route was estimated by fitting the data to an exponential trend model, using software written by Thomas (1997). After estimating the variance of route trend, a coefficient of variation was calculated by dividing the square root of the variance estimate by the mean route trend.

When calculating species richness and Shannon-Weiner community metrics from the Calling Lake and BBS data sets, we used all species in the data sets. When calculating species richness of the ground-nesting guild, we included the following species: Black and White Warbler, *Mniotilta varia*; Blackpoll Warbler, *Dendroica striata*; Canada Warbler, *Wilsonia canadensis*; Clay-colored Sparrow, *Spizella pallida*; Connecticut Warbler, *Oporornis agilis*; Dark-eyed Junco, *Junca hyemalis*; Hermit Thrush, *Catharus guttatus*; LeConte's Sparrow, *Ammodramus leconteii*; Lincoln's Sparrow, *Melospiza lincolnii*; Mourning Warbler, *Oporornis philadelphia*; Northern Waterthrush, *Seiurus noveboracensis*; Orange-crowned Warbler, *Vermivora celata*; Ovenbird, *Seiurus aurocapillus*; Palm Warbler, *Dendroica palmarum*; Ruffed Grouse, *Bonasa umbellus*; Song Sparrow, *Melospiza melodia*; Swamp Sparrow, *Melospiza georgiana*; Tennessee Warbler, *Vermivora peregrina*; Wilson's Warbler, *Wilsonia pusilla*; White-throated Sparrow, *Zonotrichia albicollis*; and Yellow-bellied Flycatcher, *Empidonax flaviventris*.

Mean and variance parameter estimates

Table 2, Table 3, and Table 4 present variance parameter and mean abundance estimates for selected species and community metrics. To summarize the effect of within-site sample effort on τ_1^2 , the response of τ_1^2 to number of surveys, number of stations, and timing of surveys was calculated for each species and community metric (Table 5).

Assumptions and data limitations

Our results are contingent on assumptions made during variance parameter estimation. Within-site temporal variance was estimated by calculating variance in abundance across years at three control sites. Control sites were assumed devoid of population trends, which would confound variance. This assumption appears to be valid, because a high level of sample effort failed to detect significant trends over the 6-year period for the eight species selected. As an alternative, Gibbs et al. (1998) estimated temporal variance from the residuals of a linear regression of counts against time, in order to remove trend from the counts and minimize the effect of non-natural variation. This method is only valid if a significant trend exists. Otherwise, it can underestimate temporal variance by allocating natural variation to a nonexistent trend. This may explain why the estimate of Gibbs et al. (1998) of temporal variation for small birds (coefficient of variation = 0.57) was lower than ours (mean coefficient of variation across species = 1.083). Underestimating temporal variance results in overly optimistic predictions of the effectiveness of proposed monitoring strategies. We evaluated how τ_1^2 was affected by within-site sample effort. Two other factors that will also influence τ_1^2 are consistency in sampling methodology and the person collecting data at the site. Both the Calling Lake study (Schmiegelow et al. 1997) and the BBS (Droege 1990) use consistent sampling methodology that includes specification of weather conditions and the time of day and year during which data can be collected. Similar strict requirements should be a component of all monitoring programs because they reduce sample error by reducing variation in the proportion of present birds that are singing. The ability to detect birds that are singing varies across people. Although some of this variation can be reduced through training, consistency in data collectors should be maintained as much as possible. Such consistency will be difficult to maintain in long-term monitoring programs, perhaps causing an increase in τ_1^2 .

Because sample error was accounted for in τ_1^2 , it was extracted from τ_2^2 to avoid double counting, and therefore exaggeration, of sample error. Sample error was not extracted from τ_3^2 , however, because no estimates were available to do so. Although this may have caused exaggeration of τ_3^2 , the influence on power estimates is thought to be minimal due to its insensitivity to τ_3^2 (see Table 6).

Data limitations required us to use BBS data to estimate between-site variance parameters. The BBS design differs from the sampling design assumed here, which may reduce the accuracy of related estimates due to differences in sample error associated with the different methodologies. However, results should be insensitive to this difference because, of the two between-site variance parameters, one had little influence on power (τ_3^2) and sample error was extracted from the other (τ_2^2). Nevertheless, the most effective way to improve the data would be to implement a long-term monitoring program with a design based on estimates presented here. Over time, data from the monitoring program could be used to re-estimate variance parameters and reassess program design.

One final caveat is that our power simulations assumed that the coefficient of variation remained stable across abundance levels within a species. Power will be exaggerated if coefficients of variation increase, or will be underestimated if they decrease, as abundance declines. However, sampling efficiencies will remain valid, unless spatial and temporal coefficients of variation are differentially affected.

APPENDIX 2. Power Simulation Methodology

Monte Carlo simulations were performed using numerical recipes from Press et al. (1992) and route regression code by Thomas (1997). We simulated 20-year trajectories for each site following an exponential model, which required specification of site-specific initial abundance and trend magnitude. Exponential trends were based on the assumption that populations experience constant rates of decline (Caughley and Sinclair 1994). Because achieving consistent timing of surveys from year to year is logistically difficult, randomly timed surveys were assumed. Initial abundance was determined by randomly selecting a deviate from $N(\theta_0, \tau_2^2)$, truncated at zero. Similarly, trend magnitude was determined by randomly selecting a deviate from $N(-3\%, \tau_3^2)$ when generating population trajectories, and $N(-1\%, \tau_3^2)$ when generating community metric trajectories. Population trends of -3% per year were simulated because this is the effect size goal of the Alberta Forest Biodiversity Monitoring Program (Farr et al. 1999). Over 20 years, this effect size equates to a 45% decline. Simulated community trends were reduced to -1% per year to account for the reduced sensitivity of community metrics to disturbance. Over 20 years, this effect size equates to an 18% decline. Negative trends were simulated to achieve conservative power estimates and because they are of greater conservation interest. For a given effect size, positive exponential trends result in larger overall changes and are therefore easier to detect.

Monitoring data collected from a site within a given year were simulated as gaussian deviates (truncated at zero) with mean equal to the point on the site-specific trajectory for that year, and variance τ_1^2 . Although a discrete distribution, such as the Poisson, is often used to generate count data, a continuous distribution was required to accommodate counts from multiple surveys that were averaged prior to analysis. Although the lognormal and zero-truncated normal distributions were considered when generating abundances, the zero-truncated normal was selected because the lognormal underrepresented zero counts. The zero-truncated normal was also used to generate community metrics because they are calculated by summing random variables and, according to the central limit theorem (Hilborn and Mangel 1997), should therefore be distributed normally.

To analyze the simulated data, we applied route regression (Geissler and Sauer 1990). Route regression is a two-stage estimator, whereby trends are first estimated at each site by fitting the data to the exponential model and are then combined to estimate regional trend. Trend significance was estimated following the method used by the U.S. National Biological Service when analyzing BBS data, as described in Thomas (1997). This involved taking 400 bootstrap samples of the site trends and using the bootstrap mean and variance in a two-sided z test to test the alternative hypothesis that regional trend differed from zero. Significance was set at 80%, following the

recommendation of Gibbs et al. (1998) that power and significance be set at 90% and 80%, respectively, for ecological monitoring programs. The simulation process was repeated 1000 times (Link and Hatfield 1990) and power was estimated as the proportion of simulations for which trend was determined significant and trend direction was correctly identified.

APPENDIX 3. Monitoring Program Cost Model

The cost model accounts for labor, equipment, and terrestrial and aerial access costs. Cost estimates were provided by the Alberta Forest Biodiversity Monitoring Program (D. Farr, *personal communication*). Sampling of sites is completed by two-person field crews. Within our simulated range, each crew samples one site per day, regardless of point-count station density. Based on an annual crew cost of \$26,667, and assuming 30 site visits per crew in a year, the cost per site sampling visit is \$889. Sites are also visited once prior to initiation of the monitoring program, to locate sites and identify point-count stations, at a cost of \$1,333 per site. Because of the remoteness of much of Alberta's forest, it is assumed that 50% of the sites require aerial (helicopter) access. Each site requires two hours of helicopter time, at a cost of \$750 per hour. Total program cost is then:

$$\text{cost} = s [fr (\$1,639) + \$2,083] \quad [\text{A3.1}]$$

The equation underestimates true monitoring cost because it accounts only for field and access costs, and not for costs associated with administration, data analysis and management, training, and report production. However, because the magnitude of these excluded costs is relatively independent of sample effort, the model should provide a reasonable portrayal of the relative costs of the sample effort parameters.

LITERATURE CITED

Bernstein, B. B., and J. Zalinski. 1983. An optimum sampling design and power tests for environmental biologists. *Journal of Environmental Management* 16:35–43.

Bibby, C. J., N. D. Burgess, and D. A. Hill. 1992. *Bird census techniques*. Academic Press, San Diego, California, USA.

Burbidge, A. A. 1991. Cost constraints on surveys for nature conservation. Pages 3–6 in C. R. Margules and M. P. Austin, editors. *Nature conservation: cost effective biological surveys and data analysis*. CSIRO, Melbourne, Australia.

Caughley, G., and A. R. E. Sinclair. 1994. *Wildlife ecology and management*. Blackwell Science, Cambridge, Massachusetts, USA.

Committee on Environment and Natural Resources. 1996. *Integrating environmental monitoring and research in the mid-Atlantic Region*. Proceedings of a Workshop, College park, maryland, 10–12 April 1996. Committee on Environment and Natural Resources, National Science and Technology Council, Washington, D.C., USA. [Available

online <http://www.epa.gov/cludygxb/Pubs/marrpt.pdf>.]

Committee on Environment and Natural Resources. 1997. *Integrating the nation's environmental monitoring and research networks and programs: a proposed framework*. Environmental Monitoring Team, Committee on Environment and Natural Resources, National Science and Technology Council, Washington, D.C., USA. [Available online <http://www.epa.gov/cludygxb/Pubs/framewrk.pdf>.]

Croonquist, M. J., and R. P. Brooks. 1991. Use of avian and mammalian guilds as indicators of cumulative impacts in riparian-wetland areas. *Environmental Management* 15:701–714.

Dixon, M., A. R. Olsen, and B. M. Kahn. 1998. Measuring trends in ecological resources. *Ecological Applications* 8(2):225–227.

Droege, S. 1990. The North American Breeding Bird Survey. Pages 1–4 in J. R. Sauer and S. Droege, editors. *Survey designs and statistical methods for the estimation of avian population trends*. Biological Report 90(1). U.S. Fish and Wildlife Service, Washington, D.C., USA.

Fairweather, P. G. 1991. Statistical power and design requirements for environmental monitoring. *Australian*

Journal of Marine and Freshwater Research 42:555–567.

Farr, D. R., S. E. Franklin, E. E. Dixon, G. Scrimgeour, S. Kendall, P. Lee, S. Hanus, N. N. Winchester, and C. C. Shank. 1999. *Monitoring forest biodiversity in Alberta: program framework*. Alberta Forest Biodiversity Monitoring Program Technical Report 3. Draft Report. [Available online <http://www.fmf.ab.ca/bm/reports.htm>.]

Geissler, P. H., and J. R. Sauer. 1990. Topics in route-regression analysis. Pages 54–57 in J. R. Sauer and S. Droege, editors. *Survey designs and statistical methods for the estimation of avian population trends*. Biological Report 90(1). U.S. Fish and Wildlife Service, Washington, D.C., USA.

Gibbs, J. P., S. Droege, and P. Eagle. 1998. Monitoring populations of plants and animals. *BioScience* 48:935–940.

Gibbs, J. P., and S. M. Melvin. 1997. Power to detect trends in waterbird abundance with call–response surveys. *Journal of Wildlife Management* 61(4):1262–1267.

Halbert, C. L. 1993. How adaptive is adaptive management? Implementing adaptive management in Washington State and British Columbia. *Reviews in Fisheries Science* 1:261–283.

Hilborn, R., and M. Mangel. 1997. *The ecological detective*. Princeton University Press, Princeton, New Jersey, USA.

Hinds, W. T. 1984. Towards monitoring of long-term trends in terrestrial ecosystems. *Environmental Conservation* 11:11–17.

Johnson, S. P. 1993. *The Earth Summit: the United Nations Conference on Environment and Development*. Graham and Trotman/Martinus Nijhoff, London, UK.

Krinsky, I., and A. L. Robb. 1986. On approximating the statistical properties of elasticities. *Review of Economics and Statistics* 68:715–719.

Lewis, S. A., and W. R. Gould. 2000. Survey effort effects on power to detect trends in raptor migration counts. *Wildlife Society Bulletin* 28:317–329.

Link, W. A., R. J. Barker, J. R. Sauer, and S. Droege. 1994. Within-site variability in surveys of wildlife populations. *Ecology* 75(4):1097–1108.

Link, W. A., and J. S. Hatfield. 1990. Power calculations and model selection for trend analysis: a comment. *Ecology* 71:1217–1220.

Margalef, R. 1958. Information theory in ecology. *General Systems* 3:36–71.

MathSoft. 1999. *S-Plus 2000*. Data Analysis Products Division, MathSoft, Seattle, Washington, D.C., USA.

Millard, S. P., and D. P. Lettenmaier. 1986. Optimal design of biological sampling programs using the analysis of variance. *Estuarine, Coastal and Shelf Science* 22:637–656.

Minister of Supply and Services Canada. 1995. *Canadian biodiversity strategy: Canada's response to the Convention on Biological Diversity*. Catalogue Number En21-134/1995E.

Myers, R. H. 1990. *Classical and modern regression with applications*. Duxbury Press, Belmont, California, USA.

National Research Council. 1995. *Review of EPA's Environmental Monitoring and Assessment Program: overall evaluation*. National Academy Press, Washington, D.C., USA.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. *Canadian Journal of Fisheries and Aquatic Sciences* 47:2–15.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical recipes in C: the art of scientific computing*. Cambridge University Press, New York, New York, USA.

Schmiegelow, F. K., C. S. Machtans, and S. J. Hannon. 1997. Are boreal birds resilient to forest fragmentation? An experimental study of short-term community responses. *Ecology* 78:1914–1932.

Schneider, R., P. Lee, C. Shank, D. Farr, and B. Stelfox. 1999. Conceptual framework and rationale for monitoring forest biodiversity in Alberta. Pages 1–45 in D. R. Farr, S. E. Franklin, E. E. Dixon, G. Scrimgeour, S. Kendall, P. Lee, S. Hanus, N. N. Winchester, and C. C. Shank. *Monitoring forest biodiversity in Alberta: program framework*. Alberta Forest Biodiversity Monitoring Program Technical Report (Draft) 3. [Available online <http://www.fmf.ab.ca/bm/reports.htm>.]

Scott, C. T. 1998. Sampling methods for estimating change in forest resources. *Ecological Applications* 8(2):228–233.

Silvey, S. D. 1980. *Optimal design: an introduction to the theory for parameter estimation*. Chapman and Hall, London, UK.

Spellerberg, I. F. 1991. *Monitoring ecological change*. Cambridge University Press, Cambridge, UK.

Steele, B. B., R. L. Bayn, Jr., and C. Val Grant. 1984. Environmental monitoring using populations of birds and small mammals: analyses of sampling effort. *Biological Conservation* 30:157–172.

Stork, N. E., and M. J. Samways. 1995. Inventorying and monitoring: executive summary. Pages 457–458 in V. H. Heywood and R. T. Watson, editors. *Global biodiversity assessment*. Cambridge University Press, Cambridge, UK.

Thomas, L. 1997. *Evaluation of statistical methods for estimating long-term population change from extensive Wildlife Surveys*. Dissertation. University of British Columbia, Vancouver, British Columbia, Canada.

Thompson, W. L., G. C. White, and C. Gowan. 1998. *Monitoring vertebrate populations*. Academic Press, San Diego, California, USA.

U.S. Environmental Protection Agency. 2000. EMAP. [Available online <http://www.epa.gov/emap/> "Environmental Monitoring and Assessment Program Homepage."]

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

USGS Patuxent Wildlife Research Center. 2000. *North American Breeding Bird Survey Internet data set*. [Available online <http://www.mp2-pwrc.usgs.gov/bbs/retrieval/disclaim.cfm>.]

Verner, J. 1988. Optimizing the duration of point counts for monitoring trends in bird populations. *Research Note PSW-395*. Pacific Southwest Forest and Range Experiment Station, U. S. Department of Agriculture, Forest Service, Berkeley, California, USA.

Versyck, K. J., K. Bernaerts, A. H. Geeraerd, and J. F. van Impe. 1999. Introducing optimal experimental design in predictive modeling: a motivating example. *International Journal of Food Microbiology* 51:39–51.

Walters, C. J., and C. S. Holling. 1990. Large-scale management experiments and learning by doing. *Ecology* 71:2060–2068.