

Power to detect trends in abundance within a distance sampling framework

Erik M. Andersen  | Robert J. Steidl 

School of Natural Resources and the Environment, University of Arizona, Tucson, AZ, USA

Correspondence

Erik M. Andersen
Email: erikandersen@email.arizona.edu

Funding information

Arizona Game and Fish Department, Grant/Award Number: I14008; U.S. Bureau of Land Management, Grant/Award Number: 3007730; Audubon Apacheria Fellowship

Handling Editor: Fernanda Michalski

Abstract

1. Ensuring that inferences from biological monitoring are reliable requires a level of sampling effort that is commensurate with programmatic objectives and reflects attributes of target populations. Consensus guidelines have emerged to guide allocation of sampling effort for programmes designed to detect trends in occupancy but not for abundance, despite its prevalence as a target parameter.
2. We evaluated the influence of population attributes (density, availability, detection probability and magnitude of trend) and sampling design features (number of sites, number of repeat surveys, and survey-year interval) on a Bayesian analogue of statistical power to detect declines in abundance estimated using distance sampling methods. For a range of values common to terrestrial vertebrates, we simulated spatially and temporally replicated populations from which we generated survey data. We then analysed each dataset with a hierarchical open-population model that allowed for temporary emigration to estimate power for periods of 5–20 years.
3. For a given amount of sampling effort, power to detect trends was highest when effort was allocated to maximizing the number of sites by decreasing the number of repeat surveys within a year and increasing the interval between survey years. For example, to have an 80% chance of detecting a 3% annual decline required 40% longer when 67 sites were surveyed three times per year compared to 600 sites surveyed once every three years despite both allocations requiring 4,000 surveys over 20 years. Notably, these patterns were independent of density or detectability of the target species, which contrasts with occupancy studies where optimal allocation shifts from surveying more sites to more repeat surveys when detectability is low or occupancy is high.
4. *Synthesis and applications.* Our findings provide guidance for allocating resources efficiently for distance sampling studies focused on terrestrial vertebrates. By comparing the approximate density and detectability of target populations to those we considered, monitoring programmes can balance the amount of survey effort allocated to sites, surveys per site, and annual revisits to help ensure sufficient power to meet objectives effectively and efficiently. These decisions are increasingly important as budgets for conservation decrease and consequences of inaction continue to increase.

KEYWORDS

distance sampling, monitoring, occupancy, power analysis, sample size, sampling design, statistical power, trend detection

1 | INTRODUCTION

To be effective, ecological monitoring programmes must detect trends in target attributes reliably and in sufficient time to develop and implement conservation and recovery strategies (Field, Tyre, & Possingham, 2005; Nichols & Williams, 2006; Wintle, Runge, & Bekessy, 2010). To be efficient, programmes must balance the effort and costs associated with data collection with the risk of failing to detect meaningful trends in target populations (Steidl & Thomas, 2001). If a population attribute such as abundance is declining over time and that decline goes undetected (i.e. a Type II error is made), consequences for species of conservation concern can be substantial (Steidl, Hayes, & Schaubert, 1997). Prospective statistical power analysis provides a unifying framework to compare the effectiveness and efficiency of alternative monitoring strategies and to ensure that programmes are likely to meet their objectives reliably (Guillera-Arroita & Lahoz-Monfort, 2012; Legg & Nagy, 2006; Steidl, Conway, & Litt, 2013). Therefore, designing monitoring programmes with sufficient power to meet the target objectives is essential in preventing inefficient use of resources, unreliable inferences that can have important consequences for conservation, and management decisions that are ill-informed (Legg & Nagy, 2006; Lindenmayer & Likens, 2010; Reynolds, Thompson, & Russell, 2011).

When developing programmes to monitor changes in biological populations, the primary instrument available to influence power is altering the amount and allocation of sampling effort. Deciding how to allocate effort most efficiently is challenging, however, because optimal allocation varies with attributes of the target population, including abundance and detectability during surveys (Nuno, Milner-Gulland, & Bunnefeld, 2015; Steidl et al., 2013). For example, detecting trends in populations of rare species, which are often the focus of conservation concerns, is inherently more difficult than detecting trends for common species (MacKenzie, Nichols, Sutton, Kawanishi, & Bailey, 2005). This difficulty is exacerbated when species are challenging to detect during surveys. Accounting for imperfect detection is an important component of effective monitoring (Kery & Schmid, 2004; Schmidt, McIntyre, & MacCluskie, 2013), but it can increase complexity and uncertainty because detectability is often species-specific and heterogeneous in space and time (Boulinier, Nichols, Sauer, Hines, & Pollock, 1998; MacKenzie et al., 2002). Consequently, understanding how attributes of target species affect the power of monitoring programmes to detect trends reliably is important for allocating resources in ways that are most likely to achieve programmatic goals (Steidl et al., 2013). Ultimately, comparing power prospectively among alternative sampling designs can help to ensure that effort is commensurate with these attributes and that

sampling designs are as efficient and cost-effective as possible (Field et al., 2005).

A number of studies have provided insights and guidance on how to allocate sampling effort efficiently for monitoring programmes where occupancy is the focal population attribute (Barata, Griffiths, & Ridout, 2017; Field et al., 2005; Gálvez, Guillera-Arroita, Morgan, & Davies, 2016; Guillera-Arroita & Lahoz-Monfort, 2012; MacKenzie & Royle, 2005; Sanderlin, Block, & Ganey, 2014). In settings where surveys yield only detection–nondetection data, occupancy provides a useful framework for monitoring changes in distributional attributes of animal populations. In settings where survey data can be used to estimate abundance, however, monitoring efforts are likely to detect population trends more quickly than with occupancy (Field et al., 2005). For a site that is occupied, for example, the occupancy state does not change until local abundance has already declined to zero (Gaston et al., 2000).

One of the most widely used methods for estimating abundance of animal populations is distance sampling, which accounts for imperfect surveys by characterising the detection probability of objects as a function of their distance from a transect or point (Buckland et al., 2001, 2004; Burnham, Anderson, & Laake, 1980). The method does not require individuals to be captured or marked, making it more efficient and cost-effective than many other methods. Development of the distance sampling framework as a hierarchical model has facilitated inferences about features that govern variation in abundance (Hedley & Buckland, 2004; Royle, Dawson, & Bates, 2004) and in the detection process (Silllett, Chandler, Royle, Kery, & Morrison, 2012), making this framework especially useful for monitoring programmes that operate over large spatial and temporal scales. Historically, abundance-based monitoring programmes used linear regressions of annual abundance estimates to evaluate population trends (Buckland et al., 2004; Gerrodette, 1987). The recent development of open-population distance sampling models has relaxed the closure assumption, however, and enabled trends to be estimated in a single model (Kery & Royle, 2015). A promising expansion for monitoring programmes that collect data using the robust design where spatial replicates are surveyed multiple times within and across years (Pollock, 1982) are the recently developed hierarchical distance sampling models based on multinomial N-mixture models (Chandler, Royle, & King, 2011). These models allow for temporary emigration, where individuals are temporarily outside of plot boundaries during the survey period or when they do not produce cues that permit detection. By partitioning variation attributable to availability of animals and detection probability, temporary-emigration models avoid biases that can be introduced when these processes are confounded (Chandler et al., 2011; Riddle, Stanislav, Pollock, Moorman, & Perkins, 2010).

Although monitoring programmes focused on detecting trends in abundance of animal populations are common (Marques et al., 2017; Schmidt et al., 2013), comprehensive advice for allocating sampling effort is sparse (Ficetola, Romano, Salvidio, & Sindaco, 2017; Field et al., 2005). General guidance derived from occupancy studies may apply to abundance-based monitoring, but differences in the data required to estimate occupancy versus abundance are likely to affect how to allocate sampling effort optimally. For example, estimating detection probabilities using standard occupancy designs typically requires >1 survey per site (but see Lele, Moreno, & Bayne, 2012; Peach, Cohen, & Frair, 2017), but detection probabilities can be estimated with data from a single survey when using distance sampling. Here, we expand on previous work focused on allocating survey effort for occupancy-based monitoring to help inform programmes developed for abundance-based monitoring of terrestrial vertebrates. Specifically, we address three questions pertaining to monitoring programmes that aim to evaluate trends in abundance while accounting for imperfect detection with distance sampling techniques: (a) How do population and life-history attributes of target species affect power to detect trends? (b) How do features of the sampling design affect power to detect trends? (c) How should sampling effort be allocated to maximize efficiency? To answer these questions, we used a simulation-based approach (Buckland et al., 2004) to explore how variation in abundance and detectability and different amounts and allocations of sampling effort affect the power of a monitoring programme to detect trends in abundance. Specifically, for a range of values common to terrestrial vertebrates, we simulated spatially- and temporally-replicated populations from which we generated survey data. We then analysed each dataset with a hierarchical open-population model that allowed for temporary emigration to estimate power for periods of 5–20 years. Our overarching objective is to help inform design of monitoring programmes so that they are more likely to meet programmatic objectives while optimizing the limited resources available for surveys.

TABLE 1 Values for population attributes and sampling design features used in simulations to estimate the power of monitoring programmes to detect trends in abundance. To determine the relative effect of each parameter on power, we estimated power for the range of values for each parameter independently while holding values for all other parameters constant

Attribute or feature	Definition	Values	Value when constant
Trend (τ)	Log-linear annual rate of population change (%)	-1, -2, -3, -4, -5	-3
Density (D)	Mean initial number of individuals/ha	0.1, 0.2, 0.4, 0.8, 1.6	0.3
Availability (ϕ)	Probability of an individual temporarily beyond survey limits or not producing a cue	0.4, 0.6, 0.8, 1	0.7
Scale of detection function (σ)	Scale parameter for a half-normal detection function (m)	30, 45, 75, 135	75
Sites (s)	No. of spatial replicates surveyed each survey year	25, 50, 100, 200	100
Repeat surveys (k)	No. of surveys per site per survey year	1, 2, 3, 4	2
Survey-year interval (f)	No. of years between surveys at a site (i.e. annual, biennial, and triennial survey years)	1, 2, 3	1

2 | MATERIALS AND METHODS

We explored how attributes of animal populations and features of sampling designs affect power to detect temporal trends in abundance. Specifically, we evaluated how power changed in response to variation in four parameters related to abundance and detectability: (a) population trend, τ , (b) initial density of animals, D , (c) availability of animals during surveys, ϕ , and (d) detection probability, in the form of the scale parameter for a half-normal detection function, σ (Table 1). We labelled this set of four parameters, which reflect demography or traits of a species or population, as 'population attributes'. Similarly, we evaluated how power varied in response to changes in three features of sampling designs used commonly for monitoring: (a) sites, s , (b) repeat surveys, k , and (c) survey-year interval, f (Table 1). We labelled this set of three parameters as 'design features'. We explored how these population attributes and design features affected power by simulating data for a range of 4–5 plausible values for each parameter while holding values for all other characteristics constant at intermediate levels (Table 1). Although some population attributes are inherently bounded (e.g. ϕ) or effectively governed by the spatial extent of the sample unit (e.g. σ), densities of animal populations vary widely. To select the values of density to explore, we searched the literature and determined the range of densities for 189 species of terrestrial vertebrates that were surveyed with distance sampling techniques (Appendix S1). Consequently, the values we explored are relevant to a wide array of terrestrial vertebrates and design alternatives based on distance sampling methods. Although we report density as the number of individuals per ha, our findings apply to other independent survey units, such as clusters of animals, territorial males, or burrows.

2.1 | Simulating data

We simulated spatially and temporally replicated data using the function `simHDSopen` from the R package `AHMBOOK` (Kery & Royle, 2015,

p. 499) that we modified to meet our requirements. Specifically, we specified sets of population attributes and design features through a three-level model that included an abundance model, a process model, and a sampling model. The abundance model governed initial abundances of animals at each site (s), which we established as 3.1-ha plots. Although we specified circular plots (100-m radius), our choice of plot shape is arbitrary and results apply equally to rectangular plots that are characteristic of transect surveys. We assigned a value for initial abundance at each site as

$$M_{s,1} \sim \text{Negative Binomial}(\lambda, \lambda + \lambda^2),$$

where M is true abundance of site s at year $t = 1$ and λ is mean initial abundance of animals per site across the study area (3.1 ha \times mean density, D). We chose the negative binomial because it is useful for describing distributions of animals in heterogeneous landscapes (Royle, 2004), and used $\lambda + \lambda^2$ to characterize dispersion because it produced abundance distributions with a moderate degree of heterogeneity. The process model governed persistence of individuals at each site over time, which we treated as a stochastic, Markov process to mimic the irregular dynamics common in natural populations (Field et al., 2005). Specifically, we represented persistence of individuals at each site s as

$$M_{s,t} \sim \text{Binomial}(M_{s,t-1}, 1 - \tau),$$

where τ is the annual log-linear trend in abundance. Although we simulated negative trends, patterns for positive trends would be similar (Gerrodette, 1987). We then established the number of individuals that were available to be detected (N) during each survey k from those assigned to a site as

$$N_{s,k,t} \sim \text{Binomial}(M_{s,t}, \phi),$$

where ϕ is the probability of availability. We did not modify the sampling model of the `simHDSopen` function, which represented the process of surveying animals with distance sampling methods following a point-transect protocol for k surveys in a survey-year during a period when the population is assumed to be closed (Table 1). We established distances from the survey point to each individual available to be detected with a polar-coordinate method (Kery & Royle, 2015) and aggregated detections as counts in each of 5, 20-m distance classes. We then determined whether each available individual was detected during a survey as

$$y_{s,k,t} \sim \text{Binomial}(N_{s,k,t}, p_{s,k,t}),$$

where detection probability, p , was governed by the scale parameter σ for a half-normal detection function (Buckland et al., 2001), which we allowed to vary with a survey-level covariate that we assigned as a random draw from a uniform distribution $(-2, 2)$ as

$$\log(\sigma_{s,k,t}) = \log(\sigma) + \beta(\text{covariate}_{s,k,t})$$

(Kery & Royle, 2015, p. 499); we specified the coefficient for this covariate (β) equal to 0.2.

2.2 | Analyses

We analysed each simulated dataset with a hierarchical multinomial N -mixture model for open populations that allows for temporary emigration (Chandler et al., 2011; Sillett et al., 2012). We used the Bayesian implicit dynamics formulation for distance sampling data described by Kery and Royle (2015, pp. 510–513) that includes a trend parameter, which would be the focal parameter for many long-term monitoring programmes. We note that using a different model for analysis would affect power estimates. We fit models using the Markov-chain Monte-Carlo approach in JAGS (Plummer, 2018) through the R package `JAGSUI` (Kellner, 2016). We specified uniform distributions for $\phi(0, 1)$ and $\sigma(0, 200)$, and normal distributions ($\mu = 0$, $SD = 10$) for all other parameters, assuming that designers of monitoring programmes will have some prior knowledge of their target system. We sampled the posterior distribution with three independent Markov chains, each set to 15,000 samples with a burn-in of 3,000 samples and no thinning. During preliminary analyses, MCMC algorithms converged reliably to the joint posterior distribution with this level of sampling based on Brooks–Gelman–Rubin statistics ($\hat{R} < 1.05$; Gelman, Carlin, Stern, & Rubin, 2004) and visual inspection of trace plots.

We simulated and analyzed 500 datasets for each of the 29 combinations of parameter values we explored (Table 1). For each combination, we computed the proportion of times that the 95% credible interval for the trend parameter correctly excluded 0, which is a quantity analogous to statistical power. We repeated this process for studies from 5 to 20 years in duration to evaluate changes in power with time, which required 232,000 different datasets. We report the number of years required to detect average annual decreases in abundance (hereafter, ‘declines’) with power ≥ 0.80 , the arbitrary reference level we established for this power-like statistic. We note that establishing a more conservative level, such as 0.9, might be more appropriate for analysis of monitoring data for species of conservation concern (Shrader-Frechette & McCoy, 1992), but this change would have no effect on the relative influence of different aspects of the sampling design (i.e. thresholds would increase similarly for all design combinations), which was the focus of our study.

2.3 | Grassland birds case study

To illustrate how design features and data collected during typical field surveys combine to influence power to detect trends in abundance, we used the process described above to explore parameter values relevant to four species of sparrow (Passerellidae) that breed in semidesert grasslands of the southwestern United States. Relative to each other, these species differ appreciably in density, availability, and detection probability (Andersen, 2019). Botteri’s sparrow *Peucaea botterii* is abundant and moderately conspicuous (see

TABLE 2 Values for attributes used to simulate abundance and survey data for four species of breeding birds that differ in density, availability, and detection probability. Values for Botteri's sparrow (BOSP), grasshopper sparrow (GRSP), Cassin's sparrow (CASP), and rufous-crowned sparrow (RCSP) from data collected in semidesert grasslands of southern Arizona, 2013 to 2015 (Andersen, 2019)

Attribute	BOSP	GRSP	CASP	RCSP
Density (D , males/ha)	1.10	0.48	0.13	0.06
Availability (ϕ)	0.6	0.5	0.5	0.4
Scale of half-normal detection function (σ)	88	67	150	62

Table 2 for parameter values that define these terms), grasshopper sparrow *Ammodramus saviarum* is common and inconspicuous, Cassin's sparrow *Peucaea cassinii* is uncommon and conspicuous, and rufous-crowned sparrow *Aimophila ruficeps* is rare and inconspicuous. For each species, we estimated power to detect a 4% annual decline in abundance as a function of survey years with all sampling design features held constant (Table 1).

2.4 | Allocating sampling effort

Although allocating more effort towards any component of the sampling design will always increase power, allocating a fixed amount of effort optimally requires balancing tradeoffs among the number of sites, number of repeat surveys per site per year, and the survey-year interval. To compare power among different sampling designs, we repeated the simulation process using nine different allocations of 4,000 total surveys over 20 years. Specifically, for each permutation of annual, biennial, and triennial surveys, and for 1, 2, and 3 repeat surveys, we varied the number of sites that could be surveyed with 4,000 total surveys allocated across 20 years while holding parameter values for population attributes constant (Table 3). Although we established the total number of surveys over time as the basis for comparing different allocation strategies (i.e. we assume that surveying one site twice requires the same effort as surveying two sites once), we acknowledge that this measure can be affected by other design aspects, including distances among sites and the effort required to measure important site covariates, such as vegetation.

Finally, we explored how attributes of the target species affected how to optimize allocation of effort between sites and repeat surveys. Specifically, we evaluated whether optimal allocation for detecting trends with distance sampling differed from occupancy where allocation shifts from favouring sites to increased surveys when detectability is low or occupancy is high (MacKenzie & Royle, 2005). For six hypothetical species that span a range of densities and detectabilities, we determined the number of sites required to detect a 3% annual decline in abundance with power = 0.80 within a given number of years for 1–4 repeat surveys in each year. The six hypothetical species represent combinations of three levels of

TABLE 3 Number of years required to detect a 3% annual decline in abundance with power = 0.8 for nine allocations of equal survey effort. Each simulated population was surveyed with all combinations of annual, biennial, or triennial surveys and 1, 2, or 3 repeat surveys. For each of these combinations, the number of sites surveyed was varied so that each allocation required 4,000 total surveys over a 20-year period. For all allocations, other parameters were held constant ($D = 0.3$ individuals/ha, $\tau = -3\%$ /year, $\phi = 0.7$, $\sigma = 75$ m)

Survey-year interval (f)	Repeat surveys (k)	Sites (s)	Years required to detect trend
Annual	1	200	10.9
	2	100	12.3
	3	67	13.5
Biennial	1	400	10.6
	2	200	11.6
	3	133	12.6
Triennial	1	600	9.8
	2	300	10.7
	3	200	11.0

detectability ($\phi = 0.8$, $\sigma = 135$ m; $\phi = 0.5$, $\sigma = 80$ m; and $\phi = 0.3$, $\sigma = 30$ m) and two levels of initial density (0.01 and 10 individuals/ha) that span much of the range observed for terrestrial vertebrates surveyed with distance sampling (85% of the 189 species identified in our literature search; Appendix S1). We specified 20 years to detect the target trend for combinations with $D = 0.01$ individuals/ha and 5 years for $D = 10$ individuals/ha.

3 | RESULTS

3.1 | Population attributes

Power to detect trends in abundance decreased as values for population attributes decreased, including magnitude of the trend, availability, detection probability, and initial density (Figure 1). For example, when initial density decreased from 1.6 to 0.1 individuals/ha, the number of years required to detect a 3% decline with power ≥ 0.80 increased by 171% (7 vs. 19 years; Figure 1) with all other values held constant (Table 1). Similarly, detecting a 3% decline required 36% more survey years (11 vs. 15) when availability (ϕ) decreased from 1.0 to 0.4 and 58% more survey years (12 vs. 19) when the scale parameter of the detection function (σ) decreased from 135 to 30 m (Figure 1). As expected, detecting smaller trends required substantially more time than larger declines. For example, detecting a 2% decline required nearly twice as many survey years than a 5% decline (17 vs. 9 years; Figure 1).

The range of values relevant to four grassland sparrows illustrate the effects of these attributes on power to detect trends (Table 2). Detecting a 4% decline required only 7 years for Botteri's sparrow, the species with the highest mean density and availability and with moderate detection probability, and >20 years for rufous-crowned

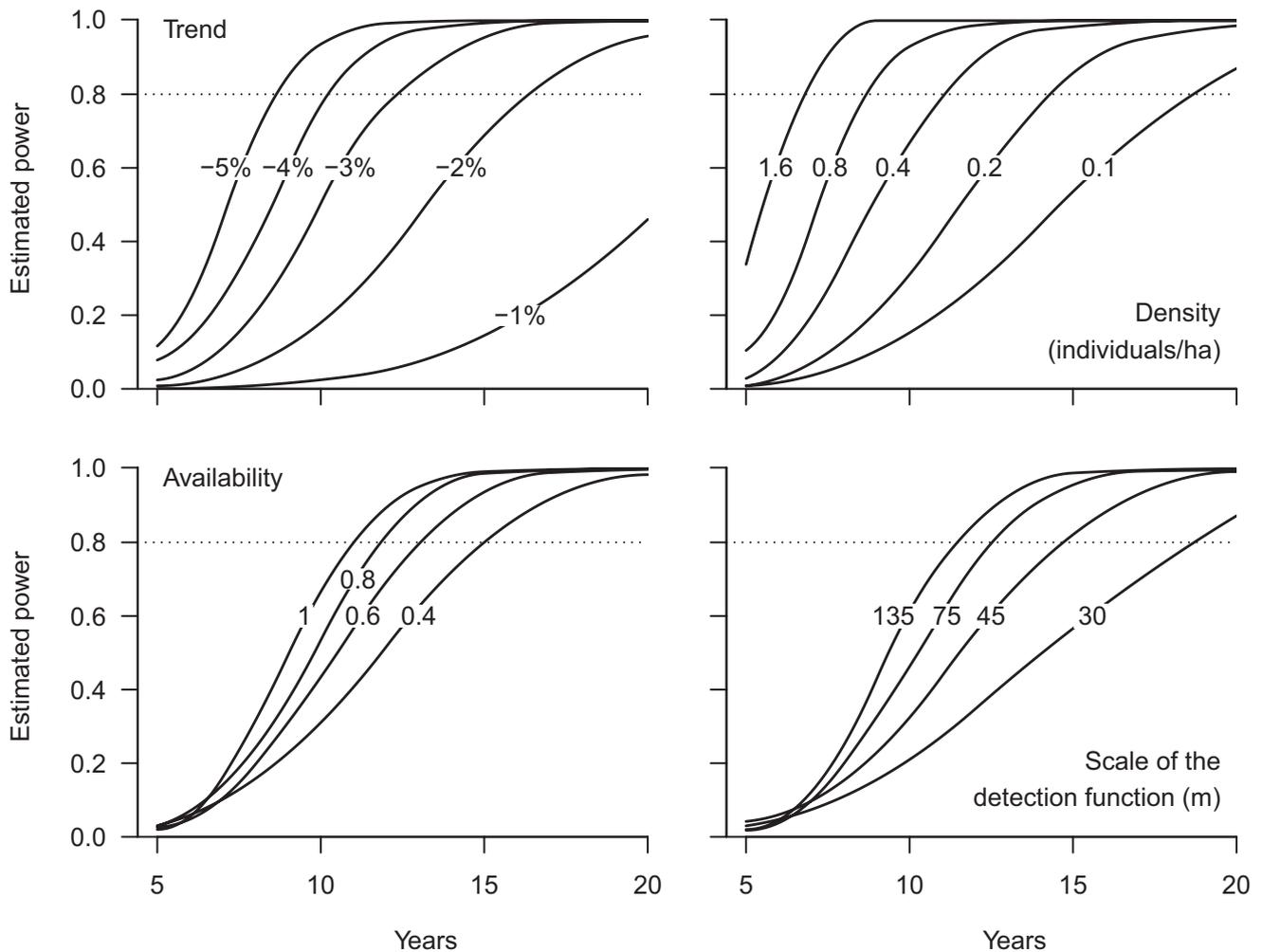


FIGURE 1 Effects of four attributes of animal species or populations on power to detect declines in abundance: population trend (τ), mean initial density (D), probability that an individual is available to be detected (ϕ), and scale parameter of the half-normal detection function (σ), which governs detection probability. Projections are based on parameter values in Table 1

sparrow, the species with the lowest values for mean density, availability, and detection probability (Figure 2). Detecting the same trend required 10 years for grasshopper sparrow, a species that is moderately common, but relatively inconspicuous, and 14 years for Cassin's sparrow, a species that is uncommon, but conspicuous (Figure 2).

3.2 | Sampling design features

The number of sites surveyed, number of repeat surveys, and survey-year interval are the principal components of the sampling design that can be manipulated to influence power. For the range of values we considered (Table 1), allocating more effort towards sites had a larger effect on power than allocating more effort towards repeat surveys (Figure 3). For example, doubling the number of sites decreased the number of years required to detect a 3% decline by 20%–25% (3–5 years), whereas doubling the number of repeat surveys decreased the number of years required by about 10% (2 years; Figure 3). For both sites and repeat surveys, power increased most

rapidly in response to increases at lower values than at higher values (Figure 3). Although increasing the survey-year interval (e.g. from annual to biennial) decreased power (Figure 3), losses in power were small relative to the reduction of sampling effort concomitant with a longer interval between survey years. For example, increasing the survey-year interval from annual to biennial increased the time required to detect a 3% decline by only 15% (13 vs. 15 years), but required 50% fewer total surveys over that time period (Figure 3).

3.3 | Allocating sampling effort

For the nine scenarios we considered to allocate 4,000 surveys over 20 years (Table 3), the time required to detect a 3% decline varied by as much as 40% (range = 10–14 years; Table 3). Given a fixed number of repeat surveys per year, longer survey-year intervals, which allows for additional sites, increased power compared to shorter intervals (Table 3). For example, with two repeat surveys per year, detecting a 3% decline required 11 years for triennial surveys with 300 sites, 12 years

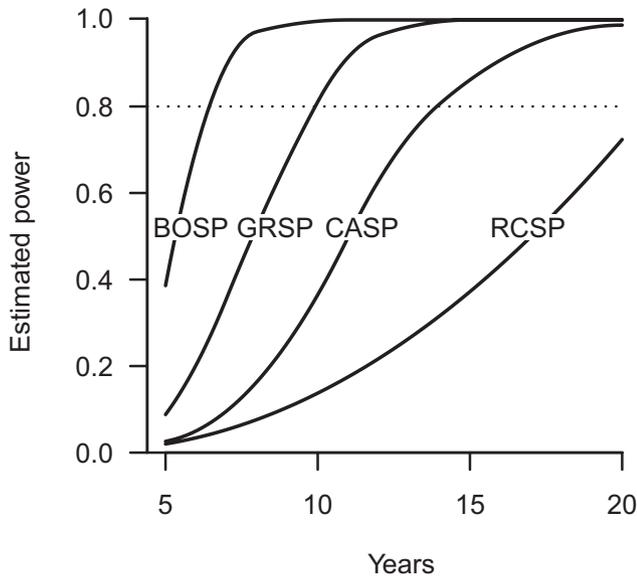


FIGURE 2 Power to detect a 4% annual decline in abundance for four breeding birds that differ in density, availability, and detection probability. Values for Botteri's sparrow (BOSP), grasshopper sparrow (GRSP), Cassin's sparrow (CASP), and rufous-crowned sparrow (RCSP) are based on the parameter values in Table 2

for biennial surveys with 200 sites, and 13 years for annual surveys with 100 sites. Similarly, power was the highest when proportionally more effort was allocated to surveying additional sites rather than increasing the number of repeat surveys, regardless of the survey-year interval (Table 3). Within these scenarios, therefore, the number of years required to detect a decline was lowest when effort was allocated towards surveying as many sites as possible with the fewest number of repeat surveys and the greatest interval between primary periods.

In contrast to occupancy studies, the advantage of allocating survey effort towards more sites at the expense of fewer repeat surveys was constant, regardless of the density and detectability of the target species (Figure 4, Appendix S2). The degree to which allocation towards sites was advantageous did not vary strongly with density, but narrowed with decreasing detectability (Figure 4). For example, for all values of detectability that we considered, the difference between $D = 0.01$ and 10 individuals/ha in percent increase in total survey effort required for 1 versus 4 repeat surveys was <30%. In contrast, for $D = 10$ individuals/ha, the increase in total survey effort required to detect the target trend with 4 repeat surveys relative to 1 was 156% when detectability was very high ($\phi = 0.8$, $\sigma = 135$ m), but only 41% when detectability was very low ($\phi = 0.3$, $\sigma = 30$ m); the difference for $D = 0.01$ individuals/ha was even more pronounced at 195% and 12%, respectively (Figure 4, Appendix S2).

4 | DISCUSSION

Ensuring that programmatic objectives for monitoring can be met reliably requires establishing a level of sampling effort commensurate

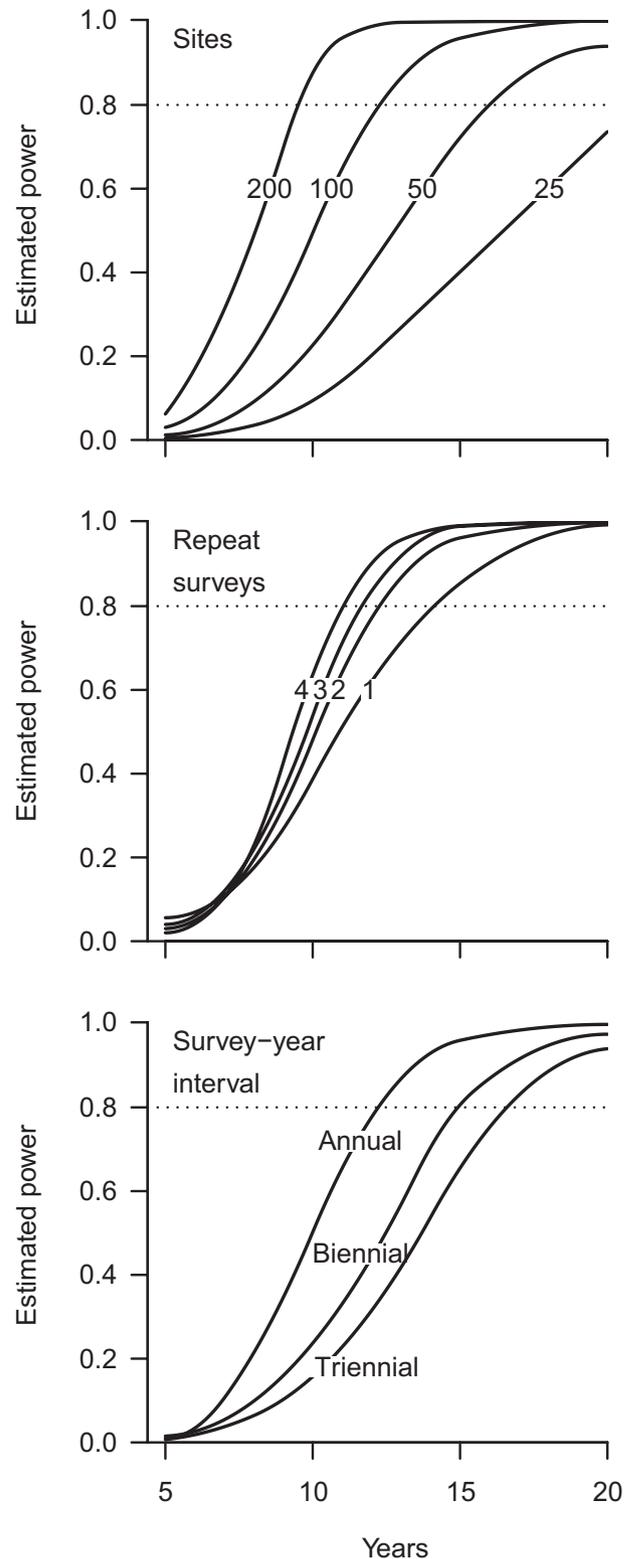


FIGURE 3 Effects of varying three features of the sampling design on power to detect a 3% annual decline in abundance: sites, repeat surveys, and survey-year interval. Projections are based on parameter values in Table 1

with attributes of the target populations and environment (Steidl et al., 2013). Some monitoring programmes, however, lack sufficient power to meet their objectives reliably (Legg & Nagy, 2006).

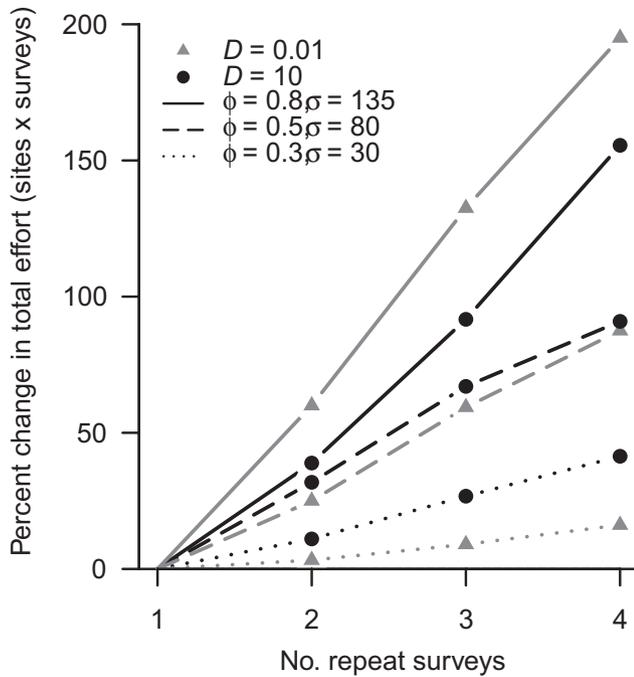


FIGURE 4 Optimal allocation of effort between the number of sites surveyed and the number of repeat surveys per site per year for species that vary in density and detectability. Points represent percent change in total survey effort (sites \times repeat surveys) required to detect a 3% annual decline in abundance with power = 0.80 in 20 years ($D = 0.01$ individuals/ha) or 5 years ($D = 10$ individuals/ha) for 2–4 repeat surveys relative to a single survey each year. Point shape indicates density and line style indicates detectability (e.g. combinations of ϕ and σ). Consequently, survey effort is best allocated towards sites when values are above zero and towards repeat surveys when values are below zero. The number of sites required for each scenario is provided in Appendix S2

Consensus guidelines for allocating sampling effort effectively have emerged from numerous studies focused on detecting trends in occupancy but not for abundance, despite their prevalence in practice (Field et al., 2005; Sanderlin et al., 2014). In this study, we demonstrate the influence of different population attributes and sampling design features on power to detect trends in abundance within a distance sampling framework allowing for temporary emigration. By comparing population attributes of target species to those we considered, our results provide general guidance on the effort required to detect population trends reliably. Because population attributes such as density and probabilities of availability and detection are intrinsic features of the target population and environment, the primary tool available to influence power is allocation of sampling effort.

In general, we found that power was maximized for studies based on distance sampling methods when a fixed amount of sampling effort was allocated towards maximizing the number of sites by decreasing the number of repeat surveys within a year and increasing the interval between survey years. This pattern was independent of density or detectability of the target species, which contrasts with occupancy studies where the optimal allocation of effort to

detect trends depends largely on attributes of the target species. Specifically, when detectability is high or occupancy is low (or both), efficiency is higher when proportionally more effort is allocated to sites than to repeat surveys; in contrast, when detectability is low or occupancy is high (or both), efficiency is higher when proportionally more effort is allocated to repeat surveys rather than to sites (Field et al., 2005; MacKenzie et al., 2005; Steidl et al., 2013). For distance sampling studies, we found that density of the target population had little effect on how to allocate survey effort more effectively, at least for the range of values we considered (e.g. the advantage of sites over repeat surveys was similar for $D = 0.01$ and 10 individuals/ha). The advantage of sites over repeat surveys becomes increasingly narrow, however, as detectability decreases. Nevertheless, this advantage held even for species that are highly challenging to detect (e.g. for the lowest values we considered, where $\phi = 0.3$ and $\sigma = 30$ m).

These differences in how to allocate effort most efficiently between sites and repeat surveys may be due in part to differences in the effort required to estimate detection probability (p) between occupancy and distance sampling frameworks. For most occupancy designs, estimating p requires two or more surveys per site per primary period, whereas p can be estimated with data from one survey when using distance sampling (Buckland et al., 2001). The temporary emigration models we used for analysis separate the detection process into two components, availability and detection probability (Chandler et al., 2011). Availability is estimated from detection–nondetection data when a site is surveyed multiple times, similar to estimating p in an occupancy framework. In contrast, detection probability is estimated from distance sampling data, which requires only one survey. Although multiple surveys can improve precision of parameter estimates in both frameworks, the ability to account for some variation in the detection process with single surveys may explain why optimal allocation shifts towards surveying sites and away from repeat surveys in frameworks incorporating distance sampling.

Although allocating survey effort towards the number of sites maximized power to detect trends in abundance for distance sampling data analysed in a temporary–emigration framework, this result may not hold in all situations. For example, when abundances on survey sites are spatially independent, optimal allocation may favour maximising the number of sites surveyed; when spatial correlation is high and temporal correlation is low, however, optimal allocation could shift towards increasing the number of repeat surveys (Rhodes & Jonzén, 2011). Similarly, if animal activity is highly variable throughout the day or across the survey season, one survey per site may yield estimates with low precision. In this situation, accounting for availability by including >1 repeat survey per site could improve precision of estimates, increasing power. Additionally, multiple repeat surveys would increase the ability of monitoring programmes to gather information on a broader range of species in regions where activity peaks of different species vary throughout the survey period. Similarly, the advantage of maximizing the number of sites surveyed by increasing the survey–year interval may not hold in systems characterized by especially high interannual variation in animal densities. One potential compromise is a rotating–panel design, where a

subset of sites are surveyed annually, but each individual site is surveyed only once every two or three years (Bailey, Hines, Nichols, & MacKenzie, 2007). This approach would allow effort to be allocated towards surveying more sites while providing data that may capture interannual variation in population attributes. Additionally, monitoring could be accomplished by employing fewer surveyors every year instead of a larger number every second or third year.

Although our focus was to explore the influence of population attributes and design features on the power of monitoring programmes to detect trends in abundance of terrestrial vertebrates, our findings are likely to be relevant to other taxa that occur within the range of densities we considered and may provide insight for distance sampling studies with alternative objectives. We note, however, that optimal sampling designs for objectives other than detecting long-term trends would reflect a different set of tradeoffs and would be different than those we identified. Similarly, we note that although the relative influence of different population attributes and design features are unlikely to be affected strongly by modest departures from the values we simulated, the absolute influence (i.e. the time required to detect a target trend at a specified power) is sensitive to specificities of the target system and species as well as the model used to analyse data.

5 | CONCLUSIONS

Ultimately, the most efficient sampling design will reflect patterns of abundance and detectability across the area of interest and require balancing the inherent set of trade-offs in temporal versus spatial replication when allocating survey effort (Bailey et al., 2007; Steidl et al., 2013). For monitoring programmes that seek to identify trends in abundance with distance sampling data, our results demonstrate the advantage of allocating effort towards maximizing the number of sites surveyed.

Although we provide general guidance for determining the number of sites required to detect a target trend reliably within a specified period of time for a range of populations that differ in their intrinsic attributes (Figures 1, 3, and 4), we stress the importance of programme-specific power analyses for determining adequate sampling effort. For monitoring programmes that target multiple species, survey effort should be established to meet objectives for species that are rarest and most challenging to survey effectively (Figure 2). This will ensure that power to detect trends for more common or conspicuous species will be high. As budgets for conservation decrease and long-term consequences of inaction increase, ensuring that monitoring programmes have sufficient power to meet objectives effectively and efficiently is increasingly important.

ACKNOWLEDGEMENTS

We appreciate the support provided by the Heritage Program of the Arizona Game and Fish Department and the input provided by Edwin Juarez and Troy Corman. This work also was made possible by funding from the U.S. Bureau of Land Management and an Audubon Apacheria Fellowship. Steven Archer, Judie Bronstein, R.

William Mannan, and two anonymous reviewers offered constructive feedback on earlier versions of the manuscript. We gratefully acknowledge the allocation of computer time from the UA Research Computing High Performance Computing (HPC) and High Throughput Computing (HTC) at the University of Arizona. The authors have no conflicts of interest.

AUTHORS' CONTRIBUTIONS

E.M.A. and R.J.S. conceived of the study; E.M.A. collected and analysed the data; E.M.A. and R.J.S. wrote the manuscript and gave final approval for publication.

ORCID

Erik M. Andersen  <https://orcid.org/0000-0002-2150-078X>

Robert J. Steidl  <https://orcid.org/0000-0001-7330-5380>

DATA AVAILABILITY STATEMENT

Data available via the OSF Digital Repository, <https://doi.org/10.17605/OSF.IO/REJQ2> (Andersen & Steidl, 2019).

REFERENCES

- Andersen, E. M. (2019). Effects of plant invasions on birds breeding in desert grasslands. Dissertation, University of Arizona, Tucson, AZ.
- Andersen, E. M., & Steidl, R. J. (2019). Data from: Power to detect trends in abundance within a distance sampling framework. *OSF Digital Repository*, <https://doi.org/10.17605/OSF.IO/REJQ2>
- Bailey, L. L., Hines, J. E., Nichols, J. D., & MacKenzie, D. I. (2007). Sampling design trade-offs in occupancy studies with imperfect detection: Examples and software. *Ecological Applications*, *17*(1), 281–290. [https://doi.org/10.1890/1051-0761\(2007\)017\[0281:SDTIOS\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2007)017[0281:SDTIOS]2.0.CO;2)
- Barata, I. M., Griffiths, R. A., & Ridout, M. S. (2017). The power of monitoring: Optimizing survey designs to detect occupancy changes in a rare amphibian population. *Scientific Reports*, *7*(1), 1–9. <https://doi.org/10.1038/s41598-017-16534-8>
- Boulinier, T., Nichols, J. D., Sauer, J. R., Hines, J. E., & Pollock, K. H. (1998). Estimating species richness: The importance of heterogeneity in species detectability. *Ecology*, *79*(3), 1018–1028. [https://doi.org/10.1890/0012-9658\(1998\)079\[1018:ESRTIO\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1998)079[1018:ESRTIO]2.0.CO;2)
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2001). *Introduction to distance sampling: Estimating abundance of biological populations*. New York, NY: Oxford University Press.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2004). *Advanced distance sampling*. New York, NY: Oxford University Press.
- Chandler, R. B., Royle, J. A., & King, D. I. (2011). Inference about density and temporary emigration in unmarked populations. *Ecology*, *92*(7), 1429–1435. <https://doi.org/10.1890/10-2433.1>
- Diggle, P. J., Burnham, K. P., Anderson, D. R., & Laake, J. L. (1980). Estimation of density from line transect sampling of biological populations. *Wildlife Monographs*, *72*, 3–202. <https://doi.org/10.2307/2981804>
- Ficetola, G. F., Romano, A., Salvidio, S., & Sindaco, R. (2017). Optimizing monitoring schemes to detect trends in abundance over broad scales. *Animal Conservation*, *21*(3), 221–231. <https://doi.org/10.1111/acv.12356>
- Field, S., Tyre, A., & Possingham, H. (2005). Optimizing allocation of monitoring effort under economic and observational constraints. *Journal of Wildlife Management*, *69*(3), 473–482. [https://doi.org/10.2193/0022-541X\(2005\)069](https://doi.org/10.2193/0022-541X(2005)069)

- Gálvez, N., Guillera-Arroita, G., Morgan, B. J. T., & Davies, Z. G. (2016). Cost-efficient effort allocation for camera-trap occupancy surveys of mammals. *Biological Conservation*, 204, 350–359. <https://doi.org/10.1016/j.biocon.2016.10.019>
- Gaston, K. J., Blackburn, T. I. M., Greenwood, J. D., Gregory, R. D., Quinn, M., & Lawton, J. H. (2000). Abundance-occupancy relationships. *Journal of Applied Ecology*, 37, 39–59. <https://doi.org/10.1046/j.1365-2664.2000.00485.x>
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Gerrodette, T. (1987). A power analysis for detecting trends. *Ecology*, 68(5), 1364–1372. <https://doi.org/10.2307/1939220>
- Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2012). Designing studies to detect differences in species occupancy: Power analysis under imperfect detection. *Methods in Ecology and Evolution*, 3(5), 860–869. <https://doi.org/10.1111/j.2041-210X.2012.00225.x>
- Hedley, S. L., & Buckland, S. T. (2004). Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics*, 9(2), 181–199. <https://doi.org/10.1198/1085711043578>
- Kellner, K. F. (2016). JAGSUI: A wrapper around rjags to streamline JAGS analyses. R package version 1.4.4. Retrieved from <https://cran.r-project.org/web/packages/jagsUI>
- Kery, M., & Royle, J. A. (2015). *Applied hierarchical modeling in ecology: Analysis of distribution, abundance and species richness in R and BUGS*. London, UK: Elsevier.
- Kery, M., & Schmid, H. (2004). Monitoring programs need to take into account imperfect species detectability. *Basic and Applied Ecology*, 5(1), 65–73. <https://doi.org/10.1078/1439-1791-00194>
- Legg, C. J., & Nagy, L. (2006). Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management*, 78(2), 194–199. <https://doi.org/10.1016/j.jenvman.2005.04.016>
- Lele, S. R., Moreno, M., & Bayne, E. (2012). Dealing with detection error in site occupancy surveys: What can we do with a single survey? *Journal of Plant Ecology*, 5(1), 22–31. <https://doi.org/10.1093/jpe/rtr042>
- Lindemayer, D. B., & Likens, G. E. (2010). The science and application of ecological monitoring. *Biological Conservation*, 143(6), 1317–1328. <https://doi.org/10.1016/j.biocon.2010.02.013>
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8), 2248–2255. [https://doi.org/10.1890/0012-9658\(2002\)083\[2248:ESORWD\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2)
- MacKenzie, D. I., Nichols, J. D., Sutton, N., Kawanishi, K., & Bailey, L. L. (2005). Improving inferences in population studies of rare species that are detected imperfectly. *Ecology*, 86(5), 1101–1113. <https://doi.org/10.1890/04-1060>
- MacKenzie, D. I., & Royle, J. A. (2005). Designing occupancy studies: General advice and allocating survey effort. *Journal of Applied Ecology*, 42(6), 1105–1114. <https://doi.org/10.1111/j.1365-2664.2005.01098.x>
- Marques, T. A., Thomas, L., Kéry, M., Buckland, S. T., Borchers, D. L., Rexstad, E., ... Camp, R. J. (2017). Model-based approaches to deal with detectability: A comment on Hutto (2016a). *Ecological Applications*, 27(5), 1694–1698. <https://doi.org/10.1002/eap.1553>
- Nichols, J. D., & Williams, B. K. (2006). Monitoring for conservation. *Trends in Ecology & Evolution*, 21(12), 668–673. <https://doi.org/10.1016/j.tree.2006.08.007>
- Nuno, A., Milner-Gulland, E. J., & Bunnefeld, N. (2015). Detecting abundance trends under uncertainty: The influence of budget, observation error and environmental change. *Animal Conservation*, 18(4), 331–340. <https://doi.org/10.1111/acv.12175>
- Peach, M. A., Cohen, J. B., & Frair, J. L. (2017). Single-visit dynamic occupancy models: An approach to account for imperfect detection with Atlas data. *Journal of Applied Ecology*, 54(6), 2033–2042. <https://doi.org/10.1111/1365-2664.12925>
- Plummer, M. (2018). JAGS: just another Gibbs sampler. version 4.2.0. Retrieved from <http://sourceforge.net/projects/mcmc-jags/files>.
- Pollock, K. H. (1982). A capture-recapture sampling design robust to unequal probability of capture. *The Journal of Wildlife Management*, 46(3), 752–757. <https://doi.org/10.2307/3808568>
- Reynolds, J. H., Thompson, W. L., & Russell, B. (2011). Planning for success: Identifying effective and efficient survey designs for monitoring. *Biological Conservation*, 144(5), 1278–1284. <https://doi.org/10.1016/j.biocon.2010.12.002>
- Rhodes, J. R., & Jonzén, N. (2011). Monitoring temporal trends in spatially structured populations: How should sampling effort be allocated between space and time? *Ecography*, 34(6), 1040–1048. <https://doi.org/10.1111/j.1600-0587.2011.06370.x>
- Riddle, J. D., Stanislav, S. J., Pollock, K. H., Moorman, C. E., & Perkins, F. S. (2010). Separating components of the detection process with combined methods: An example with northern bobwhite. *Journal of Wildlife Management*, 74(6), 1319–1325. <https://doi.org/10.2307/40801127>
- Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60(1), 108–115. <https://doi.org/10.1111/j.0006-341X.2004.00142.x>
- Royle, J. A., Dawson, D. K., & Bates, S. (2004). Modelling abundance effects in distance sampling. *Ecology*, 85(6), 1591–1597. <https://doi.org/10.1890/03-3127>
- Sanderlin, J. S., Block, W. M., & Ganey, J. L. (2014). Optimizing study design for multi-species avian monitoring programmes. *Journal of Applied Ecology*, 51(4), 860–870. <https://doi.org/10.1111/1365-2664.12252>
- Schmidt, J. H., McIntyre, C. L., & MacCluskie, M. C. (2013). Accounting for incomplete detection: What are we estimating and how might it affect long-term passerine monitoring programs? *Biological Conservation*, 160, 130–139. <https://doi.org/10.1016/j.biocon.2013.01.007>
- Shrader-Frechette, K. S., & McCoy, E. D. (1992). Statistics, costs and rationality in ecological inference. *Trends in Ecology and Evolution*, 7(3), 96–99. [https://doi.org/10.1016/0169-5347\(92\)90249-B](https://doi.org/10.1016/0169-5347(92)90249-B)
- Sillett, S. T., Chandler, R. B., Royle, A. J., Kery, M., & Morrison, S. A. (2012). Hierarchical distance sampling models to estimate population size and habitat-specific abundance of an island endemic. *Ecological Applications*, 22(7), 1997–2006. <https://doi.org/10.1890/11-1400.1>
- Steidl, R. J., Conway, C. J., & Litt, A. R. (2013). Power to detect trends in abundance of secretive marsh birds: Effects of species traits and sampling effort. *Journal of Wildlife Management*, 77(3), 445–453. <https://doi.org/10.1002/jwmg.505>
- Steidl, R. J., Hayes, J. P., & Schaubert, E. (1997). Statistical power analysis in wildlife research. *The Journal of Wildlife Management*, 61(2), 270. <https://doi.org/10.2307/3802582>
- Steidl, R. J., & Thomas, L. (2001). Power analysis and experimental design. In S. M. Scheiner & J. Gurevitch (Eds.), *Design and analysis of ecological experiments* (pp. 14–36). New York, NY: Oxford University Press.
- Wintle, B. A., Runge, M. C., & Bekessy, S. A. (2010). Allocating monitoring effort in the face of unknown unknowns. *Ecology Letters*, 13(11), 1325–1337. <https://doi.org/10.1111/j.1461-0248.2010.01514.x>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Andersen EM, Steidl RJ. Power to detect trends in abundance within a distance sampling framework. *J Appl Ecol*. 2019;00:1–10. <https://doi.org/10.1111/1365-2664.13529>