Received: 8 October 2021 | Accepted: 22 December 2021

DOI: 10.1002/eap.2621

ARTICLE

ECOLOGICAL APPLICATIONS
ECOLOGICAL SOCIETY OF AMERICA

A comparison of monitoring designs to assess wildlife community parameters across spatial scales

Alexander D. Wright[1,2] | **Evan H. Campbell Grant**[3] | **Elise F. Zipkin**[1,2]

[1]Department of Integrative Biology, Michigan State University, East Lansing, Michigan, USA

[2]Ecology, Evolution, and Behavior Program, Michigan State University, East Lansing, Michigan, USA

[3]SO Conte Anadromous Fish Research Laboratory, USGS Patuxent Wildlife Research Center, Turners Falls, Massachusetts, USA

Correspondence
Alexander D. Wright
Email: adwright@msu.edu

Funding information
National Park Service

Handling Editor: Eric J. Ward

Abstract

Dedicated long-term monitoring at appropriate spatial and temporal scales is necessary to understand biodiversity losses and develop effective conservation plans. Wildlife monitoring is often achieved by obtaining data at a combination of spatial scales, ranging from local to broad, to understand the status, trends, and drivers of individual species or whole communities and their dynamics. However, limited resources for monitoring necessitates tradeoffs in the scope and scale of data collection. Careful consideration of the spatial and temporal allocation of finite sampling effort is crucial for monitoring programs that span multiple spatial scales. Here we evaluate the ability of five monitoring designs—stratified random, weighted effort, indicator unit, rotating panel, and split panel—to recover parameter values that describe the status (occupancy), trends (change in occupancy), and drivers (spatially varying covariate and an autologistic term) of wildlife communities at two spatial scales. Using an amphibian monitoring program that spans a network of US national parks as a motivating example, we conducted a simulation study for a regional community occupancy sampling program to compare the monitoring designs across varying levels of sampling effort (ranging from 10% to 50%). We found that the stratified random design outperformed the other designs for most parameters of interest at both scales and was thus generally preferable in balancing the estimation of status, trends, and drivers across scales. However, we found that other designs had improved performance in specific situations. For example, the rotating panel design performed best at estimating spatial drivers at a regional level. Thus, our results highlight the nuanced scenarios in which various design strategies may be preferred and offer guidance as to how managers can balance common tradeoffs in large-scale and long-term monitoring programs in terms of the specific knowledge gained. Monitoring designs that improve accuracy in parameter estimates are needed to guide conservation policy and management decisions in the face of broad-scale environmental challenges, but the preferred design is sensitive to the specific objectives of a monitoring program.

KEYWORDS
Bayesian hierarchical model, biodiversity, large-scale monitoring, macrosystems ecology, occupancy, sampling design

INTRODUCTION

Monitoring programs are essential for natural resource management because they provide data to address scientific questions, develop predictive models, trigger and guide management actions, and assess the impacts of policies and interventions in support of evidence-based conservation (Nichols & Williams, 2006; Sutherland et al., 2004; Yoccoz et al., 2001). The need for ecological monitoring has increased over the last several decades as global pressures have grown in severity and biodiversity loss has accelerated (Butchart et al., 2010; Nicol et al., 2019). However, determining the efficient allocation of limited resources is a critical impediment to the development of effective monitoring programs (Buxton et al., 2020; Lindenmayer & Likens, 2010). This is particularly important when considering heterogeneous landscapes over large spatial scales and complex governance networks with multiple management jurisdictions (Bennett et al., 2018; Carlson & Schmiegelow, 2002).

Recognition that biodiversity change stems from interacting local, regional, and global drivers (Keller et al., 2008; Lindenmayer & Likens, 2010) has spurred the design of monitoring programs to match these scales (e.g., NSF NEON, Thorpe et al., 2016; TEAM, Beaudrot et al., 2016; USGS BBS, Sauer et al., 2017). Large-scale monitoring programs are used to track the status, trends, and drivers of wildlife species and communities across individual or independent spatial units. Such programs are often organized as part of a regional or national administrative network to detect and understand changes in biodiversity (Yoccoz et al., 2001). In contrast, conservation management decisions are often implemented at the local level (e.g., individual parks, refuges). As such, large-scale monitoring programs must be able to both detect biodiversity changes on broad scales while also providing specific information on local scales to inform management activities that can help mitigate declines when and where they arise (Adams & Muths, 2019).

Designing and implementing robust monitoring programs to meet these multiple priorities, often with limited resources, remains a challenge (Blanchet et al., 2020; Jones, 2011; Lindenmayer & Likens, 2010; Scholes et al., 2008). Monitoring programs typically fall into one of three categories: landscape, surveillance, and targeted (Eyre et al., 2011). These approaches range in spatial extent, information content, and purpose (Sparrow et al., 2020). Landscape monitoring primarily aims to measure population status (e.g., species distribution or abundance) through descriptive and spatially continuous information collected across broad spatial scales. Surveillance monitoring (sometimes referred to as omnibus monitoring) aims to detect and observe population trends through repeated, standardized surveys that can be conducted at local to broad spatial scales. Targeted monitoring aims to evaluate and understand the drivers of population dynamics through a hypothesis-driven approach that is often executed at small and discrete scales (Eyre et al., 2011). Yet, monitoring programs that accurately (i.e., precisely and in an unbiased manner) estimate status, trends, and drivers simultaneously across multiple scales are increasingly necessary for understanding, and reacting to, rapidly changing environmental conditions (Albert et al., 2010; Scholes et al., 2008; Sparrow et al., 2020).

Here, we evaluate the degree to which different monitoring designs allow one to make inferences about wildlife species and community status, trends, and drivers within and across multiple management units and spatial scales. To do this, we conducted a simulation study comparing the effectiveness of common designs that combine various elements of targeted, surveillance, and landscape monitoring. We used a regional amphibian monitoring program within a network of mid-Atlantic national parks as a case study (Grant & Brand, 2012; National Park Service, 2005). Like many natural resource agencies charged with large-scale monitoring initiatives, the National Capital Region Inventory and Monitoring Network (NCRN) of the US National Park Service (National Park Service, 2005) wishes to maximize the information gained from annual amphibian occupancy surveys within budget constraints. The use of a relevant case study for our analyses ensures the logistical feasibility of each strategy and the real-world applicability of our simulation results. However, our approach is general in scope, so our results should be broadly informative to researchers and managers developing sampling schemes across taxa, scales, and landscape configurations.

We reviewed existing large-scale monitoring programs to choose sampling designs that allow the simultaneous estimation of species status (e.g., occupancy), trends, and drivers of species and community changes and considered these in a hierarchical framework to develop inferences across multiple spatial scales. Our comparisons focused on five commonly used monitoring designs: stratified random, indicator, rotating panel, split panel, and weighted effort (each described in more detail in the section *Methods*). We simulated and assessed the effectiveness of the different monitoring designs (which differ in their spatial distribution of sampling sites across a network of independent units) across various levels of sampling effort (i.e., 10%, 20%, 30%, 40%, or 50% of available sites sampled across the total potential habitat). We evaluated the accuracy of multiregion community occupancy parameter estimates, including metrics of population status (mean occupancy), trend (a year-specific effect), and drivers (a site-specific effect as well as a temporal autologistic effect), across two nested spatial scales. Our results quantify the tradeoffs of common designs for large- and multiscale monitoring programs within the real-world context of allocation decisions regularly faced by management agencies.

METHODS

Large-scale monitoring designs

Five monitoring designs—stratified random, indicator, rotating panel, split panel, and weighted effort—were selected for the simulation study because they are representative of existing large-scale monitoring programs that balance some combination of targeted, surveillance, and landscape monitoring (Eyre et al., 2011). For all designs, we defined site as the sampling location within a unit, the local spatial area (e.g., park, reserve), and region as the overall geographical extent encompassing all of the local units. The monitoring designs vary in their allocation of effort across units to evaluate status, trends, and drivers of population and community change within a defined region. In what follows, we describe each design, including the pros and cons of the various approaches.

The stratified random design makes use of an approach in which a fixed percentage of sites randomly selected from each unit (weighted by the number of sites available at each unit) are surveyed periodically, usually annually (Thompson, 2012). For example, the North American Breeding Bird Survey uses a uniform number of randomly selected sites (i.e., routes) for each one degree of latitude and longitude block in every US state (and parts of Canada and Mexico), which are targeted for annual sampling (Sauer et al., 2013; Ziolkowski et al., 2010). Stratified random sampling designs are intended to provide information on each unit, including the status, trends, and drivers of monitored populations or communities, although the precision of inferences depends on the number of sites monitored in each unit. Sampling areas can be stratified by abiotic (e.g., elevation, habitat type), biotic (e.g., species richness, population density), or geographical factors (e.g., jurisdictional boundaries); here we stratify by unit because it is the desired level of inference across the study area.

The weighted effort design is an approach in which all relevant units are available for sampling; however, sampling effort is unevenly distributed across units each year (i.e., intense monitoring of sites at a subset of units, limited monitoring of sites at the remaining units). For example, amphibian monitoring of national parks in the National Capital Region Inventory and Monitoring Network (National Park Service, 2005) conducts a disproportionate level of replicate visits at select parks that have a long history of or higher need for monitoring (Wright et al., 2020). This design incorporates elements of the stratified random design described previously but also distributes effort to provide a more granular perspective at a subset of units. Thus, its ability to estimate parameters for units is not equitable in a region; high accuracy of parameter estimates is achieved in some units at the expense of others in the region.

The indicator design is an approach in which a subset of units, which are selected to represent political or biological domains, are surveyed intensively, whereas remaining available units are not sampled at all. The National Science Foundation's Long-Term Ecological Research (LTER) Network is an example of this approach, in which 28 representative units (of specific landscape types) across the United States are monitored intensely through time (Callahan, 1984). This approach ensures robust temporal coverage within each unit but limited spatial replication. The indicator unit design assumes that relevant parameter estimates of the indicator units are indicative of similar unmonitored units, that the relationship between monitored and unmonitored units is known and constant over time, or that parameter estimates at unmonitored units are not of central interest. In our simulations, we select units randomly and not to represent political or biological domains, and thus our design may show better performance (at the regional level) than a nonrandomly selected indicator design (which is more common in large-scale monitoring designs).

The rotating panel design is a spatiotemporally varying design in which units are surveyed at specific, rotating intervals (Dobbie et al., 2008; McDonald, 2003). For example, within the Alberta Biodiversity Monitoring Program, every available site across the province is sampled once in each 5-year period on a rotating basis (Stadt et al., 2006), ensuring extensive spatial coverage at the expense of limited temporal replication. The rotating panel design allocates some degree of monitoring effort in all units across the temporal extent of the monitoring program, ensuring high spatial coverage and representation across the region. However, temporal coverage is minimal because repeated visits to individual units occur infrequently.

The split panel design is a spatiotemporally varying design in which a set of core units is consistently monitored over time while the remaining sets of units are monitored on a variable or rotating basis (Dobbie et al., 2008; McDonald, 2003). The National Ecological Observatory Network (NEON) uses this approach with a set of established study areas, which are fixed and sampled every year, and relocatable units, which can be moved every 5–10 years (Keller et al., 2008; Kao et al., 2012; Thorpe et al., 2016). This design integrates elements of both the indicator and rotating panel strategies, attempting to alleviate the limited spatial coverage of the indicator design and the limited temporal coverage of the rotating panel design. As such, it emphasizes intensive monitoring at a select number of units across time while also attempting to achieve broader spatial coverage.

Data simulation

To assess the effectiveness of each of the five monitoring designs in estimating species and community status, trend, and driver metrics within and across scales, we simulated 500 data sets for each monitoring design at five sampling effort levels, defined as the percentage of sites in the region sampled (10%, 20%, 30%, 40%, 50%) for a total combination of 25 simulation scenarios and 12,500 unique data sets (Wright, 2021). For each data set, we simulated 10 years of multispecies occupancy data across 10 hypothetical spatial units. Community size varied across units and simulations and was determined by a random draw of the maximum possible number of species in the metacommunity (50 species) multiplied by the likelihood that each potential species was part of each local community (also known as the data augmentation parameter) (Royle et al., 2007), which had a mean of 0.4 (on the probability scale 0–1) and a standard deviation of 0.25. The number of available sites in each spatial unit was randomly drawn from a uniform distribution with a minimum bound of 10 sites and a maximum bound of 100 sites (16, 21, 23, 35, 40, 47, 66, 72, 90, 98). Thus, the actual number of sites sampled (sample size) at the regional level was ~50 sites (ranging from 2 to 10 sites per unit) at the 10% effort level, ~100 (3–20 per unit) at 20% effort, ~150 (5–29 per unit) at 30% effort, ~200 (6–39 per unit) at 40% effort, and ~250 (8–49 per unit) at the 50% effort level. We chose 10 units (and the corresponding number of sites at each unit) to closely resemble the network of monitoring units in our case study (NCRN) (National Park Service, 2005). Administrative evaluation is typically on 5-year cycles in US federal programs, so 10 years is a reasonable timeframe for both assessment (sufficient time series for estimation) and enactment of management activities (Government Performance and Results Act of 1993).

Data sets were simulated using occupancy and detection parameter estimates from an analysis of the long-term, regional amphibian monitoring data collected in the NCRN (Wright et al., 2020), which is characterized by a regional amphibian community with low mean detection ($p = 0.3$) and low mean occupancy ($\psi = 0.3$). We assumed moderate heterogeneity (SD $= 0.5$) across unit-level parameter means (i.e., metacommunity) and moderate heterogeneity (SD $= 0.5$) across all species-level parameter means (i.e., community). By incorporating a reasonable range of variability in the generation of simulation parameters across runs, there is considerable heterogeneity in the simulation of unit- and species-level parameters, resulting in a broader parameter space. We categorized covariate effects as moderate $|0.4|$ or strong $|0.8|$. We used the same model to both generate and analyze the latent occurrence state for species and the community (described in what follows in full detail in the "*Analysis*" section).

Following simulation of the latent occurrence state for all species in all units and sites on 12,500 occasions, we then simulated sampling according to one of the unique monitoring design and sampling effort combinations (e.g., stratified random design with 10% of sites across the region sampled) to obtain corresponding data sets. We assumed that each site selected for sampling was surveyed on four replicate visits per year (unless otherwise indicated), which is sufficient for detecting declines in occupancy (Field et al., 2005) and consistent with the current protocol in the NCRN. Within each sampling effort level, the total number of sites visited annually within the region across the monitoring designs was consistent, which ensured that our results were comparable. Thus, designs only differed in which sites, across spatial units, were sampled, not in how many total sites were sampled at the regional level.

To implement sampling in the stratified random design, the same proportion of randomly selected sites at each unit, relative to each effort level, was sampled continuously for all 10 years (Appendix S1: Table S1). For the weighted effort design, all units were sampled similarly to the stratified random design; however, the number of replicate visits per site in each year varied among units (six replicate visits for sites in half of the units, and two for the remaining half of units) (Appendix S1: Table S1). For the indicator design, the same random sample of sites at a subset of units (containing half of all available sites across the region) was sampled every year (Appendix S1: Table S2). For the rotating panel design, two sets of units containing an equal number of randomly selected sites were sampled on alternate 2-year rotations (Appendix S1: Table S3). For the split panel design, the same random sample of sites at a subset of units (containing half of all available sites across the region) was sampled every year, while the remaining units were split into two equal sets that were surveyed on alternate 2-year periods (Appendix S1: Table S4).

Analysis

Multispecies (community) occupancy models are often used in the analysis of biodiversity monitoring data to estimate richness as well as species and community dynamics (Dorazio et al., 2006; Dorazio & Royle, 2005). These models utilize replicate observations to incorporate detection probability (p) in the estimation of the true latent state of species occurrence (present or absent) at a sampling site (MacKenzie et al., 2002). By incorporating detection and assuming a shared link across species within a community, multispecies occupancy models can accommodate data from rare, cryptic, and unobserved species to produce accurate estimates of individual species occupancy probabilities (ψ) and species richness (Boulinier et al., 1998; Zipkin et al., 2010). The recent development of multiregion community occupancy models incorporates both multiple species and multiple independent spatial units through a unified statistical analysis (Sutherland et al., 2016), allowing for the investigation of community occupancy dynamics across spatial scales (e.g., ranging from local to regional levels).

We fit a multiregion community occupancy model (Sutherland et al., 2016; Wright, 2021; Wright et al., 2020) to each simulated data set to evaluate how the estimated parameters compared to the true parameter values for each of the 25 allocation strategies (five designs at five effort levels) using the same biological process model that was used to generate the data. We summarized the data into an array, $X_{i,r,j,t,k}$, with the detection ($X_{i,r,j,t,k}=1$) and nondetection ($X_{i,r,j,t,k}=0$) histories for each species i within unit r at site j during year t on replicate k. We assumed the detection of a species was conditional on the presence of species i within unit r at site j during year t ($Z_{i,r,j,t}=1$ if the species was there and a structural 0 otherwise) and the probability of detecting species i within unit r at site j during year t on replicate k ($p_{i,r,j,t,k}$) according to a Bernoulli process:

$$X_{i,r,j,t,k} \sim \text{Bernoulli}\left(Z_{i,r,j,t} \times p_{i,r,j,t,k}\right).$$

We then modeled detection probability assuming that detection could change by species or unit, where $\text{logit}\left(p_{i,r,j,t,k}\right)=\beta_{i,r}$, in which $\beta_{i,r}$ is an intercept term indicating the detection probability for each species i in each unit r on the logit scale.

We similarly modeled species occupancy state, $Z_{i,r,j,t}$, with a Bernoulli random process:

$$Z_{i,r,j,t} \sim \text{Bernoulli}\left(\psi_{i,r,j,t}\right),$$

where $\psi_{i,r,j,t}$ is the occupancy probability of species i within unit r at site j of year t. We incorporated covariates on species occupancy probability using a logit link function:

$$\text{logit}\left(\psi_{i,r,j,t}\right)=\alpha0_{i,r}+\alpha1_{i,r}\times \text{Year}_t+\alpha2_{i,r} \times \text{Site Covariate}_{j,r}+\alpha3_{i,r}\times \text{Z}_{j,t-1,i,r}.$$

We included species- and unit-specific intercept terms for mean occupancy ($\alpha0_{i,r}$) and effects for year ($\alpha1_{i,r}$), a spatially varying covariate ($\alpha2_{i,r}$), and an autologistic process ($\alpha3_{i,r}$). The covariate that influences species occupancy probabilities (Site Covariate$_{j,r}$) varies by site and was randomly generated during the data simulation process (from a normal distribution with a mean of 0 and a SD of 1). The autologistic term incorporates the processes of colonization (when $Z_{j,t-1,i,r}=0$) and extinction (when $Z_{j,t-1,i,r}=1$) that drive occupancy patterns for many species, including amphibians (Dorazio et al., 2010; Zipkin et al., 2012).

We categorized status as mean occupancy ($\alpha0$, the spatial distribution of occupancy in a moment of time), trend as the effect of year ($\alpha1$, the increase or decrease of occupancy over time), and drivers as the effects of the spatially varying covariate ($\alpha2$) and the autologistic process ($\alpha3$, the underlying processes that can influence occupancy status and trend). To link the single-species occupancy models at a community level, we assumed that each parameter was drawn from a common unit-level normal distribution, e.g., $\alpha0_{i,r} \sim \text{Normal}\left(\mu_{\alpha0,r},\sigma^2_{\alpha0}\right)$, and each unit-level distribution was drawn from a common region-level normal distribution, $\mu_{\alpha0,r} \sim \text{Normal}\left(\overline{\mu}_{\alpha0},\overline{\sigma}^2_{\alpha0}\right)$, matching the data generation process. This allowed us to estimate and compare parameters at both the unit (e.g., for status: $\mu_{\alpha0,r}$) and region (e.g., for status: $\overline{\mu}_{\alpha0}$) levels.

We ranked the five monitoring designs in terms of their ability to accurately recover estimates of the status, trend, and driver parameters across the two spatial scales—for local units individually and the region collectively. To measure accuracy, we calculated the root-mean-square error (RMSE)—a metric that accounts for both precision and bias, with the lowest values indicating highest accuracy—for all parameters (using the values estimated by the model and the known values used to simulate the data) in each monitoring design and sampling effort level combination. We estimated the parameters in our models for each simulated data set using a Bayesian framework in R (R Core Team, 2016) with the program JAGS and corresponding jagsUI package (Kellner, 2015; Plummer, 2003; see Wright, 2021 for code). We set vague priors for each parameter: mean regional-level intercept parameters for occupancy ($\overline{\mu}_{\alpha0}$)

and detection ($\overline{\mu}_{b0}$) had normal prior distributions with a mean of 0 and a variance of 2.70 (Lunn et al., 2012) and variance parameters with gamma prior distributions with shape and scale parameters of 0.1. The mean regional-level slope parameters (e.g., $\overline{\mu}_{\alpha1}$) had normal prior distributions with a mean of 0 and a variance of 10 and prior distributions for the variance term similar to those of the intercept parameters. Convergence for each parameter was assessed using the Gelman and Rubin convergence diagnostic ($\widehat{R}$ statistic <1.1) (Gelman & Rubin, 1992; Gelman & Shirley, 2011).

RESULTS

Status

At the regional level, the stratified random design had the lowest RMSE (and, thus, the highest accuracy) in estimating mean community-level occupancy ($\overline{\mu}_{\alpha0}$) across all effort levels, followed closely by the split panel design (average RMSE was 2% higher compared to the stratified random design), rotating panel design (6%), and weighted-effort design (13%), with the indicator unit design performing much more poorly than the other four approaches (290%) (Table 1). For the stratified random design, RMSE decreased with increased effort (by as much as half when going from 10% to 50% effort); however, the gains were substantially larger when the effort was low (i.e., a change from 10% to 20% effort yielded more improvement than a change from 40% to 50% effort). Across sampling designs, the relative decrease in RMSE was tempered as effort increased, indicating a general decrease in returns of estimation accuracy for the higher effort levels. The differences in RMSE across monitoring designs were most pronounced when effort was low (Figure 1a), indicating that the differences in performance among monitoring designs diminished as effort increased.

At the unit level, the stratified random design again had the lowest mean RMSE (across all units) in estimating the mean occupancy across species within a unit ($\mu_{\alpha0,r}$) for the individual r units across all effort levels and the lowest variation of RMSE across all units in each effort level (Figure 1b and Table 1). However, while the mean and variance were low, the lower bound of unit-level RMSE values was highest in the stratified random design (RMSE = 0.126), compared with the weighted

TABLE 1 Comparison of root-mean-square error (RMSE) estimates for each monitoring design and parameter of interest

Monitoring design and scale	Parameter of interest: Status — Mean occupancy, α0	Trends — Year-specific effect, α1	Trends — Site-specific effect, α2	Drivers — Autologistic effect, α3
Stratified random				
Region	0.0607	0.0477	0.0593	0.0800
Unit	0.155 (0.126–0.197)	0.128 (0.110–0.159)	0.152 (0.112–0.231)	0.223 (0.171–0.301)
Weighted effort				
Region	0.0684	0.0491	0.0605	0.0929
Unit	0.175 (0.117–0.257)	0.135 (0.108–0.169)	0.159 (0.109–0.251)	0.248 (0.155–0.338)
Indicator unit				
Region	0.177	0.168	0.171	0.187
Unit	0.348 (0.110–0.568)	0.336 (0.102–0.560)	0.337 (0.0999–0.556)	0.377 (0.148–0.566)
Rotating panel				
Region	0.0643	0.0518	0.0547	0.106
Unit	0.162 (0.120–0.203)	0.136 (0.102–0.182)	0.143 (0.111–0.190)	0.268 (0.182–0.344)
Split panel				
Region	0.0619	0.0505	0.0550	0.0955
Unit	0.159 (0.120–0.206)	0.135 (0.106–0.186)	0.147 (0.110–0.212)	0.248 (0.175–0.336)

Notes: Root-mean-square error (RMSE) values are summarized across effort levels. Region-level estimates for RMSE were characterized for the regional mean parameter (e.g., for status: $\overline{\mu}_{\alpha0}$). Unit-level estimates were characterized for the unit mean parameters (e.g., for status: $\mu_{\alpha0,r}$) and include the average RMSE of all units and the lower and upper bounds (in parentheses) of the distribution of unit-level RMSE estimates for each parameter.

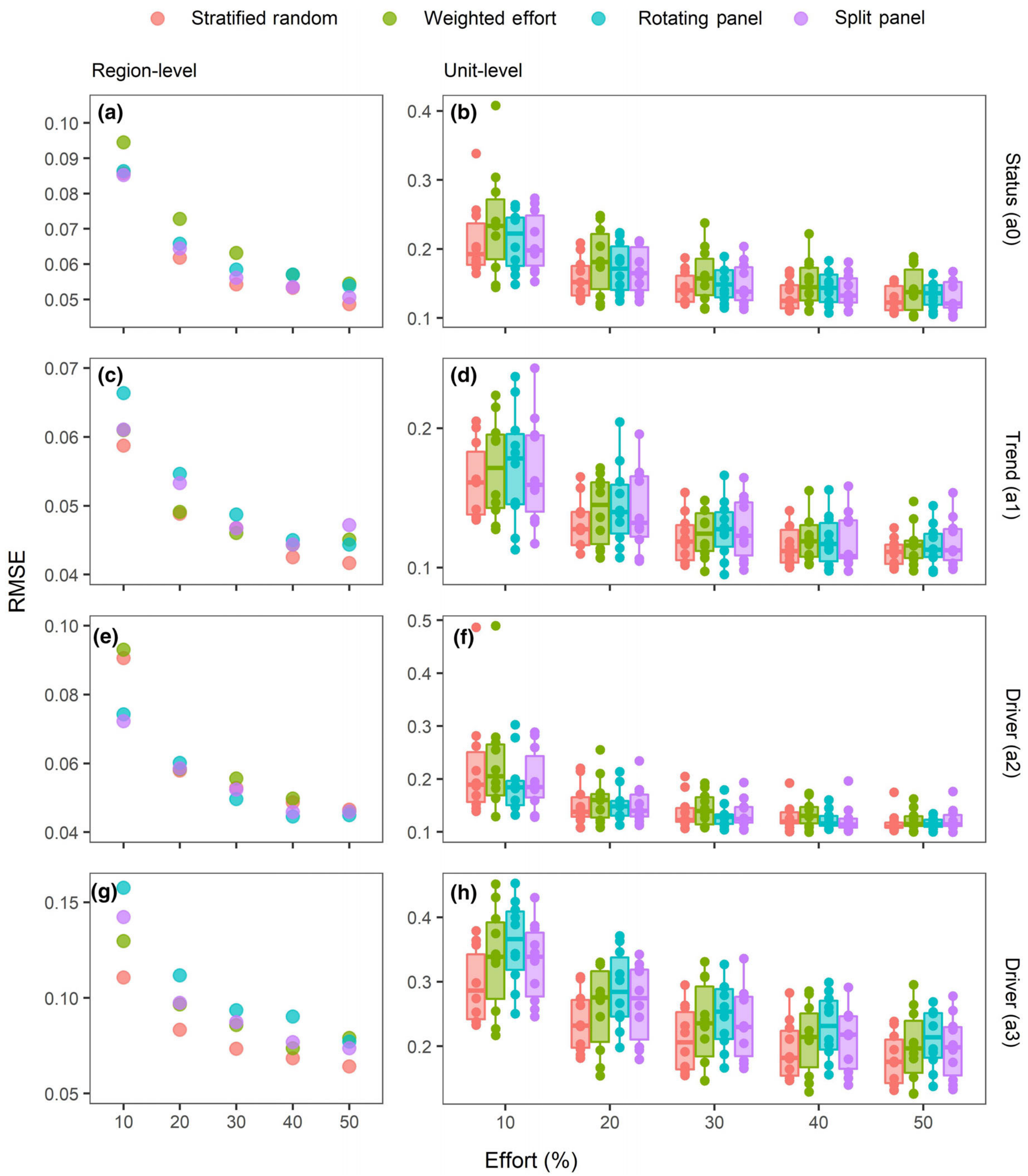

FIGURE 1 (a, c, e, g) Estimated root-mean-square error (RMSE) for each monitoring design and effort combination for each parameter at regional scale. (b, d, f, h) The unit-level RMSE estimates are organized by individual units (dots) and box plots describing the distribution of RMSE values across the units. The sample sizes for each effort level are 10%, ~50 sites at regional level (2–10 sites/unit); 20%, ~100 (3–20); 30%, ~150 (5–29), 40%, ~200 (6–39); and 50%, ~250 (8–49). The indicator unit design is not shown as its RMSE values were much higher than the others (see Table 1)

(0.117), rotating panel (0.120), split panel (0.120), and indicator unit designs (0.110). Thus, while stratified random design provided the most equitable estimates across all units in a region, other designs resulted in more accurate estimates of some individual units at the expense of parameter accuracy in other units (Table 1). The importance or significance of individual units may vary according to management or monitoring objectives, and equitability in parameter accuracy or across units may not be necessary in every monitoring scenario.

Trends

Unsurprisingly, monitoring design performance for estimating trends at the regional level were similar to those for estimating status. The stratified random design had the lowest RMSE in estimating a linear year effect on occupancy ($\overline{\mu}_{\alpha 2}$) across effort levels (Figure 1c), followed closely by the weighted effort design (average RMSE was 3% higher compared to the stratified random design), split panel design (6%), rotating panel design (9%), and, lastly, the indicator unit design (350%) (Table 1). For the stratified random design, RMSE decreased by 17% when effort was increased from 10% to 20%, 20% when effort was increased to 30%, 28% when effort was increased to 40%, and 29% when effort was increased to 50% (Figure 1c). Other designs showed a similar plateau of increased accuracy as effort increased.

At the unit level, the stratified random design again had the lowest mean RMSE in estimating the year effect parameter ($\mu_{\alpha 2,r}$) across all effort levels and the lowest variation of RMSE across all units in each effort level (Figure 1d and Table 1). However, though RMSE estimates per unit were more equitable for the stratified random design, the lower bounds of the unit-specific RMSE distributions for the other four designs were again less than that of the stratified random design. Of those other designs, the indicator unit and rotating panel designs had the lowest individual unit-specific RMSE estimates (0.102) (Table 1). This again reveals that while the stratified random design performs better for average trend estimates, the other designs are capable of estimating trends more accurately for a subset of units.

Drivers

Our results on drivers differed somewhat from the status and trend parameters. At the regional level, the rotating panel design had the lowest RMSE in estimating the effect of a site-specific covariate on occupancy ($\overline{\mu}_{\alpha 1}$) across effort levels (Figure 1e), which was comparable to estimates for the split panel design (average RMSE was <1% higher compared to the rotating panel design), followed by the stratified random design (8%), weighted effort design (11%), and indicator unit design (313%) (Table 1). Similarly, at the unit level, the rotating panel design had the lowest mean RMSE in estimating the spatially varying covariate parameter ($\mu_{\alpha 1,r}$) across all effort levels and the lowest variation of RMSE across all units in each effort level (Figure 1f and Table 1). However, again, the lowest individual unit RMSE estimate was from the indicator unit design.

In estimating the autologistic effect on occupancy ($\overline{\mu}_{\alpha 3}$) at the regional level, the stratified random design had the lowest RMSE across effort levels (Figure 1g), followed by the weighted effort design (average RMSE was 16% higher compared to the stratified random design), split panel design (19%), rotating panel design (33%), and, finally, the indicator unit design (234%) (Table 1). Likewise, at the unit level, the stratified random design had the lowest mean RMSE in estimating the autologistic slope parameter ($\mu_{\alpha 3,r}$) and the lowest variation of RMSE across all units in each effort level (Figure 1h and Table 1). However, both the indicator unit and weighted effort designs had lower individual unit RMSE estimates (0.148 and 0.155, respectively) relative to the stratified random design (0.171).

DISCUSSION

Our results suggest that stratified random sampling remains the most accurate monitoring approach for understanding wildlife occupancy at multiple spatial scales. With the exception of the spatially varying parameter on occupancy ($\alpha 2$), the stratified random design consistently had the lowest RMSE estimate across parameters at the regional level and the lowest mean and variation of RMSE estimates at the unit level. Thus, if decision makers have multiple, or dynamic, monitoring objectives across parameters of interest (i.e., status, trends, and drivers) or scales (i.e., locally and regionally), the stratified random design will likely be the preferred monitoring design. This is not unexpected; indeed, it is one of the reasons that stratified random sampling is so widely used. A stratified random design ensures that data come from a representative sample that accounts for spatial heterogeneity, leading to an efficient use of monitoring effort (Schreuder et al., 2004). Further, stratified random sampling avoids subjective decision-making and potential biases in site or unit representation in a monitoring program (Dobbie et al., 2008).

Despite the many advantages of stratified random sampling, other monitoring designs may be preferable if inference across parameters or scales is not a primary goal of a monitoring program. For example, the rotating panel design outperformed all other designs in estimating the site-specific effect, suggesting that the choice of monitoring design depends on the parameters of interest to managers. Though the other designs (rotating panel, split panel, and weighted effort) had higher RMSE values relative to the stratified random design for most parameters, that difference was marginal in many instances (i.e., <15% difference in RMSE) (Table 1), particularly when effort was high. Additionally, the stratified random design had the lowest mean and variation of unit-level estimates, but other designs typically performed better for individual units (most consistently the indicator unit design) (Table 1). Thus, the preferred design for data collection depends on the monitoring objectives and spatial scale of interest, and there will necessarily be tradeoffs in parameter accuracy (Figure 2).

Our analyses were motivated by our work with the NCRN Inventory and Monitoring program of the US National Park Service. The NCRN Vital Signs monitoring program seeks to provide an understanding of the condition of national parks in the Washington, DC, metropolitan area and identify appropriate management actions necessary to maintain natural resources in the network of parks (Fancy et al., 2009). As the program considers different monitoring strategies to meet their objectives and budget constraints, we aimed to evaluate the effectiveness of multiple proposed monitoring designs to inform one of their key vital signs, amphibian occurrence and distribution (National Park Service, 2005). For the NCRN and other hierarchically organized systems, the stratified random monitoring design performs best across their primary objectives of understanding the status, trends, and drivers of amphibian occurrence at individual parks and across the network. However, the allocation of monitoring resources must also consider whether equal information is needed at all parks (i.e., units), which may not be the case for decision-making. For example, parks with amphibian populations near an ecological or management threshold (Martin et al., 2011) may require increased information when deciding whether to implement management interventions. While the stratified random design did perform marginally better than the weighted effort design at the regional-level (and across the average of unit-level estimates), the weighted effort design had a lower bound to unit-specific estimates across all four parameters. Thus, the selection of a monitoring design will depend on the information needed across scales and among individual parks in the network. Importantly, we found that the return on monitoring investment was not linear, meaning that the magnitude of increase in accuracy declined as additional sites were sampled. Though our results provide valuable information concerning the tradeoffs of different sampling designs applicable to real-world decision-making (e.g., in the NCRN and other hierarchically structured amphibian networks), these results are also generalizable because the parameter space we use is relevant to a variety of taxa and systems (Sanderlin et al., 2014; Sutherland et al., 2016).

Monitoring objectives and constraints will vary across programs, so a balance of scale and parameters of focus may not always be necessary, beneficial, or efficient. Here, we evaluated the performance across three parameter estimates common to published monitoring programs—status, trends, and drivers—at two management-relevant scales to identify and understand tradeoffs that might arise in large-scale and long-term monitoring programs. Although we

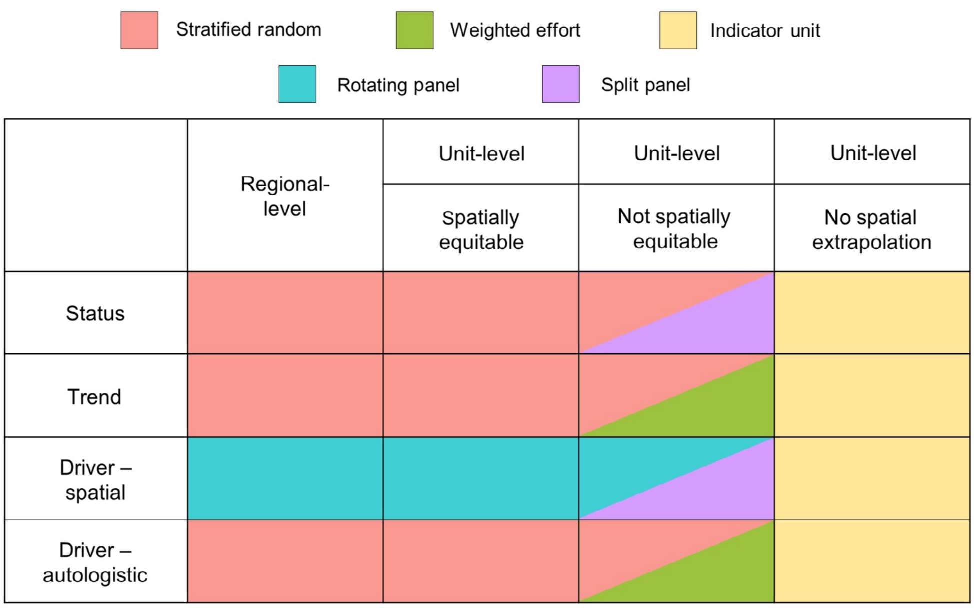

FIGURE 2 The recommended monitoring design(s) for each spatial scale, objective, and parameter combination. Recommendations were determined by which design(s) had the lowest root-mean-square error (RMSE) estimates within the simulation study. Parameters evaluated were α0 (mean occupancy) for status, α1 (year-specific effect) for trend, α2 (site-specific effect) for driver–spatial, and α3 (autologistic effect) for driver–autologistic. Parameters were assessed at the regional (e.g., for status: $\bar{\mu}_{\alpha 0}$) and unit (e.g., for status: $\mu_{\alpha 0,r}$) levels

demonstrated that the stratified random design was preferable in most cases, real-world factors may support the implementation of other sampling designs. For example, if information is needed primarily to understand the drivers of species distributions across space, a hypothesis-based approach (Nichols & Williams, 2006) that emphasizes spatial replication and increased sample size may be more desirable. In this case, the rotating panel design, which sacrifices temporal replication (across years) for an increased number of sites, may be the preferred design. As a tradeoff, this design choice would limit the ability of the monitoring program to detect and respond to declines as they arise.

The importance of information for individual management or governance units (e.g., refuges, states) may not be equal across a broad geographical extent. This can occur when the system at an individual unit is far from the decision or management threshold or the decision at that unit is insensitive to the system state (Martin et al., 2011). The stratified random design had the lowest mean and variation among RMSE estimates across all units, but other designs had lower bounds to the RMSE estimates at the unit level (e.g., Figure 1b). In cases in which regional-level estimates are of lesser importance relative to management objectives at select units, designs such as the weighted effort or split panel, in which a subset of the units receives a disproportionate amount of effort, may be preferred. Selection of one of those two designs will vary depending on the focus of the monitoring program: The weighted effort design better captured temporal variation (e.g., trend and autologistic effect), whereas the split panel design better captured spatial variation (e.g., status and site-specific effect) (Figure 2). Further, not all large-scale monitoring programs aim to extrapolate findings to a broader spatial extent. For example, the NSF LTER Network is designed to provide highly detailed information to understand long-term ecological phenomena at spatially independent locations (Callahan, 1984). In such cases, when the regional-level estimates are not of primary importance, the indicator unit design, in which resources are targeted at select units instead of balancing at the regional level, may be most efficient.

Evaluating monitoring program designs is important as we seek to understand, manage, and conserve the world's ecosystems. The use of evidence-based decision-making to guide the design and objectives of large-scale monitoring programs is necessary to ensure justification and accountability of relevant information-gathering investments (Wintle et al., 2010). A number of considerations should be taken into account in designing and implementing effective programs aimed at monitoring biological communities across spatial scales (Olsen et al., 1999). Past research focused on developing monitoring approaches that account for observation biases (Guillera-Arroita et al., 2010; MacKenzie & Royle, 2005), spatial variation in species distributions or abundance (Pollock et al., 2002), and species rarity (Pacifici et al., 2012; Sanderlin et al., 2014). However, less research has focused on dealing with monitoring objectives that differ across and within scales in a collaborative monitoring network.

While our results inform the tradeoffs of various monitoring objectives in multiscale systems, future research and development can further enhance the practice of monitoring wildlife communities across large scales. First, our analyses assumed that the per-wetland cost of monitoring was the same across all designs, although we recognized that the associated costs with each design could vary across different real-world scenarios. Future research that incorporates cost explicitly can inform monitoring expenditure decisions more optimally. Second, the efficacy or tradeoffs among different sampling designs could shift if spatial units of interest were considerably larger or heterogeneous (beyond that described with the covariate used in our study). In such cases, additional simulations may be necessary to fully address these issues, particularly for communities with many habitat specialists. Third, the size and composition (e.g., rare vs. common species) of biological communities might exert an influence on the preferred sampling design, as well as the temporal extent (years of sampling) and grain of sampling (number of replicates per year). Finally, adjusting the design of large-scale monitoring programs may be logistically challenging or infeasible. Thus, optimizing the data collection process may not always be the appropriate response. With increased access to data from other monitoring programs, various research labs, and volunteer-collected data, future research that leverages integrative analyses and multiple data sources (e.g., integrated population models) can enhance existing and future monitoring programs (Saunders et al., 2019; Zipkin et al., 2021).

Only 25% of the budget necessary to implement threatened species recovery plans is allocated annually in the United States (Gerber, 2016). Approximately half of these recovery resources are dedicated to research and monitoring—not on-the-ground management actions (Buxton et al., 2020). Hence, increasing the efficiency of monitoring programs has the potential to free up resources for management activities. At the center of this issue, though, is how and where in a landscape to most efficiently use available resources (e.g., targeted vs. surveillance monitoring) (Nichols & Williams, 2006; Wintle et al., 2010). Careful consideration of the management context, objectives, and specification of desired accuracy of various parameters can help achieve large-scale monitoring objectives that aim to inform and guide science, management, and policy at multiple scales.

ACKNOWLEDGMENTS

This paper was greatly improved by comments from Kayla Davis, Matt Farr, and Sarah Saunders. This project was funded by the NPS Inventory and Monitoring Program. This work was supported in part by Michigan State University through computational resources provided by the Institute for Cyber-Enabled Research. This is Contribution Number 830 of the US Geological Survey Amphibian Research and Monitoring Initiative (ARMI). Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the US government.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Empirical data were not used for this research. All code is available from Zenodo (Wright, 2021) https://doi.org/10.5281/zenodo.4577521.

ORCID

Alexander D. Wright https://orcid.org/0000-0002-0151-9248

Elise F. Zipkin https://orcid.org/0000-0003-4155-6139

REFERENCES

Adams, M. J., and E. Muths. 2019. "Conservation Research across Scales in a National Program: How to Be Relevant to Local Management Yet General at the Same Time." *Biological Conservation* 236: 100–6.

Albert, C. H., N. G. Yoccoz, T. C. Edwards, Jr., C. H. Graham, N. E. Zimmermann, and W. Thuiller. 2010. "Sampling in Ecology and Evolution–Bridging the Gap between Theory and Practice." *Ecography* 33: 1028–37.

Beaudrot, L., J. A. Ahumada, T. O'Brien, P. Alvarez-Loayza, K. Boekee, A. Campos-Arceiz, D. Eichberg, et al. 2016. "Standardized Assessment of Biodiversity Trends in Tropical Forest Protected Areas: The End Is Not in Sight." *PLoS Biology* 14: e1002357.

Bennett, J. R., S. L. Maxwell, A. E. Martin, I. Chadès, L. Fahrig, and B. Gilbert. 2018. "When to Monitor and when to Act: Value of Information Theory for Multiple Management Units and Limited Budgets." *Journal of Applied Ecology* 55: 2102–13.

Blanchet, F. G., K. Cazelles, and D. Gravel. 2020. "Co-Occurrence Is Not Evidence of Ecological Interactions." *Ecology Letters* 23: 1050–63.

Boulinier, T., J. D. Nichols, J. R. Sauer, J. E. Hines, and K. H. Pollock. 1998. "Estimating Species Richness: The Importance of Heterogeneity in Species Detectability." *Ecology* 79: 1018–28.

Butchart, S. H., M. Walpole, B. Collen, A. Van Strien, J. P. Scharlemann, R. E. Almond, J. E. Baillie, et al. 2010. "Global Biodiversity: Indicators of Recent Declines." *Science* 328: 1164–8.

Buxton, R. T., S. Avery-Gomm, H. Y. Lin, P. A. Smith, S. J. Cooke, and J. R. Bennett. 2020. "Half of Resources in Threatened Species Conservation Plans Are Allocated to Research and Monitoring." *Nature Communications* 11: 4668.

Callahan, J. T. 1984. "Long-Term Ecological Research." *Bio Science* 34: 363–7.

Carlson, M., and F. Schmiegelow. 2002. "Cost-Effective Sampling Design Applied to Large-Scale Monitoring of Boreal Birds." *Conservation Ecology* 6: 11.

Dobbie, M. J., B. L. Henderson, and D. L. Stevens, Jr. 2008. "Sparse Sampling: Spatial Design for Monitoring Stream Networks." *Statistics Surveys* 2: 113–53.

Dorazio, R. M., M. Kery, J. A. Royle, and M. Plattner. 2010. "Models for Inference in Dynamic Metacommunity Systems." *Ecology* 91(8): 2466–75.

Dorazio, R. M., and J. A. Royle. 2005. "Estimating Size and Composition of Biological Communities by Modeling the Occurrence of Species." *Journal of the American Statistical Association* 100: 389–98.

Dorazio, R. M., J. A. Royle, B. Söderström, and A. Glimskär. 2006. "Estimating Species Richness and Accumulation by Modeling Species Occurrence and Detectability." *Ecology* 87: 842–54.

Eyre, T. J., A. Fisher, L. P. Hunt, and A. S. Kutt. 2011. "Measure it to Better Manage it: A Biodiversity Monitoring Framework for the Australian Rangelands." *The Rangeland Journal* 33: 239–53.

Fancy, S. G., J. E. Gross, and S. L. Carter. 2009. "Monitoring the Condition of Natural Resources in US National Parks." *Environmental Monitoring and Assessment* 151: 161–74.

Field, S. A., A. J. Tyre, and H. P. Possingham. 2005. "Optimizing Allocation of Monitoring Effort under Economic and Observational Constraints." *The Journal of Wildlife Management* 69: 473–82.

S.20 - 103rd Congress (1993-1994): Government Performance and Results Act of 1993. Library of Congress, 3 August 1993, https://www.congress.gov/bill/103rd-congress/senate-bill/20.

Gerber, L. R. 2016. "Conservation Triage or Injurious Neglect in Endangered Species Recovery." *Proceedings of the National Academy of Sciences* 113: 3563–6.

Gelman, A., and D. B. Rubin. 1992. "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science* 7: 457–72.

Gelman, A., and K. Shirley. 2011. "Inference from Simulations and Monitoring Convergence." In *Handbook of Markov Chain Monte Carlo*, edited by S. Brooks, A. Gelman, G. L. Jones, and X. L. Meng, 163–74. Boca Raton, FL: CRC Press.

Grant, E. H. C., and A. B. Brand. 2012. "National Capital Region Network Amphibian Monitoring Protocol: Revision 1.4 10 January 2012." Natural Resource Technical Report. Fort Collins, CO: National Park Service.

Guillera-Arroita, G., M. S. Ridout, and B. J. Morgan. 2010. "Design of Occupancy Studies with Imperfect Detection." *Methods in Ecology and Evolution* 1: 131–9.

Jones, J. P. 2011. "Monitoring Species Abundance and Distribution at the Landscape Scale." *Journal of Applied Ecology* 48: 9–13.

Kao, R. H., C. M. Gibson, R. E. Gallery, C. L. Meier, D. T. Barnett, K. M. Docherty, K. K. Blevins, et al. 2012. "NEON Terrestrial Field Observations: Designing Continental-Scale, Standardized Sampling." *Ecosphere* 3: 1–7.

Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. "A Continental Strategy for the National Ecological

Observatory Network." *Frontiers in Ecology and the Environment* 6: 282–4.

Kellner, K. 2015. "jags UI: a wrapper around rjags to streamline JAGS analyses." R Package Version 1.1. https://cran.r-project.org/web/packages/jagsUI/index.html.

Lindenmayer, D. B., and G. E. Likens. 2010. "The Science and Application of Ecological Monitoring." *Biological Conservation* 143: 1317–28.

Lunn, D., C. Jackson, N. Best, D. Spiegelhalter, and A. Thomas. 2012. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton: CRC Press.

MacKenzie, D. I., and J. A. Royle. 2005. "Designing Occupancy Studies: General Advice and Allocating Survey Effort." *Journal of Applied Ecology* 42: 1105–14.

Mac Kenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. "Estimating Site Occupancy Rates when Detection Probabilities Are Less than One." *Ecology* 83: 2248–55.

Martin, J., P. L. Fackler, J. D. Nichols, B. C. Lubow, M. J. Eaton, M. C. Runge, B. M. Stith, and C. A. Langtimm. 2011. "Structured Decision Making as a Proactive Approach to Dealing with Sea Level Rise in Florida." *Climatic Change* 107: 185–202.

McDonald, T. L. 2003. "Review of Environmental Monitoring Methods: Survey Designs." *Environmental Monitoring and Assessment* 85: 277–92.

National Park Service. 2005. *Long-term Monitoring Plan for Natural Resources in the National Capitol Region Network*. Washington, D.C.: Center for Urban Ecology.

Nicol, S., J. Brazill-Boast, E. Gorrod, A. McSorley, N. Peyrard, and I. Chadès. 2019. "Quantifying the Impact of Uncertainty on Threat Management for Biodiversity." *Nature Communications* 10: 1–14.

Nichols, J. D., and B. K. Williams. 2006. "Monitoring for Conservation." *Trends in Ecology & Evolution* 21: 668–73.

Olsen, A. R., J. Sedransk, D. Edwards, C. A. Gotway, W. Liggett, S. Rathbun, K. H. Reckhow, and L. J. Young. 1999. "Statistical Issues for Monitoring Ecological and Natural Resources in the United States." *Environmental Monitoring and Assessment* 54: 1–45.

Pacifici, K., R. M. Dorazio, and M. J. Conroy. 2012. "A Two-Phase Sampling Design for Increasing Detections of Rare Species in Occupancy Surveys." *Methods in Ecology and Evolution* 3: 721–30.

Plummer, M. 2003. "JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling." In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, Vienna, Austria, Vol 124, edited by K. Hornik, F. Leisch, and A. Zeileis, 1–10. http://www.ci.tuwien.ac.at/Conferences/DSC-2003/.

Pollock, K. H., J. D. Nichols, T. R. Simons, G. L. Farnsworth, L. L. Bailey, and J. R. Sauer. 2002. "Large Scale Wildlife Monitoring Studies: Statistical Methods for Design and Analysis." *Environmetrics* 13: 105–19.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Royle, J. A., R. M. Dorazio, and W. A. Link. 2007. "Analysis of Multinomial Models with Unknown Index Using Data Augmentation." *Journal of Computational and Graphical Statistics* 16: 67–85.

Sanderlin, J. S., W. M. Block, and J. L. Ganey. 2014. "Optimizing Study Design for Multi-Species Avian Monitoring Programmes." *Journal of Applied Ecology* 51: 860–70.

Sauer, J. R., W. A. Link, J. E. Fallon, K. L. Pardieck, and D. J. Ziolkowski. 2013. "The North American Breeding Bird Survey 1966-2011: Summary Analysis and Species Accounts." *North American Fauna* 79: 1–32.

Sauer, J. R., K. L. Pardieck, D. J. Ziolkowski, A. C. Smith, M. A. R. Hudson, V. Rodriguez, H. Berlanga, D. K. Niven, and W. A. Link. 2017. "The First 50 Years of the North American Breeding Bird Survey." *The Condor: Ornithological Applications* 119: 576–93.

Saunders, S. P., M. T. Farr, A. D. Wright, C. A. Bahlai, J. W. Ribeiro, S. Rossman, A. L. Sussman, T. W. Arnold, and E. F. Zipkin. 2019. "Disentangling Data Discrepancies with Integrated Population Models." *Ecology* 100: e02714.

Scholes, R. J., G. M. Mace, W. Turner, G. N. Geller, N. Jürgens, A. Larigauderie, D. Muchoney, B. A. Walther, and H. A. Mooney. 2008. "Toward a Global Biodiversity Observing System." *Science* 321: 1044–5.

Schreuder, H. T., R. Ernst, and H. Ramirez-Maldonado. 2004. *Statistical Techniques for Sampling and Monitoring Natural Resources*. General Technical Report RMRS-GTR-126. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Sparrow, B. D., W. Edwards, S. E. Munroe, G. M. Wardle, G. R. Guerin, J. F. Bastin, B. Morris, R. Christensen, S. Phinn, and A. J. Lowe. 2020. "Effective Ecosystem Monitoring Requires a Multi-Scaled Approach." *Biological Reviews* 95: 1706–19.

Stadt, J. J., J. Schieck, and H. A. Stelfox. 2006. "Alberta Biodiversity Monitoring Program–Monitoring Effectiveness of Sustainable Forest Management Planning." *Environmental Monitoring and Assessment* 121: 33–46.

Sutherland, C., M. Brambilla, P. Pedrini, and S. Tenan. 2016. "A Multiregion Community Model for Inference about Geographic Variation in Species Richness." *Methods in Ecology and Evolution* 7: 783–91.

Sutherland, W. J., A. S. Pullin, P. M. Dolman, and T. M. Knight. 2004. "The Need for Evidence-Based Conservation." *Trends in Ecology & Evolution* 19: 305–8.

Thompson, S. K. 2012. *Stratified Sampling*. Sampling; Wiley Series in Probability and Statistics 139–56. Hoboken, NJ: John Wiley & Sons.

Thorpe, A. S., D. T. Barnett, S. C. Elmendorf, E. L. S. Hinckley, D. Hoekman, K. D. Jones, K. E. LeVan, C. L. Meier, L. F. Stanish, and K. M. Thibault. 2016. "Introduction to the Sampling Designs of the National Ecological Observatory Network Terrestrial Observation System." *Ecosphere* 7: e01627.

Wintle, B. A., M. C. Runge, and S. A. Bekessy. 2010. "Allocating Monitoring Effort in the Face of Unknown Unknowns." *Ecology Letters* 13: 1325–37.

Wright, A. D. 2021. "zipkinlab/Wright_etal_mon-simul: Wright et al., A Comparison of Monitoring Designs to Assess Wildlife Community Parameters across Spatial Scales." Version v1.0. Zenodo. https://doi.org/10.5281/zenodo.4577522.

Wright, A. D., E. H. C. Grant, and E. F. Zipkin. 2020. "A Hierarchical Analysis of Habitat Area, Connectivity, and Quality on Amphibian Diversity across Spatial Scales." *Landscape Ecology* 35: 529–44.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. "Monitoring of Biological Diversity in Space and Time." *Trends in Ecology & Evolution* 16: 446–53.

Ziolkowski, D., K. Pardieck, and J. R. Sauer. 2010. "On the Road Again for a Bird Survey that Counts." *Birding* 42: 32–41.

Zipkin, E. F., E. H. C. Grant, and W. F. Fagan. 2012. "Evaluating the Predictive Abilities of Community Occupancy Models Using AUC while Accounting for Imperfect Detection." *Ecological Applications* 22: 1962–72.

Zipkin, E. F., E. R. Zylstra, A. D. Wright, S. P. Saunders, A. O. Finley, M. C. Dietze, M. S. Itter, and M. W. Tingley. 2021. "Addressing Data Integration Challenges to Link Ecological Processes across Scales." *Frontiers in Ecology and the Environment* 19: 30–8.

Zipkin, E. F., J. A. Royle, D. K. Dawson, and S. Bates. 2010. "Multi-Species Occurrence Models to Evaluate the Effects of Conservation and Management Actions." *Biological Conservation* 143: 479–84.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Wright, Alexander D., Evan H. Campbell Grant, and Elise F. Zipkin. 2022. "A Comparison of Monitoring Designs to Assess Wildlife Community Parameters across Spatial Scales." *Ecological Applications* 32(6): e2621. https://doi.org/10.1002/eap.2621