Estimating population impacts via dynamic occupancy analysis of Before–After Control–Impact studies

4 authors, including:

Viorel D Popescu
Ohio University

100 PUBLICATIONS **1,489** CITATIONS

Perry de Valpine
University of California, Berkeley

109 PUBLICATIONS **3,965** CITATIONS

Douglas Tempel
University of Wisconsin–Madison

33 PUBLICATIONS **725** CITATIONS

Some of the authors of this publication are also working on these related projects:

Extinction Dynamics and Environmental Change View project

Is spread of invasive species regulated? Using ecological theory to interpret statistical analysis View project

Estimating population impacts via dynamic occupancy analysis of Before–After Control–Impact studies

Viorel D. Popescu,[1,2,5] Perry de Valpine,[1] Douglas Tempel,[3] and M. Zachariah Peery[4]

[1]*Department of Environmental Science, Policy and Management, University of California Berkeley, 130 Mulford Hall #3114, Berkeley, California 94720-3114 USA*
[2]*Centre for Environmental Research (CCMESI), University of Bucharest, 1 N. Balcescu Blvd., 010041 Bucharest, Romania*
[3]*Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, 200 Hodson Hall, 1980 Folwell Avenue, St. Paul, Minnesota 55108 USA*
[4]*Department of Forest and Wildlife Ecology, University of Wisconsin–Madison, A233 Russell Labs, 1630 Linden Drive, Madison, Wisconsin 53706 USA*

Abstract. Estimating environmental impacts on populations is one of the main goals of wildlife monitoring programs, which are often conducted in conjunction with management actions or following natural disturbances. In this study we investigate the statistical power of dynamic occupancy models to detect changes in local survival and colonization from detection–nondetection data, while accounting for imperfect detection probability, in a Before–After Control–Impact (BACI) framework. We simulated impacts on local survival and/or detection probabilities, and asked questions related to: (1) costs and benefits of different analysis models, (2) confounding changes in detection with changes in local survival, (3) sampling design trade-offs, and (4) species with low vs. high rates of turnover. Estimating seasonal effects on local survival and colonization, as opposed to estimating Before–After effects, had little effect on the power to detect changes in local survival. Estimating a parameter that accounted for pretreatment differences in local survival between Control and Impact sites decreased power by 50%, but it was critical to include when such differences existed. When the experimental treatment had a negative impact on species detectability but analysis assumed constant detection, the Type I error rates were dramatically inflated (0.20–0.33). In general, there was low power (<0.5) to detect a 50% decrease in local survival for all combinations of sites ($N = 50$ vs. 100), seasons sampled (8 vs. 12), and visits per site per season (4 vs. 6). Unbalanced designs performed worse than balanced designs, with the exception of the case of treatments being implemented in different seasons at different sites. Adding more control sites improved the ability to detect changes in local survival. Surveying more seasons after impact resulted in modest power gains, but at least three seasons before impact were required to successfully implement BACI occupancy studies. Turnover rates had a low impact on power. Occupancy studies conducted in a BACI design offer the opportunity to detect environmental impacts on wildlife populations without the costs of intensive studies. However, given the low power to detect small changes (20%) in local survival, these studies should be used when researchers are confident that major treatment impacts will occur or very large sample sizes are obtainable.

Key words: BACI; colonization; detection; dynamic occupancy; environmental impact; local survival; sampling design; simulation; statistical power; turnover; wildlife monitoring.

Introduction

Detecting the effects of anthropogenic disturbance on animal and plant populations is the overarching goal of many environmental monitoring programs. Data from wildlife surveys pose special challenges because of imperfect observations of animal populations and lack of understanding of processes driving population dynamics (Williams 2001). Detection–nondetection data can be analyzed using occupancy models that incorpo-

rate imperfect detection and covariates that might affect site occupancy dynamics (MacKenzie et al. 2002, 2003, Tyre et al. 2003). These models have been extended to include community-level (Dorazio and Royle 2005, Kéry and Royle 2008) and multi-state approaches (Nichols et al. 2007, Royle and Kéry 2007, MacKenzie et al. 2009). Thus, researchers now have the opportunity to implement robust, broadscale studies and monitoring programs without the costs associated with data-intensive mark–recapture studies (DeWan and Zipkin 2010; for examples see Muths et al. 2005, Joseph et al. 2006, Collier et al. 2010). Along with detecting changes in occupancy, these models can be used to assess habitat suitability and population dynamics, and to estimate

abundance and/or demographic parameters (Martin et al. 2009, 2010, MacKenzie et al. 2010, Roth and Amrhein 2010, Risk et al. 2011). In particular, dynamic occupancy modeling allows the detection of population trends via estimation of extinction and colonization rates (Hanski 1994, 1998, Moilanen 1999) based on robust sampling design (repeated visits to each site each year or season; Kendall et al. 1997, MacKenzie et al. 2003) for closed populations (Dail and Madsen 2011).

Many wildlife monitoring programs are conducted, by design or coincidence, in conjunction with management actions (treatments or impacts) at some sites (Marsh and Trenham 2008). Researchers then evaluate treatment impacts by assessing two contrasts in the data: before vs. after measurements in the treatment sites, and control sites vs. treatment sites. Designs with a before–after factor crossed with a control–treatment factor are known as Before–After Control–Impact (BACI) (Green 1979, Stewart-Oaten et al. 1986, Underwood 1991, Osenberg and Schmitt 1996). Traditional BACI designs range from measurements taken at one impacted site over time to more complicated, asymmetric (beyond BACI) designs that ensure adequate spatial replication (Underwood 1993). BACI analysis approaches include linear and generalized linear mixed models (McDonald et al. 2000), with a significant treatment (Control–Impact) × time (Before–After) interaction indicating that the experimental treatment truly has had an effect on the Impact sites. Considering the interaction term in this manner allows treatment impacts to be distinguished from background time effects shared by all sites and from background differences between control and treatment sites.

Integrating BACI study designs into dynamic occupancy analysis offers a potentially general approach for determining management (or other) effects on wildlife. However, dynamic occupancy studies involve estimation of extinction, colonization, and detection rates, as well as initial occupancy probability, each of which may depend on BACI factors. In addition, occupancy studies may be conducted over many seasons, so that modeling seasonal effects beyond simply "before" and "after" may be necessary. Given these complications, consideration of statistical power and optimal study design is crucial to evaluate the costs and benefits of such an approach. The statistical power of monitoring programs to detect population trends has been a long-term focus for environmental research (Gerrodette 1987, Fairweather 1991, Kendall et al. 1992, Taylor and Gerrodette 1993), but the lack of power is still a major weakness of wildlife monitoring programs (Gibbs et al. 1998). In particular, simple presence–absence data have been shown to have low power (i.e., to require large sample sizes) for detecting population changes (Strayer 1999, Pollock 2006; but see Zielinski and Stauffer 1996), although incorporating imperfect detection can improve statistical power (Rhodes et al. 2006). How these results will translate to BACI occupancy scenarios is unclear.

Because BACI and dynamic occupancy approaches each raise a host of study design questions, their combination is particularly complicated. For example, design of occupancy studies involves trade-offs among the number of sites, number of seasons, and number of surveys per site per season (MacKenzie and Royle 2005, Bailey et al. 2007, Mattfeldt et al. 2009). BACI design and analysis issues include the amount of spatial and temporal replication and the treatment of spatial and temporal variation in the statistical model (Conquest 2000, Benedetti-Cecchi 2001, Hewitt et al. 2001). Widespread application of BACI approaches in ecology was not always accompanied by the implementation of designs with sufficient temporal and spatial replication (Stewart-Oaten and Bence 2001). As such, BACI experiments were found to overstate the significance of the association between the impact and ecological response, whereas any attempts to account for temporal and spatial sources of variability resulted in overall low power (Murtaugh 2002).

However, formal BACI analysis previously has not been built into dynamic occupancy models, with the exception of Russell et al. (2009). Here we present a more general framework using frequentist methods, and systematically investigate model performance and study design questions. We allow that treatments may affect extinction or colonization, and/or detection probabilities, in contrast to Russell et al. (2009), who assumed similar treatment effects on extinction and colonization. We also allow the possibility of background differences between Control and Impact sites, which were omitted by Russell et al. (2009), and between times, which were included only for detection probabilities by Russell et al. (2009). We take advantage of recently developed software to fit the dynamic occupancy models of MacKenzie et al. (2003), specifically the package *unmarked* 0.9-2 (Fiske and Chandler 2011) for program R 2.13 (R Development Core Team 2011).

In this paper we introduce a general approach to BACI occupancy analysis in the context of detecting changes caused by human impacts to the presence–absence of species by directly affecting extinction and/or colonization. To systematically simulate design trade-offs while avoiding an unwieldy combination of possible scenarios, we established a baseline scenario and addressed specific questions using various levels and combinations of model parameters. In the baseline scenario, we simulated impacts on site extinction only, but included possible impacts on both extinction and colonization in the analysis models. Estimating the impacts on extinction (or mortality) is one of the main goals of wildlife monitoring programs and it is deemed critical for determining population trends (Saracco et al. 2008). We only use scenarios where the treatment impact is detrimental rather than beneficial (i.e., a restoration treatment), because they should have largely similar design issues (within the range of impacts used in our study). Specifically, we ask the following questions. (1)

What are the costs and benefits of modeling time as a Before–After factor vs. as a seasonal effect, and of accounting for potential background differences between Control and Impact sites? (2) What is the potential for treatment effects on detection probabilities to be confounded with effects on extinction or colonization rates? (3) If the analysis model is reasonable, what are the power trade-offs involved in choosing numbers of control and treatment sites, surveys per season, and number of seasons before and after treatment? (4) How do low-turnover vs. high-turnover dynamics affect power?

METHODS

Occupancy modeling

MacKenzie et al. (2002) proposed a single-season occupancy model that accounts for imperfect detection when estimating site occupancy. MacKenzie et al. (2003) then extended the single-season model to dynamic multi-season models by introducing site colonization and extinction rates. In this model, occupancy dynamics can be described as a first-order Markov process, in which the probability that a site is occupied at time t depends on the state of the site at $t - 1$ (MacKenzie et al. 2006). Many natural processes, such as ecological succession, are characterized by higher-order Markov dynamics (Usher 1979), but incorporating this level of complexity was beyond the scope of this study. The colonization, extinction, and/or detection probabilities can depend on covariates, including dummy variables for categories, allowing flexible formulation of BACI analysis using dynamic occupancy models.

Specifically, let $z(i, j)$ be the true occupancy status of site i in season j. If the site is occupied, $z(i, j) = 1$, otherwise $z(i, j) = 0$. For the first season,

$$z(i, 1) \sim \text{Bernoulli}(\psi_i) \qquad (1)$$

where ψ_i is the probability of a site being initially occupied (in this study, fixed at equilibrium occupancy probability). The occupancy dynamics in the subsequent seasons ($t = 2, 3, \ldots$) are described by

$$z(i, t) \,|\, z(i, t - 1) \sim \text{Bernoulli}$$
$$\{ z(i, t - 1)(1 - \varepsilon_{i,t-1}) + [1 - z(i, t - 1)]\gamma_{i,t-1} \} \qquad (2)$$

where $\varepsilon_{i,t-1}$ and $\gamma_{i,t-1}$ are site- and time-specific probabilities of local extinction and colonization, respectively.

The observation process is conditional upon the state of the site and is modeled as

$$y(i, t) \,|\, z(i, t) \sim \text{Bernoulli}[z(i, t)p_t] \qquad (3)$$

where, p_t is the probability of detection at time t, given that the site is truly occupied. This state-space model contains two processes: (1) the unobserved ecological (or state) process, $z(i, t)$, and (2) the observation process, $y(i, t)$, which depends on the ecological process, where i

$= 1, 2, \ldots$, and $t = 1, 2, \ldots$ (Royle and Kéry 2007). The components of the observation process are not completely independent (i.e., if the site is truly unoccupied, then the successive observations of absences are non-independent), and similar concerns have been raised for individual-level detectability (Martin et al. 2011).

Implementing dynamic occupancy models in a BACI framework

With the general framework just outlined, it is possible to model occupancy at time $t = 1$ as well as detection, colonization, and extinction probabilities at time $t = 1, 2, \ldots$ as functions of measured covariates using a logit link function (MacKenzie et al. 2006). Note that colonization and extinction "at $t = 2$" refer to the rates affecting changes from $t = 2$ to $t = 3$, for example. Detection may be modeled using survey-specific, site-specific, and/or seasonal site-specific covariates (i.e., site covariates that vary on a seasonal basis), whereas occupancy, extinction, and colonization may be modeled using site- and seasonal site-specific covariates. In general, the model for one of these rates is

$$\theta_{i,t} = \frac{\exp[\mathbf{X}(i, t)\boldsymbol{\beta}]}{1 + \exp[\mathbf{X}(i, t)\boldsymbol{\beta}]} \qquad (4)$$

where $\theta_{i,t}$ represents $p_{i,t}$, $\varepsilon_{i,t}$, or $\gamma_{i,t}$; $\mathbf{X}(i, t)$ represents the combination of covariates used to model $p_{i,t}$, $\varepsilon_{i,t}$, or $\gamma_{i,t}$, and $\boldsymbol{\beta}$ is the column vector of the coefficients to be estimated.

For our BACI models, the $\mathbf{X}(i, t)$ vector comes from a set of categorical variables. First, we indicated Control and Impact sites with a site-specific variable, SiteTreat = "Control" and "Impact," respectively. Second, we indicated seasons that were either Before or After the treatment time with a seasonal site-specific variable, SeasonBA = "Before" or "After," respectively. Third, we allowed colonization and extinction rates to vary across each season by using a different seasonal site-specific variable, Season = "1", "2," ... "N". The SiteTreat and SeasonBA factors had only two levels, whereas the Season variable had as many levels as there were seasons in the study. Each categorical variable is turned into $\mathbf{X}(i, t)$ entries in a standard way by one or more dummy variables, such as 0 and 1 to distinguish two levels. Note that our models do not really need the logit link, because they include only categorical variables, but we formulate our models in this general way to allow the matrix notation and to accommodate continuous variables in real studies.

We denoted models according to the combination of variables used for detection, colonization, and extinction probabilities (Table 1), but did not explore explanatory variables for initial occupancy probability. For example, [γ(SiteTreat $\times$ SeasonBA), ε(SiteTreat $\times$ SeasonBA), $p(.)$] is a model in which colonization and extinction probabilities depend on SiteTreat, SeasonBA, and the SiteTreat $\times$ SeasonBA interaction, and detection probability is constant. This represents classic BACI concepts

TABLE 1. Null and alternative models for estimating the statistical power of occupancy BACI designs for detecting treatment effects on local survival (1 − extinction).

Question, generating data, and analysis	Null model	Alternative model
Q1: Seasonal vs. Before–After effects on γ and ε (with and without seasonal effects on γ and ε and without treatment effects on p)		
With pretreatment site differences	γ(SiteTreat+SeasonBA), ε(SiteTreat+SeasonBA), p(.)	γ(SiteTreat+SeasonBA+SeasonSiteTreat), ε(SiteTreat+SeasonBA+SeasonSiteTreat), p(.)
	γ(SiteTreat+Season), ε(SiteTreat+Season), p(.)†	γ(SiteTreat+Season+SeasonSiteTreat), ε(SiteTreat+Season+SeasonSiteTreat), p(.)†
Without pretreatment site differences	γ(SeasonBA), ε(SeasonBA), p(.)	γ(SeasonBA+SeasonSiteTreat), ε(SeasonBA+SeasonSiteTreat), p(.)
	γ(Season), ε(Season), p(.)†	γ(Season+SeasonSiteTreat), ε(Season+SeasonSiteTreat), p(.)†
Q2: Effects of treatment on p (with and without treatment effects on p and without seasonal effects on γ and ε)		
With treatment effect on p	γ(SiteTreat+Season), ε(SiteTreat+Season), p(SeasonSiteTreat)	γ(SiteTreat+Season+SeasonSiteTreat), ε(SiteTreat+Season+SeasonSiteTreat), p(SeasonSiteTreat)
Without treatment effect on p	γ(SiteTreat+Season), ε(SiteTreat+Season), p(.)	γ(SiteTreat+Season+SeasonSiteTreat), ε(SiteTreat+Season+SeasonSiteTreat), p(.)
Q3: Balanced and unbalanced BACI designs (without seasonal effects on γ and ε and without treatment effects on p)	γ(SiteTreat+SeasonBA), ε(SiteTreat+SeasonBA), p(.)	γ(SiteTreat×SeasonBA), ε(SiteTreat×SeasonBA), p(.)
Q3e: Impact occurs at different times at each treated site (without seasonal effects on γ and ε and without treatment effects on p)	γ(SiteTreat+Season), ε(SiteTreat+Season), p(.)	γ(SiteTreat+Season+SeasonSiteTreat), ε(SiteTreat+Season+SeasonSiteTreat), p(.)
Q4: Low vs. high turnover species (without seasonal effects on γ and ε and without treatment effects on p)	γ(SiteTreat+SeasonBA), ε(SiteTreat+SeasonBA), p(.)	γ(SiteTreat×SeasonBA), ε(SiteTreat×SeasonBA), p(.)

Note: We omitted the occupancy parameter (ψ) because we did not explore explanatory occupancy-related variables; γ is colonization, ε is extinction, and p is detection probability.

† This set of models was also used for examining the effects of background differences in local survival between Control and Impact sites.

applied to both colonization and extinction rates under the assumption that treatments do not affect detection. The main null hypothesis of no treatment effects corresponds to no interactions between SiteTreat and SeasonBA, so the corresponding null model is [γ(SiteTreat + SeasonBA), ε(SiteTreat + SeasonBA), p(.)], which has two fewer parameters (one for each interaction).

Finally, we indicated site–season combinations that represented Impact sites After treatment application with a seasonal site-specific variable, SeasonSiteTreat. This variable had level Control for the three subsets of the data set that did not receive a treatment (i.e., Control–Before, Control–After, and Impact–Before), and level Impact for the subset that received the treatment (i.e., Impact–After). This variable provides an equivalent way to set up a classic BACI model without an interaction term [γ(SiteTreat +

SeasonBA + SeasonSiteTreat), ε(SiteTreat + SeasonBA + SeasonSiteTreat), p(.)]. This model is equivalent to the model containing the Site × SeasonBA interaction term because the SeasonSiteTreat variable indicates Impact–After combinations of SiteTreat and SeasonBA (i.e., their interaction). Using this model notation will facilitate some useful comparisons.

Categorizing seasons as Before and After may be unrealistic for a long-term study with multiple seasons before and/or after treatment application, due to inherent variation between seasons. Furthermore, in some studies treatments may be applied in different seasons to different sites, so a simple variable such as SeasonBA is inappropriate. The Season variable provides more realism by allowing an effect for each season, but it comes at the cost of estimating more parameters. Using Season instead of SeasonBA, a model assuming

constant detection probability would be [γ(SiteTreat + Season + SeasonSiteTreat), ϵ(SiteTreat + Season + SeasonSiteTreat), $p(.)$], which would be compared to the null model [γ(SiteTreat + Season), ϵ(SiteTreat + Season), $p(.)$].

Likelihood ratio tests

Because maximum likelihood methods are used to estimate the parameters of interest, likelihood ratio tests (LRT) are useful for testing hypotheses regarding the relative level of support for certain parameters using nested models (MacKenzie et al. 2006, Royle and Dorazio 2008). Nesting means that a less general model (associated with the null hypothesis H_0) can be obtained by restricting some parameters of a more complex model (associated with the alternative hypothesis H_a). The likelihood ratio is

$$\Lambda = \frac{L(\theta_0 \,|\, y)}{L(\theta_a \,|\, y)}. \qquad (5)$$

The likelihood ratio statistic is then computed as $-2\log(\Lambda)$, and if the null is true, it asymptotically follows a chi-squared distribution, $\chi^2(\nu)$, where degrees of freedom ν is the difference in the number of parameters to be estimated under H_a and H_0 (Royle and Dorazio 2008). If H_0 rejection occurred, we concluded that the more general model better described the ecological process, and thus the additional single or interaction effects had a significant impact on estimating extinction and/or colonization. In all of our tests (Table 1), H_0 represented no treatment impact for extinction or colonization, whereas H_a allowed treatment impacts, so ν was always 2.

Simulations and data analysis

We used program R 2.13 (R Development Core Team 2011) for both data simulations and statistical analyses. We simulated multi-season presence–absence data sets following a robust sampling design with multiple visits per season (sensu Kendall et al. 1997), with actual detections simulated according to the detection probability (see Supplement). The data sets expressed site occupancy dynamics as a state process based on: (1) the probability of an occupied site continuing to be occupied from one season to the next, and (2) the probability of an unoccupied site becoming colonized. We simulated the probability that a species that occupied a site at time t continues to occupy the site time $t + 1$ using the probability of local survival [i.e., local survival (φ) = 1 − extinction (ϵ); Royle and Kéry 2007]. In this study, we are using the terms local mortality and extinction interchangeably; i.e., we simulate impacts and provide results in the context of incremental declines in local survival, but refer to extinction in the context of analysis models according to the occupancy terminology of MacKenzie et al. (2003). The simulations rely on several critical assumptions: (1) detections occur independently at sites, (2) occupancy and detection probabilities are

similar across sites and time unless they can be modeled using covariates, (3) the species is correctly identified (i.e., no false-positive observations), and (4) no changes in site occupancy occur within a season ("closure assumption"). Failure to meet these assumptions could bias parameter estimates (Bailey et al. 2007). The violation of the last assumption can lead to overestimating or underestimating the probability of occurrence, so special attention must be paid to correctly defining the "season" based on the ecology of the study organism during both sampling design and data analysis (Rota et al. 2009, Dail and Madsen 2011). If a violation of the closure assumption is suspected (i.e., "open" populations), then explicit estimations of immigration/emigration rates can be obtained using the models proposed by Dail and Madsen (2011) and implemented in the R package *unmarked* (function *pcountOpen*; Fiske and Chandler 2011).

Because occupancy dynamics are expressed as a linear first-order Markov process, in which occupancy state in season t is conditional upon the occupancy state in season $t - 1$, we modeled the occupancy history assuming a true equilibrium occupancy state in season 1, $\psi_1 = 0.66$ (i.e., stationarity). This assumption draws on metapopulation dynamics theory (Hanski 1994), which states that if ψ_i is the stationary probability that site i is occupied, then this probability can be expressed as

$$\psi_i = \frac{\gamma_i}{\gamma_i + \epsilon_i} \qquad (6)$$

where ϵ_i and γ_i are site-specific probabilities of local extinction and colonization, respectively; in our particular case, $\epsilon_1 = 0.2$ and $\gamma_1 = 0.4$.

Because we are jointly testing for changes in extinction and colonization rates, and these two parameters alone determine equilibrium occupancy, we are in effect testing for changes in equilibrium occupancy. Conditional on the true site occupancy state, we simulated the observed occupancy using the probability of detection, p (Royle and Kéry 2007). We simulated 300 data sets for each question and scenario to be described. We considered 300 simulations per scenario for two reasons. First, the processing time for the more complicated models was long; for example, simulating data with the four levels of survival [$300 \times 4 = 1200$ cases], and running models that estimated parameters for each Season, with and without variable SiteTreat for each of the 1200 cases, took >40 h on a 2-GHz AMD Opteron 246 processor (Penguin Computing, Fremont, California, USA). Second, the standard deviations (SD) of our simulated estimates of the Type I error and power values were within acceptable limits; for example, we were able to estimate the Type I error rate (0.05) with SD = 0.013, and a power value of 0.5 with SD = 0.028.

We performed all statistical analyses using package *unmarked* 0.9-2 (Fiske and Chandler 2011), which fits hierarchical models of animal occurrence and abun-

Table 2. Input parameters for the baseline simulation scenario and extensions outlining effort trade-offs for dynamic occupancy study design (eight possible combinations of low/high values).

Parameter	Abbreviation	Level	Value
Total number of sites†	Nsites	low	50
		high	100
Number of visits/site/season	Nvisits	low	4
		high	6
Detection probability‡	p	low	0.4
		high	0.8
Total number of seasons§	Nseasons		8
Colonization, Control¶	γ_C		0.4
Colonization, Impact¶	γ_I		0.4
Local survival = $1 -$ extinction, Control¶,#	$\varphi_C = 1 - \varepsilon_C$		0.8
Local survival = $1 -$ extinction, Impact#	$\varphi_I = 1 - \varepsilon_I$		$\varphi_C \times [0.2, 0.5, 0.8, 1.0]$
Initial site occupancy	ψ_1		0.66

† Except for Question 3c (adding more Controls, when there are a limited number of Impact sites).
‡ Except for Question 2 (baseline $p = 0.6$).
§ We simulated 12-season data sets for the baseline scenario (Question 3a) only.
¶ Except for Question 4 (γ_C, γ_I, and ε_C were multiplied by 0.5 and 1.5 to simulate low and high turnover species, respectively).
# Except for Question 1, local survival was simulated using low and high seasonal effects by using an additive term on a logit scale, drawn from a normal distribution of mean = 0 and standard deviation = 0.2, and 0.4, respectively).

dance to data from survey methods that are subject to imperfect detection, such as occupancy sampling (detected/undetected), distance sampling, and point counts. Specifically, we used the *colext* function, which fits the dynamic occupancy model formulated by MacKenzie et al. (2003) to detection–nondetection data. In *colext*, first-season occupancy can be modeled using site-specific covariates that do not vary across seasons (e.g., SiteTreat), whereas colonization and extinction can be modeled using site-specific and seasonal site-specific covariates (e.g., SeasonBA, Season, SeasonTreatSite). Detection can be modeled using both site-specific and seasonal site-specific covariates, as well as survey covariates specific to each sampling occasion (not used in our analysis).

Baseline scenario

We started by simulating a baseline scenario in which the population response to a hypothetical experimental treatment (i.e., impact) was a decrease in local survival (equivalent to an increase in extinction) at the Impact sites during the After period compared to Control conditions (Impact sites are also in Control conditions prior to the experimental treatment). We considered 20%, 50%, and 80% decreases in local survival from the baseline constant condition ($\varphi_C = 0.8$, $\varphi_I = 0.64$, 0.4, and 0.16, respectively; Table 2). In addition, we simulated data sets with no change in local survival ($\varphi_C = \varphi_I = 0.8$), which allowed us to investigate whether or not the nominal Type I error was within an acceptable range ($\sim$0.05). For simplicity, we simulated the data sets with constant colonization ($\gamma_C = \gamma_I = 0.4$) and detection probability ($p_C = p_I$; Table 2) and no seasonal effects on local survival, colonization, and detection probabilities.

The baseline scenario included 25 each of Control and Impact sites, and was simulated across eight seasons, divided equally between the Before and After periods, with four visits per site per season. The sampling was

assumed to occur each season, resulting in three time steps before impacts and four time steps after. We extended the baseline scenario by considering combinations of low and high levels for (1) the total number of sites (Nsites), (2) the number of visits per site per season (Nvisits), and (3) detection probability (p; Table 2), outlining trade-offs in sampling effort for designing dynamic occupancy studies. This approach yielded eight combinations of low/high levels; however, the combination high Nvisits (6)/high p (0.8) was unnecessary because the combination low Nvisits (4)/high p (0.8) already yielded a cumulative detection probability of 0.998, enough for detecting the species if it was indeed present. The choice of baseline values for detection, colonization, and local survival (Table 1) was based on several criteria. First, these values are commonly used in simulation studies for exploring trade-offs in occupancy and metapopulation dynamics studies (e.g., Vuilleumier and Possingham 2006, Mattfeldt et al. 2009, Archaux et al. 2012). Second, we selected high detection and local survival probabilities because we wanted the simulated changes in these parameters to encompass a wide range of values (e.g., $p = 0.3$–0.8, $\varphi = 0.16$–0.8), representing small to severe impacts, thus increasing our ability to provide general recommendations for analysis and study design. Third, these values loosely reflected occupancy dynamics of long-lived territorial species, such as California Spotted Owls (Blakesley et al. 2010). The caveat for using these baseline parameters is that the results of our simulations are not likely to be applicable for rare or cryptic species with detection probabilities of $p < 0.3$, which are also commonly the focus of monitoring studies. Further simulations pertaining to our questions of interest (see *Introduction*) incorporated these extensions (except for Question 2, which investigated confounding effects of Impact-induced changes in detection probability on detecting changes in local survival).

Scenarios exploring seasonal effects on local survival

For Question 1, we evaluated the costs and benefits of using only a Before–After category (i.e., SeasonBA) to model temporal changes in extinction or colonization. This was done by simulating data with and without seasonal effects on local survival and testing for treatment effects with each model (i.e., testing whether SeasonSiteTreat is significant; Table 1). For data simulated without seasonal effects, both the SeasonBA and Season models should be valid for analysis, so comparing their power shows the cost of using more parameters in the Season model. For data simulated with seasonal effects, models that use SeasonBA (instead of Season) are "wrong" for analysis. In general, this would be revealed by inaccurate Type I error rate, i.e., more or fewer Type I errors than the intended Type I error rate of 0.05. We introduced low and high seasonal effects on local survival by using an additive term on a logit scale, drawn from a normal distribution of mean = 0 and standard deviation = 0.2 or 0.4. Although we simulated the seasonal effects randomly, the analysis models handle them as fixed effects, similar to a great deal of mark–recapture modeling (e.g., Kendall et al. 1997, Nichols et al. 1998, Blakesley et al. 2010).

In addition, we explored the statistical power trade-offs of including or ignoring background differences between Control and Impact sites. To do so, we simulated data with and without differences in local survival between Control and Impact sites prior to the experimental treatments, and fitted models with and without the fixed effect SiteTreat for both colonization and extinction (Table 1). For data simulated without background differences, we looked only at the cost of estimating the extra parameter SiteTreat in terms of statistical power. For data simulated with background differences, models containing variable SiteTreat should be valid for analysis, while models without variable SiteTreat are "wrong" and should result in biased parameter estimates and incorrect Type I error rates. We simulated pretreatment background differences by assigning high local survival (0.8) to Control sites and low local survival (0.6 and 0.4) to Impact sites, and vice versa. Similar to the other scenarios, we further considered three levels of decrease in local survival due to an experimental treatment at the Impact sites (20%, 50%, and 80%). When the Impact sites started at low local survival compared to Control sites, we investigated whether we were able detect a further decrease in local survival (when variable SiteTreat was included). All sites were simulated starting from their equilibrium occupancy probability.

Scenarios exploring treatment effects on detection

For Question 2, we investigated the potential confounding of treatment effects on detection probability with effects on colonization or local survival. We simulated data sets with and without an Impact effect on the probability of detection, and then tested for effects on colonization or extinction using models that omit or allow possible effects on detection (Table 1). When data are simulated with treatment effects on detection changes, we can see if any decrease in power occurs by needing to accommodate detection changes in the analysis. When data are simulated with treatment effects on detection changes and analyzed without them (the "wrong" model), we can see if this leads to incorrect Type I error rates. We considered a baseline detection probability of $p = 0.6$, four visits per site per season, and assumed that the Impact had a positive effect (+0.3 on the probability scale), a negative effect (–0.3), or no effect on detection. For this question we omitted seasonal effects from the simulations for simplicity, but we included Season terms in the analysis models.

Scenarios exploring trade-offs in sampling design

For Question 3, we investigated the influence of sampling design on the power to detect changes in colonization or local survival at Impact sites during the After period by varying the sampling effort (number of sites, number of visits per site per season, and number of seasons), as well as the probability of detection (Table 2). We simulated the data sets with constant detection probability and no seasonal effects on local survival, colonization, and detection probabilities (i.e., baseline scenario). To explore design trade-offs, we conducted five separate comparisons, in each of which we varied the magnitude of the treatment effect on local survival (Table 2).

1) Question 3a: Balanced (basic) occupancy BACI design, in which the number of seasons and sites was equally divided between Before–After and Control–Impact, respectively. This is the baseline scenario with the high and low sampling effort extensions previously described, and the only additional scenario, which was simulated across 12 seasons (high Nseason).

2) Question 3b: Unbalanced occupancy BACI design in which the Control and Impact sites were unequally represented; as such, we varied the percentage of sites that fell in either category while keeping the total number of sites constant (30% Control and 70% Impact and vice versa, rounding the values to the nearest integer).

3) Question 3c: Unbalanced design in which the number of potential Impact sites is outside of the researcher's control. We explored the benefits (in terms of statistical power) of adding more Control sites: four times the number of Control sites (using 25 reference Impact sites), and two times the number of Control sites (using 50 reference Impact sites).

4) Question 3d: Unbalanced design in which the seasons (Before and After) were unequally represented. We varied the proportion of seasons that fell in either category while keeping the total number of seasons constant ($N = 8$). We considered one scenario for
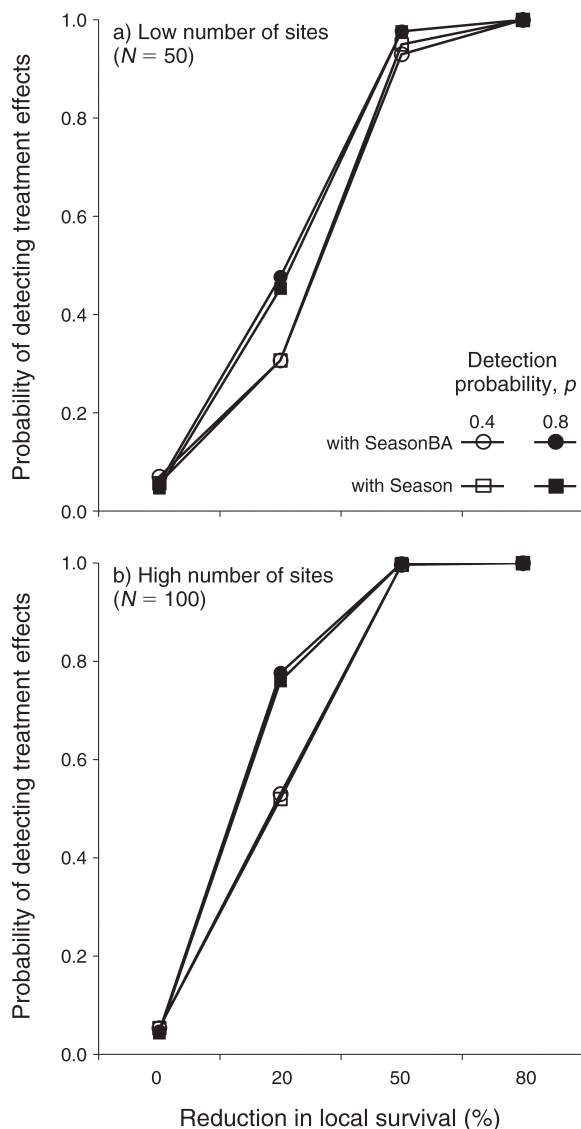
FIG. 1. Probability of detecting treatment effects for occupancy BACI designs when data were simulated without seasonal effects on local survival and were analyzed with Before–After effects (SeasonBA) or seasonal effects (Season) for (a) low and (b) high numbers of sites sampled. Two detection probability (p) levels were tested: low (0.4) and high (0.8). Type I error rates occur for 0% reduction in survival.

fewer seasons After ($N = 2$), and two scenarios for fewer seasons Before ($N = 2$ and $N = 3$).

5) Question 3e: Unbalanced design in which the simulated experimental treatment is implemented during different seasons at each Impact site. This scenario is likely to better represent the real-world situation in which large habitat manipulations, such as silvicultural treatments, cannot be applied simultaneously across a large number of sites. We assumed an equal number of Control and Impact sites and randomly varied the timing of treatment implemen-

tation between seasons 4 and 7 (i.e., some Impact sites are sampled only one season posttreatment).

Scenarios exploring effects of low vs. high turnover

For Question 4, we used the basic parameters described in Table 2 (seasons and sites equally divided between Before–After and Control–Impact, respectively), but varied the probability of colonization and extinction to simulate low and high fluctuations (turnover) in site occupancy by multiplying both parameters by 0.5 and 1.5, respectively. Multiplying both parameters by the same amount allowed us to compare the results to the basic scenario while keeping the stationary occupancy probability constant (i.e., $\psi_1 = 0.66$).

For each analysis we then used LRT's to test for the relative support of the extra parameters in the more general alternative models (Table 1). In our case, the extra parameters always involved testing the effect of the experimental treatment jointly on extinction and/or colonization, as described in section *Implementing dynamic occupancy models in a BACI framework* (i.e., SeasonSiteTreat for the additive models fitted for Questions 1 and 2, and SiteTreat × SeasonBA for the models fitted for Questions 3 [except 3e] and 4). By repeating the entire procedure 300 times (i.e., data set simulations, model fitting under H_a and H_0, and computing LRT's), we were able to approximate statistical power as the proportion of the repetitions in which we rejected H_0, assuming $\alpha = 0.05$ (Gibbs et al. 1998). In general, the ecological literature agrees that a power value >0.8 (or ideally >0.9) is indicative of high performance to detect environmental changes or population trends; values <0.8 are less desirable and reflect a lower performance (Gibbs et al. 1998).

RESULTS

Modeling seasonal effects on local survival

Estimating extra parameters associated with seasonal effects on local survival and colonization had no discernible cost in terms of statistical power when the data sets were simulated without random effects on local survival (Fig. 1; Appendix). Contrary to our expectations, there was virtually no penalty when using the "wrong" model containing Before–After effects (i.e., SeasonBA) to analyze data containing seasonal effects on local survival (the Type I error did not deviate from the nominal value of 0.05).

Effects of estimating background differences in local survival

Adding site-specific parameters SiteTreat when no background differences exist between Control and Impact sites prior to the application of the experimental treatment decreased the power to detect changes in local survival. For example, when the number of sites and detection probability were high, including the parameter
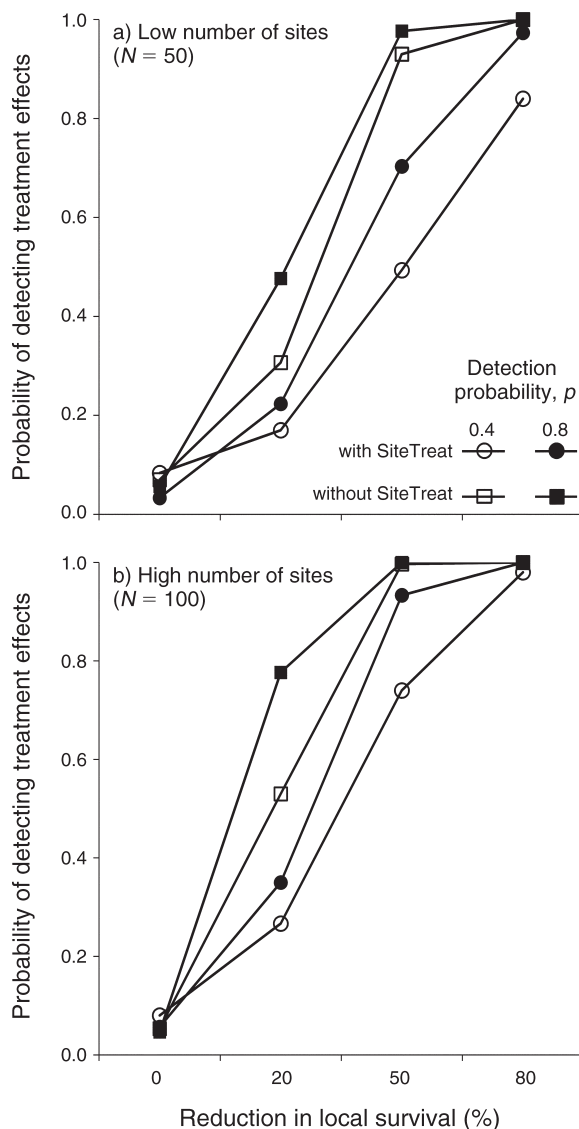
FIG. 2. Effects of estimating or not estimating a site-level covariate expressing potential background differences between Control and Impact sites (SiteTreat) on detecting treatment effects for occupancy BACI designs when no such differences are simulated. Two detection probability levels (0.4 and 0.8) were tested for (a) low and (b) high numbers of sites sampled. Type I error rates occur for 0% reduction in survival.

SiteTreat reduced power to detect a 20% change in local survival from 0.78 to 0.35 (Fig. 2a). A similar decrease in power from 0.93 to 0.43 occurred for detecting a 50% change in local survival when both the number of sites and detection were low (Fig. 2b).

When background differences in local survival between Control and Impact sites were simulated, it was critically important to include the site-specific parameter SiteTreat. Overall, the power to detect changes in local survival when the Impact sites started at low φ values was lower compared to the baseline scenario (Fig. 3a; Appendix); this is probably because power depends on

the true parameter magnitudes and is not due to having different survivals per se. For example, when pretreatment local survival at Impact sites (φ_I) was 0.6, the power to detect a 50% change was 0.33 and 0.64 for low and high number of sites, respectively, compared to 0.70 and 0.93 for the baseline scenario. The power further
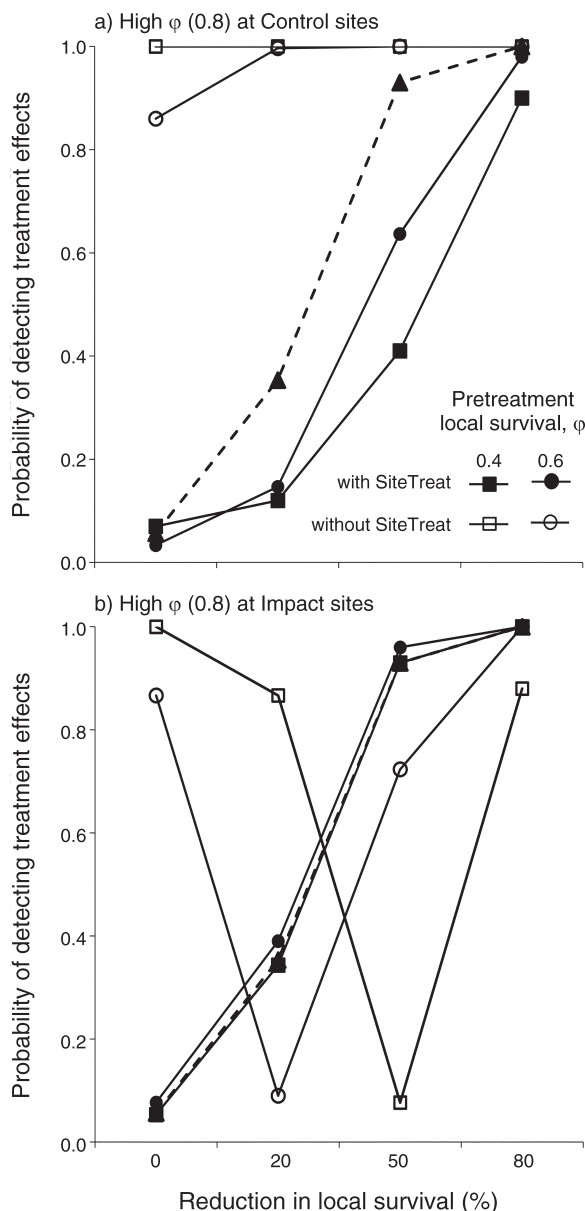

FIG. 3. Effects of estimating or not estimating a site-level covariate expressing potential background differences between Control and Impact sites (SiteTreat) on detecting treatment effects for occupancy BACI designs when such differences are simulated. Results are shown for $N = 100$ sites with high detection ($p = 0.8$) only, for (a) Control and (b) Impact sites. Lower pretreatment local survival (φ) at Control or Impact sites is either 0.4 or 0.6. Triangles (and the dashed line) depict the baseline scenario (no background differences in local survival). Type I error rates occur for 0% reduction in survival.
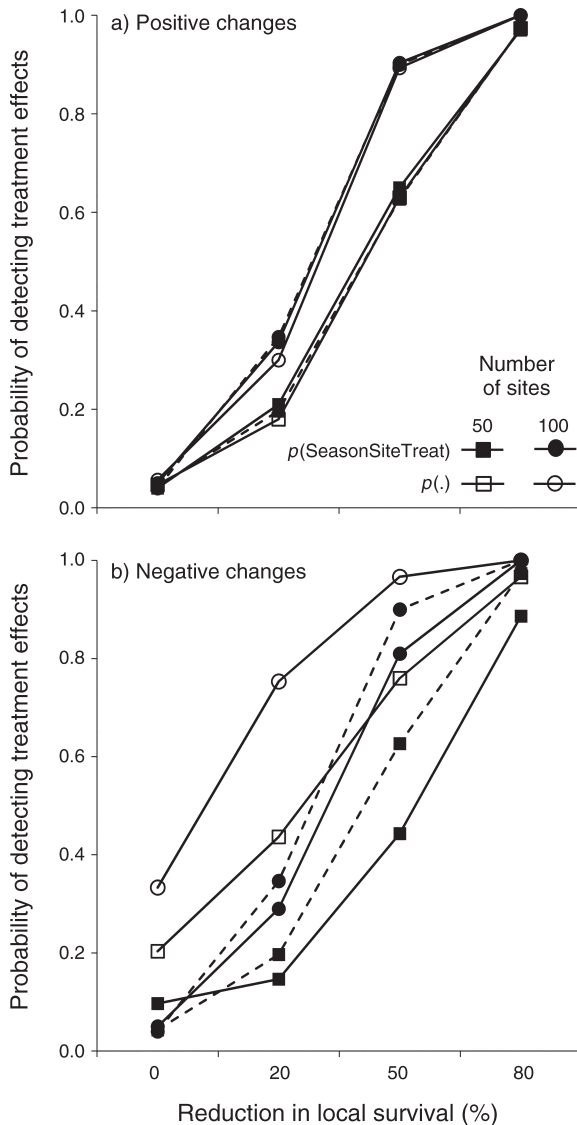
FIG. 4. Probability of detecting treatment effects for occupancy BACI designs when the experimental treatment induces (a) positive changes in detection probability (i.e., from 0.6 to 0.9), and (b) negative changes in detection probability (i.e., from 0.6 to 0.3) for low ($N = 50$) and high ($N = 100$) numbers of sites sampled and for "wrong" models, $p(.)$, vs. "correct" models, $p(SeasonSiteTreat)$. For comparison, the dashed lines denote the "no treatment effect on detection probability" scenario ($p_C = p_I = 0.6$, where subscripts C and I denote Control and Impact). Type I error rates occur for 0% reduction in survival.

decreased when φ_I was 0.4. When variable SiteTreat was not included, the models falsely detected changes in local survival when, in fact, they did not occur (the Type I error rates were grossly inflated up to 1.0 when φ_I was 0.4; Fig. 3a; Appendix).

When the Control sites had overall low local survival ($\varphi = 0.6$ or 0.4) compared to Impact sites ($\varphi = 0.8$), the correct models accounting for such differences yielded

power values similar to the baseline scenario (Fig. 3b; Appendix). However, erroneously assuming no background differences affected the Type I error rates (Fig. 3b; Appendix) and caused failure to detect a change in local survival when Control and Impact values were similar. For example, for local survival at Control sites (φ_C) = 0.6 and local survival at Impact sites (φ_I) = 0.64 (i.e., a 20% decrease from 0.8), the power value was 0.05 for the low number of sites. Similar values were recorded when $\varphi_C = 0.4$ and $\varphi_I = 0.4$ (i.e., a 50% decrease from 0.8) (Fig. 3b; Appendix).

Treatment effects on detection

Estimating changes in detection probability had no discernable cost in terms of statistical power when there was a positive treatment-induced effect on detection. For example, when the simulated data included a 50% increase in detection probability from the treatment (from 0.6 to 0.9), Type I error rates and power were similar to those of the baseline model, regardless of whether or not a treatment effect on detection was included in the model. This result held true for both low and high number of sites, with at least 100 sites needed to achieve a 0.9 power to detect a 50% decrease in local survival (Fig. 4a; Appendix). However, erroneously assuming constant detection probability in the analysis model yielded dramatically inflated Type I error rates when there was a negative treatment-induced effect on detection (from 0.6 to 0.3). The actual Type I error rates were 0.20 and 0.33 for the low and high number of sites, respectively, much greater than the intended value of 0.05 (0% change in local survival; Fig. 4b). Thus, the apparent increase in power when there was a negative treatment-induced effect on detection (Fig. 4b) was an artifact of an overly permissive test.

Design trade-offs of BACI occupancy studies

There was overall higher power to detect 20–80% changes in local survival posttreatment when we sampled a high total number of sites, over a longer period of time (12 seasons), for highly detectable species. Low detection probability had a substantial impact on the power at all levels of decrease in local survival when sites were sampled four times per season, but using six surveys per season per site increased the power to levels similar to the high-detection scenarios. There was overall very low power (<0.5) to detect a 20% decrease in local survival, even for a high number of sites over a 12-season sampling period (Fig. 5). Sampling a low number of sites yielded a moderate power (0.7) to detect a 50% decrease in local survival. In general, an 80% decrease in local survival was well captured by every BACI occupancy design (power >0.9, Fig. 5; Appendix).

An unbalanced design in terms of both sites and seasons sampled (i.e., unequally divided between Control–Impact and Before–After, respectively; Questions 3b, d) performed worse compared to a balanced design
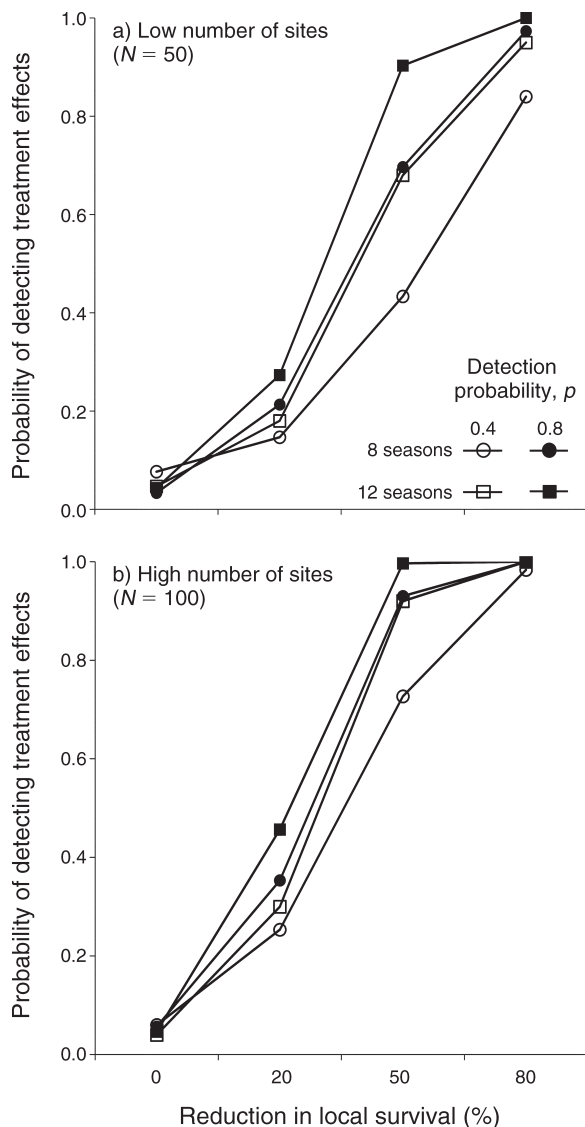
Fig. 5. Probability of detecting treatment effects for balanced occupancy BACI designs (4 visits per site per season). The number of sites was equally divided between Control and Impact sites; the number of seasons (8 or 12) was equally divided between Before and After periods. Two detection probability levels (0.4 and 0.8) were tested for (a) low and (b) high numbers of sites sampled over 8 seasons vs. 12 seasons. Type I error rates occur for 0% reduction in survival.

(i.e., three seasons prior and five seasons after) resulted in power levels similar to those of the balanced design (Fig. 6).

Sampling four times as many Control sites when the potential number of Impact sites was low ($N = 25$; Question 3c) improved the power to detect a 50% decrease in local survival by >0.2. However, the increase in power was similar to the increase expected when three times as many Control sites were sampled (covered under Question 3b: high number of sites, 30% of them

(Question 3a). Specifically, sampling substantially fewer Control sites (Question 3b) lowered power to detect 50% and 80% decreases in local survival (for the low number of sites), and 50% decrease in local survival (for the high number of sites) by 0.05–0.19. When sampling was mostly conducted after the treatment (i.e., two seasons prior and six seasons after; Question 3d), this design typically failed to correctly identify an 80% decrease in local survival (for the low number of sites), and a 50% decrease (for the high number of sites) (Fig. 6). However, adding one more sampling season Before
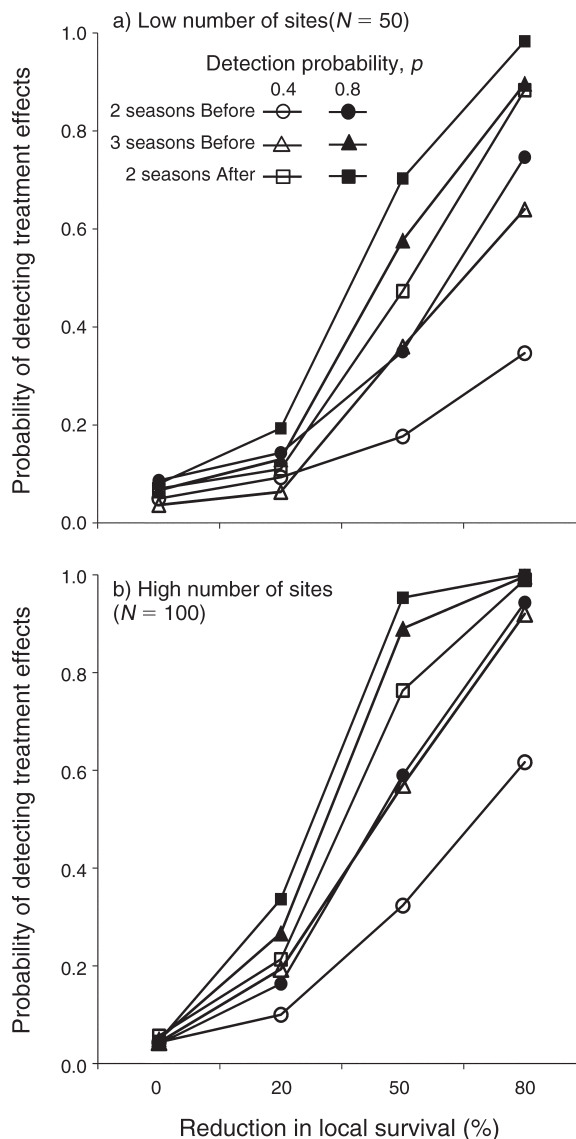

Fig. 6. Probability of detecting treatment effects for unbalanced occupancy BACI designs (eight seasons with four visits per site per season) when fewer seasons ($N = 2$ and $N = 3$) are sampled pretreatment, and fewer seasons ($N = 2$) are sampled posttreatment, for (a) low and (b) high numbers of sites sampled. The baseline scenario is similar to the "fewer seasons posttreatment" scenario. Type I error rates occur for 0% reduction in survival.
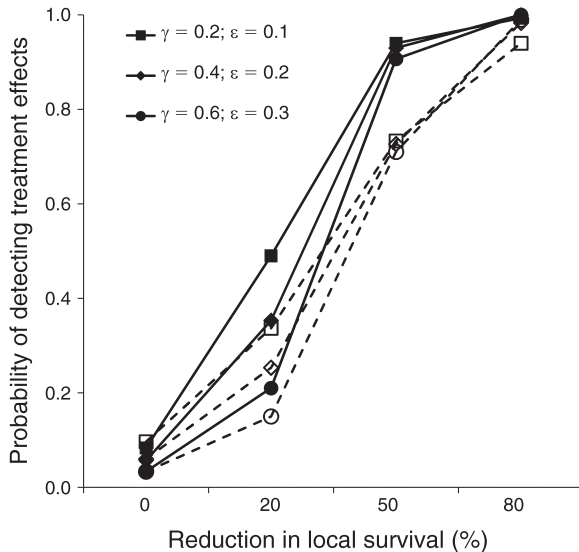
FIG. 7. Probability of detecting treatment effects when using occupancy BACI designs (high number of sites) for species with different turnover rates, for different levels of colonization (γ) and extinction (ε). All three scenarios yield an initial occupancy in season 1 of $\psi_1 = 0.66$. Solid symbols indicate high detection ($p = 0.8$); open symbols indicate low detection ($p = 0.4$). Type I error rates occur for 0% reduction in survival.

Impact), suggesting that an addition of 25 Control sites is not required to improve power. When twice as many Control sites were sampled for the high number of Impact sites scenario ($N = 50$), yielding a total of 150 sites, there was only a minor (<0.1) power increase (Appendix).

Implementing experimental treatments in different seasons at each site yielded higher power to detect 50% changes in local survival at low number of sites compared to the baseline (balanced) scenario (e.g., 0.88 vs. 0.70 for highly detectable species). For the high number of sites, the increase in power compared to the baseline scenario was most substantial for a 20% change in local survival, but the overall power was still low (e.g., 0.49 vs. 0.35 for highly detectable species).

Low vs. high turnover rates

Considering species with different turnover rates had only a moderate impact on the power to detect changes in local survival posttreatment. As expected, there was overall lower power for high-turnover species (extinction, $\varepsilon = 0.3$; colonization, $\gamma = 0.6$), and higher power for low-turnover species ($\varepsilon = 0.1$; $\gamma = 0.2$), compared to our basic scenario ($\varepsilon = 0.2$; $\gamma = 0.4$), but the differences were mostly manifested for a 20% decrease in local survival and the high number of sites (Fig. 7; Appendix). The design trade-offs in terms of low vs. high number of sites, visits per season, and detection probability were maintained for both low- and high-turnover species.

Discussion

Occupancy studies conducted in a BACI experimental design offer the opportunity to detect environmental impacts on wildlife populations without the costs of intensive studies (e.g., mark–recapture). Some of the design trade-offs for dynamic occupancy BACI studies are relatively straightforward to interpret and common to general occupancy sampling designs (MacKenzie and Royle 2005, Bailey et al. 2007). In general, a larger number of sites and monitoring seasons will yield better power to detect small changes in local survival, with low detectability having a substantial negative impact on power (Figs. 2 and 5). There was generally lower power to detect changes in local survival for species with high rates of turnover (i.e., high extinction and colonization rates) compared to low-turnover species. However, our simulations yielded several results that offer insight into additional pitfalls or advantages of using dynamic occupancy studies to investigate impacts on wildlife populations under a BACI framework.

Our study showed that dynamic occupancy BACI studies are highly flexible in terms of both experimental design and data analysis. Logistic, financial, or socio-political issues can interfere with the implementation of large-scale treatments in field studies, even with the a priori establishment of a balanced sampling design. Under these potential restrictions, alternatives such as implementing treatments at different times or adding more Control sites (when there are a limited number of potential Impact sites) can retain satisfactory power to detect changes in local survival of 50% or more. From an analytical perspective, we also found that using seasonal effects to parameterize the colonization and extinction components of dynamic occupancy models (MacKenzie et al. 2003) came at little or no cost in terms of statistical power compared to the Before–After effects (i.e., variable SeasonBA). This result depends on the seasonal variation in our simulated data sets; the mean standard deviation of local survival ranged between 0.03 and 0.05 (low seasonal effects) and between 0.05 and 0.09 (high seasonal effects) on the probability scale. This suggests some robustness to how seasonal differences are handled, but one should nevertheless be cautious using the more simplistic SeasonBA approach. Moreover, the flexible Season model allowed us to investigate the power to detect changes in local survival for scenarios when the experimental treatments are applied during different seasons at each Impact site, which is also likely for many real-world studies. Interestingly, the power to detect 50% and 20% changes in local survival at the low and high number of sites sampled, respectively, was higher compared to the basic, balanced BACI design scenario. This might be an outcome of combining sites with different combinations of pretreatment and posttreatment information.

Another notable result was the sensitivity of tests to how site-specific treatment effects on extinction and colonization were modeled (Fig. 2). Estimating the effect

of treatment (e.g., site-specific variable SiteTreat, representing Control vs. Impact sites) when no pretreatment (background) differences between Control and Impact sites were simulated considerably lowered the power to detect changes in local survival. This finding held true for both seasonal-effect and Before–After effect models. Note that our scenarios explored only systematic differences between site categories rather than random variation between individual sites, although the latter could obviously be important. However, estimating the effects of site categories (i.e., variable SiteTreat) was critical when background differences in local survival and equilibrium occupancy existed between Control and Impact sites. More specifically, for models that lacked this variable, the pretreatment differences were confounded with changes in local survival due to treatment, or failed to detect a change in local survival when changes did occur (Fig. 3). Our findings suggest that when differences are suspected to occur (e.g., small sample size, nonrandom allocation of Control and Impact sites), it is always beneficial to estimate site category effects to avoid spurious results, although it will always come at a cost in terms of statistical power (Fig. 2). For cases in which the number of sites is high and the allocation of sites into Control and Impact categories is completely randomized, potential pretreatment difference should not be an issue. As such, models that do not estimate site effects and yield higher statistical power can be used (Fig. 2). Researchers may also include any particular site-specific covariates (e.g., habitat quality) suspected to account for these differences.

The probability of detection affected the power to detect changes in local survival. Less detectable species ($p = 0.4$) required an extra two visits per site per season ($N = 6$ vs. $N = 4$ visits) to achieve the same power to detect changes in local survival as for highly detectable species ($p = 0.8$). However, a high number of sites sampled fewer times ($N = 4$ visits) yielded higher power for less detectable species, than visiting a low number of sites six times. This result was consistent across all scenarios and questions (Appendix). This suggests that, for less detectable species, effort in designing BACI occupancy studies should be directed toward increasing the number of sites sampled (which would also improve spatial replication) and not visits per site per season.

Addressing trade-offs between site replication and number of surveys per site is particularly important for resource managers designing occupancy studies and can affect bias and precision. On one hand, sampling more sites using a single survey per site could potentially increase power to detect potential differences, but resolving those differences in terms of detection, colonization, and/or extinction may be difficult. On the other hand, using a robust design with fewer sites sampled allows better estimation of detection probabilities, which, if they are not constant, could clarify differences in demographic rates and occupancy. Al-

though it is outside the scope of this manuscript, an in-depth exploration of such trade-offs is extremely important for evaluating the performance of broadscale monitoring programs based on a single survey per site, which could also account for imperfect detection (Lele et al. 2012, Sólymos et al. 2012).

The statistical power to detect changes in local survival was strongly affected by changes in detection probability associated with the experimental treatment. Specifically, a decrease in detection probability (from 0.6 to 0.3) was easily confounded with a decrease in local survival when the effect of treatment on detection was not modeled (Fig. 4b). In contrast, for the simulation scenarios denoting no change or an increase in detection probability (from 0.6 to 0.9), estimating an additional seasonal-site parameter SeasonSiteTreat for the detection probabilities had no cost in terms of statistical power (Fig. 4a). However, it is important to acknowledge that this finding is likely to be sensitive to the baseline parameters for detection probability used in this simulation, and may not hold true for rare or cryptic species with low detection probabilities (e.g., $p \leq 0.1$). More specifically, both baseline detection probability (0.6) and the 50% increase to 0.9 result in a cumulative detection probability close to 1.0. For low-detection species, the cumulative detection probability is always considerably less than 1.0 (unless each site is surveyed >30 times per season); not accounting for increased detectability as a result of an experimental treatment might also result in significantly higher Type I error rates. This suggests that researchers should include a treatment effect on detection to avoid spurious conclusions regarding changes in local survival at the impacted sites. More generally, changes in detection can be attributed to changes in local densities, and modeling heterogeneity in detection can be used to assess species abundance (Royle and Nichols 2003). This type of approach potentially could be extended to BACI models, with the caveat that it involves another layer of assumptions that may be hard to assess.

Despite the flexibility in design and analysis that we have outlined, dynamic occupancy BACI studies will probably fail (power <0.5) to detect small changes in local survival (20%), even for a large number of sites sampled (i.e., 150 total sites;, Question 3c). Therefore, when researchers have prior knowledge about the magnitude of potential treatment effects, they should only consider implementing occupancy studies when substantial impacts are expected (>50% decrease in local survival).

Russell et al. (2009) suggested that monitoring more seasons posttreatment is needed to detect the effect of environmental impacts. Although this is a rational statement, it is highly dependent on the response variable through which environmental impacts are assessed, as well as the time frame of sampling pretreatment. In many large-scale BACI studies, emphasis is placed on gathering data on the effects of

disturbance posttreatment, and minimal baseline data are usually gathered (e.g., one or two seasons prior to treatment). In general, there is a trade-off between the intensity of sampling methodology and the feasibility of obtaining extensive pretreatment baseline data. For example, Lindenmayer et al. (2009) used detection–nondetection methods (low-intensity sampling) to estimate community-level changes in bird and mammal species richness over broad areas. In contrast, Rittenhouse et al. (2009) used telemetry (highly intensive sampling) to estimate amphibian survival in relation to forest management practices over a more limited area.

Our power analyses showed that unless data are gathered for at least three seasons before the treatments are implemented, monitoring more seasons After the Impact using dynamic occupancy methods may not provide accurate estimation of changes in local survival, even when the number of sites sampled is high (Fig. 6). This corresponds to a minimum number of two "transition" periods between sampling seasons corresponding to the number of potential colonization or extinction events. Consequently, the explanation for the substantial difference in power values between scenarios with two vs. three seasons sampled Before (Fig. 6) is probably an outcome of modeling one vs. two state transitions (i.e., occupancy in season 1 is "fixed" because there are no data available to estimate colonization or extinction rates before the first sampling season). Simply, one transition period (i.e., two seasons sampled Before) is not enough to accurately describe the ecological processes of extinction and colonization. The same rule is likely to apply if sampling occurs for only one season After, which also yields a single transition period (e.g., from season 7 to 8 in our case). Notably, these results are dependent on the assumption that, in our study, occupancy dynamics was expressed as a first-order Markov process.

We limited our simulation and analysis choices in several respects. Model selection might be used to analyze real data. Model selection and likelihood ratio hypothesis tests are intimately linked: both use the log likelihood ratio difference to compare two models, and the validity of both derives from asymptotic (large-sample) theory of maximum likelihood estimation. However, model selection introduces a host of more complex issues, so here we limited ourselves to asking how "correct" and "incorrect" models perform for likelihood ratio hypothesis tests. The single-state approach adopted in this study can be extended to dynamic multi-state occupancy (Nichols et al. 2007, MacKenzie et al. 2009, 2010, Green et al. 2011), in which BACI occupancy studies can be used to assess changes in reproductive status or reproductive output. Along with detection–nondetection data, another type of data used in occupancy studies is count data, which allows estimation of local densities and abundance of species while incorporating imperfect detection probability (Royle 2004). Such data have been used widely for determining population trends and impacts on local abundances, and have potential for higher statistical power (Kéry and Royle 2010, Chelgren et al. 2011, Moore and Barlow 2011). However, detecting relatively small changes in demographic parameters and/or local densities is likely to be subject to the same limitations as detecting environmental impacts on local survival.

We assumed spatial independence among sites and omitted an explicit treatment of spatial patterns (Risk et al. 2011). Even in carefully designed sampling schemes, latent spatial relationships between sampling units (e.g., determined by the distribution of a certain habitat type or by the location of the sampling plot or point in relation to the edge or core of the species range) can bias the analysis outcome. Recent studies using detection–nondetection, count, or mark–recapture data (e.g., Thogmartin et al. 2006, Saracco et al. 2010, Aing et al. 2011, Bled et al. 2011, Chelgren et al. 2011, Kéry et al. 2011, Royle et al. 2011) showed that accounting for the spatial arrangement of sampling site or plots is essential to obtain unbiased coefficient estimates. However, these studies were conducted in a Bayesian framework to facilitate spatial correlation structures, whereas the software used in our study did not (Fiske and Chandler 2011). Moreover, including spatial information in the data simulation was beyond the objective of our study.

In conclusion, well-designed dynamic occupancy studies provide a great opportunity to implement low-cost single- or multispecies monitoring programs (DeWan and Zipkin 2010), and adopting the BACI analysis framework described in our study can improve the detection of environmental impacts on wildlife populations. However, it is important to acknowledge the limitations of detection–nondetection data for dynamic occupancy studies for detecting relatively small changes in site occupancy.

Literature Cited

Aing, C., S. Halls, K. Oken, R. Dobrow, and J. Fieberg. 2011. A Bayesian hierarchical occupancy model for track surveys conducted in a series of linear, spatially correlated, sites. Journal of Applied Ecology 48:1508–1517.

Archaux, F., P.-Y. Henry, and O. Gimenez. 2012. When can we ignore the problem of imperfect detection in comparative studies? Methods in Ecology and Evolution 3:188–194.

Bailey, L. L., J. E. Hines, J. D. Nichols, and D. I. MacKenzie. 2007. Sampling design trade-off in occupancy studies with imperfect detection: examples and software. Ecological Applications 17:281–290.

Benedetti-Cecchi, L. 2001. Beyond BACI: optimization of environmental sampling designs through monitoring and simulation. Ecological Applications 11:783–799.

Blakesley, J. A., et al. 2010. Population dynamics of Spotted Owls in the Sierra Nevada, California. Wildlife Monographs 174:1–36.

Bled, F., J. A. Royle, and E. Cam. 2011. Hierarchical modeling of an invasive spread: the Eurasian Collared-Dove *Streptopelia decaocto* in the United States. Ecological Applications 21:290–302.

Chelgren, N. D., M. J. Adams, L. L. Bailey, and R. B. Bury. 2011. Using multilevel spatial models to understand salamander site occupancy patterns after wildfire. Ecology 92:408–421.

Collier, B. A., M. L. Morrison, S. L. Farrell, A. J. Campomizzi, J. A. Butcher, K. B. Hays, D. I. MacKenzie, and R. N. Wilkins. 2010. Monitoring golden-cheeked warblers on private lands in Texas. Journal of Wildlife Management 74:140–147.

Conquest, L. L. 2000. Analysis and interpretation of ecological field data using BACI designs: discussion. Journal of Agricultural, Biological, and Environmental Statistics 5:293–296.

Dail, D., and L. Madsen. 2011. Models for estimating abundance from repeated counts of an open metapopulation. Biometrics 67:577–587.

DeWan, A. A., and E. F. Zipkin. 2010. An integrated sampling and analysis approach for improved biodiversity monitoring. Environmental Management 45:1223–1230.

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 100:389–398.

Fairweather, P. G. 1991. Statistical power and design requirements for environmental monitoring. Australian Journal of Marine and Freshwater Research 42:555–567.

Fiske, I., and R. Chandler. 2011. unmarked: an R package for fitting hierarchical models of wildlife occurrence and abundance. Journal of Statistical Software 43:1–23.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology 68:1364–1372.

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

Green, A. W., L. L. Bailey, and J. D. Nichols. 2011. Exploring sensitivity of a multistate occupancy model to inform management decisions. Journal of Applied Ecology 48:1007–1016.

Green, R. H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley, New York, New York, USA.

Hanski, I. 1994. A practical model of metapopulation dynamics. Journal of Animal Ecology 63:151–162.

Hanski, I. 1998. Metapopulation dynamics. Nature 396:41–49.

Hewitt, J. E., S. E. Thrush, and V. J. Cummings. 2001. Assessing environmental impacts: effects of spatial and temporal variability at likely impact scales. Ecological Applications 11:1502–1516.

Joseph, L. N., S. A. Field, C. Wilcox, and H. P. Possingham. 2006. Presence–absence versus abundance data for monitoring threatened species. Conservation Biology 20:1679–1687.

Kendall, K. C., L. H. Metzgar, D. A. Patterson, and B. M. Steele. 1992. Power of sign surveys to monitor population trends. Ecological Applications 2:422–430.

Kendall, W. L., J. D. Nichols, and J. E. Hines. 1997. Estimating temporary emigration using capture–recapture data with Pollock's robust design. Ecology 78:563–578.

Kéry, M., B. Gardner, T. Stoeckle, D. Weber, and J. A. Royle. 2011. Use of spatial capture–recapture modeling and DNA data to estimate densities of elusive animals. Conservation Biology 25:356–364.

Kéry, M., and J. A. Royle. 2008. Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys. Journal of Applied Ecology 45:589–598.

Kéry, M., and A. J. Royle. 2010. Hierarchical modelling and estimation of abundance and population trends in metapopulation designs. Journal of Animal Ecology 79:453–461.

Lele, S. R., M. Moreno, and E. Bayne. 2012. Dealing with detection error in site occupancy surveys: what can we do with a single survey? Journal of Plant Ecology 5:22–31.

Lindenmayer, D., J. Wood, R. Cunningham, M. Crane, C. Macgregor, D. Michael, and R. Montague-Drake. 2009. Experimental evidence of the effects of a changed matrix on conserving biodiversity within patches of native forest in an industrial plantation landscape. Landscape Ecology 24:1091–1103.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84:2200–2207.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. P. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling. Inferring patterns and dynamics of species occurrence. Academic Press, San Diego, California, USA.

MacKenzie, D. I., J. D. Nichols, M. E. Seamans, and R. J. Gutierrez. 2009. Modeling species occurrence dynamics with multiple states and imperfect detection. Ecology 90:823–835.

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

MacKenzie, D. I., M. Seamans, R. Gutiérrez, and J. Nichols. 2010. Investigating the population dynamics of California spotted owls without marked individuals. Journal of Ornithology. http://dx.doi.org/10.1007/s10336-010-0544-6

Marsh, D. M., and P. C. Trenham. 2008. Current trends in plant and animal population monitoring. Conservation Biology 22:647–655.

Martin, J., S. Chamaille-Jammes, J. D. Nichols, H. Fritz, J. E. Hines, C. J. Fonnesbeck, D. I. MacKenzie, and L. L. Bailey. 2010. Simultaneous modeling of habitat suitability, occupancy, and relative abundance: African elephants in Zimbabwe. Ecological Applications 20:1173–1182.

Martin, J., C. L. McIntyre, J. E. Hines, J. D. Nichols, J. A. Schmutz, and M. C. MacCluskie. 2009. Dynamic multistate site occupancy models to evaluate hypotheses relevant to conservation of Golden Eagles in Denali National Park, Alaska. Biological Conservation 142:2726–2731.

Martin, J., J. A. Royle, D. I. MacKenzie, H. H. Edwards, M. Kéry, and B. Gardner. 2011. Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. Methods in Ecology and Evolution 2:595–601.

Mattfeldt, S. D., L. L. Bailey, and E. H. C. Grant. 2009. Monitoring multiple species: Estimating state variables and exploring the efficacy of a monitoring program. Biological Conservation 142:720–737.

McDonald, T. L., W. P. Erickson, and L. L. McDonald. 2000. Analysis of count data from Before–After Control–Impact studies. Journal of Agricultural, Biological, and Environmental Statistics 5:262–279.

Moilanen, A. 1999. Patch occupancy models of metapopulation dynamics: efficient parameter estimation using implicit statistical inference. Ecology 80:1031–1043.

Moore, J. E., and J. Barlow. 2011. Bayesian state-space model of fin whale abundance trends from a 1991–2008 time series of line-transect surveys in the California Current. Journal of Applied Ecology 48:1195–1205.

Murtaugh, P. A. 2002. On rejection rates of paired intervention analysis. Ecology 83:1752–1761.

Muths, E., et al. 2005. Amphibian research and monitoring initiative (ARMI): a successful start to a national program in the United States. Applied Herpetology 2:355–371.

Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998. Estimating rates of local species extinction, colonization, and turnover in animal communities. Ecological Applications 8:1213–1225.

Nichols, J. D., J. E. Hines, D. I. MacKenzie, M. E. Seamans, and R. J. Gutierrez. 2007. Occupancy estimation and modeling with multiple states and state uncertainty. Ecology 88:1395–1400.

Osenberg, C. W., and R. J. Schmitt. 1996. Detecting ecological impacts caused by human activities. Pages 3–16 in R. J. Schmitt and C. W. Osenberg, editors. Detecting ecological impacts: concepts and applications in coastal habitats. Academic Press, San Diego, California, USA.

Pollock, J. F. 2006. Detecting population declines over large areas with presence–absence, time-to-encounter, and count survey methods. Conservation Biology 20:882–892.

R Development Core Team. 2011. R: A language and environment for statistical computing. Vienna, Austria. http://www.R-project.org/

Rhodes, J. R., A. J. Tyre, N. Jonzen, C. A. McAlpine, and H. P. Possingham. 2006. Optimizing presence–absence surveys for detecting population trends. Journal of Wildlife Management 70:8–18.

Risk, B. B., P. de Valpine, and S. R. Beissinger. 2011. A robust-design formulation of the incidence function model of metapopulation dynamics applied to two species of rails. Ecology 92:462–474.

Rittenhouse, T. A. G., R. D. Semlitsch, and F. R. Thompson III. 2009. Survival costs associated with wood frog breeding migrations: effects of timber harvest and drought. Ecology 90:1620–1630.

Rota, C. T., R. J. Fletcher, R. M. Dorazio, and M. G. Betts. 2009. Occupancy estimation and the closure assumption. Journal of Applied Ecology 46:1173–1181.

Roth, T., and V. Amrhein. 2010. Estimating individual survival using territory occupancy data on unmarked animals. Journal of Applied Ecology 47:386–392.

Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. Biometrics 60:108–115.

Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities. Academic Press, San Diego, California, USA.

Royle, J. A., and M. Kéry. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 88:1813–1823.

Royle, J. A., A. J. Magoun, B. Gardner, P. Valkenburg, and R. E. Lowell. 2011. Density estimation in a wolverine population using spatial capture–recapture models. Journal of Wildlife Management 75:604–611.

Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence–absence data or point counts. Ecology 84:777–790.

Russell, R. E., J. A. Royle, V. A. Saab, J. F. Lehmkuhl, W. M. Block, and J. R. Sauer. 2009. Modeling the effects of environmental disturbance on wildlife communities: avian responses to prescribed fire. Ecological Applications 19:1253–1263.

Saracco, J. F., D. F. Desante, and D. R. Kaschube. 2008. Assessing landbird monitoring programs and demographic causes of population trends. Journal of Wildlife Management 72:1665–1673.

Saracco, J. F., J. A. Royle, D. F. DeSante, and B. Gardner. 2010. Modeling spatial variation in avian survival and residency probabilities. Ecology 91:1885–1891.

Sólymos, P., S. Lele, and E. Bayne. 2012. Conditional likelihood approach for analyzing single visit abundance survey data in the presence of zero inflation and detection error. Environmetrics 23:197–205.

Stewart-Oaten, A., and J. R. Bence. 2001. Temporal and spatial variation in environmental impact assessment. Ecological Monographs 71:305–339.

Stewart-Oaten, A., W. W. Murdoch, and K. R. Parker. 1986. Environmental impact assessment: "pseudoreplication in time?" Ecology 67:929–940.

Strayer, D. L. 1999. Statistical power of presence–absence data to detect population declines. Conservation Biology 13:1034–1038.

Taylor, B. L., and T. Gerrodette. 1993. The uses of statistical power in conservation biology: the vaquita and Northern Spotted Owl. Conservation Biology 7:489–500.

Thogmartin, W. E., M. G. Knutson, and J. R. Sauer. 2006. Predicting regional abundance of rare grassland birds with a hierarchical spatial count model. Condor 108:25–46.

Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, and H. P. Possingham. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecological Applications 13:1790–1801.

Underwood, A. J. 1991. Beyond BACI: Experimental designs for detecting human environmental impacts on temporal variations in natural populations. Australian Journal of Marine and Freshwater Research 42:569–587.

Underwood, A. J. 1993. The mechanics of spatially replicated sampling programs to detect environmental impacts in a variable world. Australian Journal of Ecology 18:99–116.

Usher, M. B. 1979. Markovian approaches to ecological succession. Journal of Animal Ecology 48:413–426.

Vuilleumier, S., and H. P. Possingham. 2006. Does colonization asymmetry matter in metapopulations? Proceedings of the Royal Society B 273:1637–1642.

Williams, B. K. 2001. Uncertainty, learning, and the optimal management of wildlife. Environmental and Ecological Statistics 8:269–288.

Zielinski, W. J., and H. B. Stauffer. 1996. Monitoring Martes populations in California: survey design and power analysis. Ecological Applications 6:1254–1267.

SUPPLEMENTAL MATERIAL

Appendix

Probabilities of detecting changes in local survival for the full complement of questions and scenarios (*Ecological Archives* A022-072-A1).

Supplement

R code for simulating treatment effects on local survival and models fitted with the R package *unmarked* (*Ecological Archives* A022-072-S1).