



Realized Population Change for Long-Term Monitoring: California Spotted Owl Case Study

MARY M. CONNER,^{1,2} *USDA Forest Service, Pacific Southwest Research Station, 1731 Research Park Drive, Davis, CA 95618, USA*

JOHN J. KEANE, *USDA Forest Service, Pacific Southwest Research Station, 1731 Research Park Drive, Davis, CA 95618, USA*

CLAIRE V. GALLAGHER, *USDA Forest Service, Pacific Southwest Research Station, 1731 Research Park Drive, Davis, CA 95618, USA*

GRETCHEN JEHL, *USDA Forest Service, Pacific Southwest Research Station, 1731 Research Park Drive, Davis, CA 95618, USA*

THOMAS E. MUNTON, *USDA Forest Service, Pacific Southwest Research Station, 2081 East Sierra Avenue, Fresno, CA 93710, USA*

PAULA A. SHAKLEE, *USDA Forest Service, Pacific Southwest Research Station, 1731 Research Park Drive, Davis, CA 95618, USA*

ROSS A. GERRARD, *USDA Forest Service, Pacific Southwest Research Station, 1731 Research Park Drive, Davis, CA 95618, USA*

ABSTRACT The annual rate of population change (λ_t) is a good metric for evaluating population performance because it summarizes survival and recruitment rates and can be used for open populations. Another measure of population performance, realized population change (Δ_t) is an encompassing metric of population trend over a period of time; it is the ratio of population size at an end time period relative to the initial population size. Our first goal was to compare mean λ and Δ_t as summaries of population change over time. Our second goal was to evaluate different methods for estimating these parameters; specifically we wished to compare the value of estimates from fixed effects models, random effects estimates from mixed effects models, and Bayesian Markov chain Monte Carlo (MCMC) methods. Our final goal was to evaluate the use of the posterior distribution of Δ_t as a means for estimating the probability of population decline retrospectively. To meet these goals, we used California spotted owl (*Strix occidentalis occidentalis*) data collected on 3 study areas from 1990 to 2011 as a case study. The estimated MCMC median λ s for 2 of the study areas were 0.986 and 0.993, indicating declining populations, whereas median λ was 1.014 for the third study area, indicating an increasing population. For 2 of the study areas, estimated MCMC median Δ_t s over the 18-year monitoring period were 0.78 and 0.89, suggesting 21% and 11% declines in population size, whereas the third study area was 1.22 suggesting a 22% increase. Results from Δ_t analyses highlight that small differences in mean λ from 1.0 (stationary) can result in large differences in population size over a longer time period; these temporal effects are better depicted by Δ_t . Fixed effects, random effects, and MCMC estimates of mean and median λ and of Δ_t were similar ($\leq 9\%$ relative difference). The estimate of temporal process variance was larger for MCMC than the random effects estimates. Results from a Bayesian approach using MCMC simulations indicated that the probabilities of a $\geq 15\%$ decline over 18 years were 0.69, 0.40, and 0.04 for the 3 study areas, whereas the probabilities the populations were stationary or increasing were 0.07, 0.22, and 0.82. For retrospective analyses of monitored populations, using Bayesian MCMC methods to generate a posterior distribution of Δ_t is a valuable conservation and management tool for robustly estimating probabilities of specified declines of interest. © 2013 The Wildlife Society.

KEY WORDS Bayesian MCMC approach, California, California spotted owl, hierarchical model monitoring, Pradel's temporal symmetry model, random effects estimator, rate of population change, realized population change, *Strix occidentalis occidentalis*.

Long-term monitoring programs are often focused on species of concern, from game to threatened species. Management of these species is often controversial and sociopolitical challenges are often as important as biological

ones. Controversial species require clear, simple, and objective biological metrics to inform the management decision process, particularly for regular and relatively short-term intervals, such as management plans and status updates (e.g., ≤ 5 years). Population size is often used because of its ease of interpretation and because management or conservation goals often include maintaining population levels close to sustainable objectives (Thompson et al. 1998, Williams et al. 2002) or above some minimum threshold, such as the size of a minimum viable population (Soulé 1980). However, population size alone reveals little about the long-term

Received: 17 August 2012; Accepted: 15 May 2013

Published: 1 August 2013

¹E-mail: mary.conner@usu.edu

²Present address: Department of Watershed Sciences, Utah State University, 5210 Old Main Hill, Logan, UT 84322, USA

sustainability of a population (Lancia et al. 2005). Determining whether the population is increasing, decreasing, or remaining stable is also critical to ensuring future population size objectives (Thompson et al. 1998, White et al. 2002, Williams et al. 2002). The annual rate of population change (λ_t) is a good metric for evaluating population performance because it summarizes survival and recruitment rates and can be used for open populations (Nichols and Hines 2002). The value of λ_t is interpreted as the rate of annual (or other time period of interest) increase or decrease in the population, and is used more frequently for monitoring and management decisions.

Another metric used for evaluating changes in population status is realized population change (Δ_t), which is the change in population over a period of time (Franklin et al. 2004). This quantity is the ratio of population size at some point in time period K , relative to the initial population size N ($\Delta_t = N_K/N$), which reflects the cumulative effect of λ_t for a time period of interest. For example, if λ_t had a geometric mean of 0.97, a population size starting at 100 would decline to approximately 86 over 5 years (a common management time frame) and Δ_t would be 0.86. Specifically, Δ_t incorporates all estimates of λ_t and provides an intuitive number for evaluating population change over periods of time.

In the past, the confidence intervals (CI) for estimates of mean λ and of Δ_t were used to evaluate population decline. That is, if the CI included 1, even if just barely, the conclusion was that there was no evidence for a decline. However, with this methodology, the influence and probability of a Type II error (inability to detect a decline or change) cannot be eliminated from the interpretation of no decline. The key element missing in this approach was the ability to estimate the probability of decline. Using Bayesian Markov chain Monte Carlo (MCMC) methods, a posterior distribution of Δ_t can be used to estimate any probability of decline. The approach has not been widely used, although Gerrodette and Rojas-Bracho (2011) used it to evaluate declines in vaquita (*Phocena sinus*) over a 15-year monitoring period. Indeed, the posterior distribution of Δ_t provides a robust method for detecting and describing retrospective population trajectories, similar to how population viability analysis (PVA) and more generally, projection models, are used to evaluate potential prospective population trajectories.

The inclusion of sampling variation in estimates of λ_t can negatively bias estimates of Δ_t , similar to how it can negatively bias estimates of persistence in PVA models. Bayesian MCMC methods can be used to separate process variation from sampling variation during the estimation of Δ_t . In addition to Bayesian MCMC methods, a random effects estimator, also called a shrinkage or empirical Bayes estimator, is an analytical approach to separate sampling variation from the overall process variance (Efron and Morris 1977, Ver Hoef 1996, Burnham and White 2002). Depending on computing tools available, both approaches can be used for estimating λ_t and Δ_t .

Our first goal was to compare mean λ and Δ_t as summaries of population change over time. Our second goal was to

evaluate different methods for estimating mean λ and Δ_t . Specifically, we wished to compare the value of estimates based on a random effects estimator from a mixed effects model and a Bayesian MCMC approach for evaluating population performance of species with long-term monitoring data. Our final goal was to evaluate a Bayesian approach to generate a posterior distribution of Δ_t as a means to estimate the probability of population decline retrospectively. We use data collected on the California spotted owl (*Strix occidentalis occidentalis*) as a case study. The California spotted owl remains a focal species of conservation concern, and knowledge of population trends is an important component of assessing status and informing forest management planning efforts.

STUDY AREA

We used data from spotted owls on 3 study areas in the Sierra Nevada and southern Cascade Mountains, California, collected from 1990 to 2011 (Fig. 1). Although we did not randomly select the study areas, they spanned the length of the contiguous California spotted owl range in the Sierra Nevada and encompassed all habitat types known to be used by spotted owls in the Sierra Nevada. The Lassen (LAS) study area was in the southern Cascades, but it was included in the Sierra Nevada province by the United States Forest Service for management purposes (U.S. Forest Service 2004). Most of the LAS and Sierra (SIE) study areas were located on public land managed by the United States Forest Service, whereas the Sequoia and Kings Canyon (SKC) study area was in 2 national parks. Franklin et al. (2004) and Blakesley et al. (2010) described in detail the LAS, SIE, and SKC study areas; however, SKC has changed from Blakesley et al. (2010). In 2006, the study area for SKC was reduced; all estimates of λ_t were based on data from this reduced study area. Long-term mark-resight data were available for each study population. The number of marked birds varied each year, but generally ranged between 40 and 90, with 4–14 new birds marked each year (Table 1).

METHODS

Field Surveys

Field methods for the study of spotted owl population dynamics have been well described elsewhere (Forsman 1983; Franklin et al. 1996, 2004; Anthony et al. 2006). In particular, the field methods and protocols used to collect the data are detailed by Franklin et al. (2004). Therefore, we present only a brief summary of the methods used to capture and recapture (by resighting) owls.

We conducted surveys to find and locate California spotted owls from 1 April to 31 August on LAS and from 1 March to 30 September on SIE and SKC. We identified owls by capturing and banding, or by resighting owls previously banded with unique color band combinations. We surveyed spotted owls primarily at night by vocally imitating spotted owl calls or by broadcasting recordings of spotted owl calls. We completely surveyed the core study areas on ≥ 3 occasions throughout each field season. When we detected owls, we



Figure 1. Outline of southern Cascade and Sierra Nevada Mountains, California, showing the location of 3 study areas for California spotted owls, 1990–2011.

conducted walk-in surveys during daylight hours to visually locate owls, identify color bands, or capture unbanded owls.

We identified spotted owls to sex (male, female) by vocalization and behavior, and determined age by plumage characters (Franklin et al. 1996). Owls can be identified to 4 age classes (juvenile [young of the year], 2 subadult ages [1–2 years old; 2–3 years old], and adult [≥ 3 years old]; Moen et al. 1991). We use the term non-juvenile owl to refer to subadult and adult age classes; we used only non-juvenile owls for our analyses.

Rate of Population Change

We used each owl's encounter history to estimate the annual rate of population change in territorial owls (λ_t) using Pradel's temporal symmetry model (Pradel 1996, Nichols and Hines 2002, Franklin et al. 2004) in Program MARK (White and Burnham 1999). We refer to Pradel's temporal

symmetry model as the Pradel model. All 3 study areas had territorial survey areas, in which some areas were not consistently surveyed, and a subset core study area, which was surveyed consistently for estimation of λ_t . In the core areas, coverage was complete, in that each year, we surveyed areas without known owl territories, as well as areas with known owl territories. For population growth rate to represent changes in the number of owls for the Pradel model, the study area size and boundary must remain unchanged through time (Hines and Nichols 2002, Franklin et al. 2004). Consequently, we used data only from the core areas to interpret λ_t as the annual rate of population change.

We used the $(\phi\lambda p)$ structure of the Pradel model, where ϕ is apparent survival (probability that an owl alive in year t survived to the next year $t + 1$ and remained on the study area [i.e., available for recapture or resight]) and p is the resight (by recapture or resight) probability. Based on previous California spotted owl analyses (Franklin et al. 2004, Blakesley et al. 2010), we used a model with fixed effects of sex and year, $\lambda_t(t)\phi(t)p(s \times t)$, where $s =$ sex and $t =$ year as a categorical time effect. For each study area, we estimated the overdispersion parameter (\hat{c}) using the median \hat{c} procedure in Program MARK under the Cormack-Jolly-Seber (CJS) global model $\phi(s \times t)p(s \times t)$. When \hat{c} was >1 , we used \hat{c} to inflate variances of parameter estimates (Burnham and Anderson 2002).

Using the model structure above, $\lambda_t(t)\phi(t)p(s \times t)$, we estimated the mean rate of population change using a fixed effect model ($\hat{\lambda}$). Using the same model, we also used a mixed model approach to estimate the mean rate of population change ($\bar{\lambda}$) and its temporal process variation ($\hat{\sigma}^2$), as well as to generate random effects estimates of λ_t ($\tilde{\lambda}_t$; White et al. 2001). The parameters ϕ and p were fixed effects and λ_t was a random effect. To estimate $\bar{\lambda}$, we used an intercept-only (means) model for λ_t . We did not include the first 2 estimates and last estimate of λ_t to estimate mean λ because the first and last estimates were confounded with estimation of p , and the second estimate has a potential bias from trap response or a learning curve experienced by field crews at the beginning of the studies (Hines and Nichols 2002). We calculated $\hat{\lambda}$, $SE(\hat{\lambda})$, $\bar{\lambda}$, $SE(\bar{\lambda})$, and $\hat{\sigma}^2$ on the log scale and used the delta method to estimate variance of the back-transformed geometric means of $\hat{\lambda}$ and $\bar{\lambda}$ (Franklin et al. 2004). We used the log-scale because the geometric mean generates an unbiased estimate of mean λ over a finite time period, whereas the arithmetic mean generates a biased estimate (Morris and Doak 2002).

Table 1. Descriptions of the 3 study areas in the southern Cascades and Sierra Nevada, California, and sample sizes of California spotted owls.

Study area ^a	Years	Area (km ²)	Avg. total no. marked ^b	Avg. new marked/year ^c
LAS	1990–2011	1,254	88	14
SIE	1990–2011	562	76	8
SKC	1991–2011	182	59	7

^a LAS is Lassen, SIE is Sierra, and SKC is Sequoia Kings Canyon.

^b Average number of marked owls on study area; includes territorial adult and subadult owls.

^c Average number of newly marked owls on study area; includes territorial adult and subadult owls.

Both LAS and SIE had expansion areas, areas in which surveying began several years after the start of the study. If these were not accounted for, new owls found in these areas would enter the Pradel model as new recruits, and result in a positive bias in estimates of λ_t . Therefore, we accounted for these areas by grouping them separately and excluding estimates of λ_t from the year of and year following the expansion, and estimates of \hat{p} for the year of expansion (see Blakesley et al. 2010, Appendix G, for a detailed explanation of the procedure). One study area, SKC, had a year of missed data collection (2005); we used an unequal time interval to account for this in Program MARK, which resulted in an average survival over the 2-year interval. We did not use the average estimate from these 2 years (2004–2005 and 2005–2006) in any random effects analyses because random effects in Program MARK assumes equal time intervals.

Realized Population Change

We translated fixed effects estimates of $\lambda_t(\hat{\lambda}_t)$ into estimates of realized population change ($\hat{\Delta}_t$) on each study area, which is the proportional change in estimated population size, relative to the initial population size (Franklin et al. 2004). We calculated $\hat{\Delta}_t$ as the product $1 \times \hat{\lambda}_k \times \hat{\lambda}_{k+1} \times \hat{\lambda}_{k+2} \times \dots \times \hat{\lambda}_K$, where k was year of the first estimated λ_t (i.e., the third year because the first 2 estimates of λ_t were confounded or biased) and K was the last (i.e., the second to last estimate because the last estimate was confounded). For each study area, we used the delta method to estimate variance of realized change (on the natural log scale) and the associated back-transformed CIs (Franklin et al. 2004). We also used the random effects estimates of $\lambda_t(\hat{\lambda}_t)$ to estimate realized population change ($\hat{\Delta}_t$). Because the random effects variance estimate is conditionally biased, we based CIs on root mean square error (RMSE; White et al. 2001, Burnham and White 2002), which includes a term for the difference between $\hat{\Delta}_t$ and $\hat{\Delta}_t$ (e.g., $[\hat{\Delta}_t - \hat{\Delta}_t]^2$). To estimate the covariance between random effects estimates, we multiplied the correlations between fixed effects estimates by the RMSE^2 (i.e., variance) of each estimate.

To estimate probabilities of population declines or increases for each area, we used a Bayesian approach via

MCMC sampling implemented in Program MARK to estimate posterior distributions of λ_t ($\lambda_{t\text{MCMC}}$) using the same mixed effects model and years as described above. We used the posterior distributions of $\lambda_{t\text{MCMC}}$ to estimate the posterior distribution of median λ (λ_{MCMC}) and Δ_t ($\Delta_{t\text{MCMC}}$). For all MCMC simulations, we used 4,000 tuning samples, 1,000 burn in samples, and 20,000 realizations. We used vague priors for all parameters included in the model. For estimating median λ_{MCMC} and $\Delta_{t\text{MCMC}}$, we used a hyperprior for μ and σ of λ_t . For μ , we used a normally distributed prior with mean = 0 and standard deviation = 100 and for σ , we used a gamma prior to model its transformation, $1/\sigma^2$, with $\alpha = 1.00001$ and $\beta = 0.000001$. Although we did not use estimates of ϕ and \hat{p} in the hyperprior to estimate μ and σ of λ_t , we did include them in the model, $\lambda(t)\phi(t)\hat{p}(s \times t)$, and they required a prior distribution. Because they were logit transformed parameters, we used a normal prior distribution with mean of 0 and a standard deviation of 1.75, which is a vague prior when back transformed to the real scale (2.5th and 97.5th percentiles of approximately 0.02 and 0.98, with a uniform distribution between those percentiles). We determined if the Markov chains converged using the Gelman–Rubin statistic, R-hat (Gelman et al. 2004). For each parameter, we used 10 chains of 1,000 each and used a threshold of R-hat < 1.1 to indicate adequate sampling of the posterior distribution.

RESULTS

We used encounter histories for 323, 317, and 165 individual owls, for LAS, SIE, and SKC respectively, to model rate of population change. Median \hat{c} ranged from 1.09 to 1.18, suggesting some overdispersion but no serious lack of fit. We adjusted all variance estimates by site-specific estimates of c . The MCMC chains for all model parameters converged (R-hat < 1.1).

Estimates of mean $\hat{\lambda}$, $\tilde{\lambda}$, and median λ_{MCMC} were <1.0 for LAS and SIE, which suggested declining populations (Table 2). Estimates of median λ_{MCMC} were very close to $\hat{\lambda}$ and $\tilde{\lambda}$, although estimates of σ were $\geq 2.9 \times$ greater for MCMC than for the random effects estimate (Table 2).

Table 2. Estimates of mean annual rate of population change ($\bar{\lambda}$) and its temporal process standard deviation ($\hat{\sigma}$). Data are from California spotted owls on 3 study areas (Lassen [LAS], Sierra [SIE], and Sequoia Kings Canyon [SKC]) in the southern Cascades and Sierra Nevada, California, for 1990–2011. Random effect estimates are based on mixed effects models using a means model for time-specific estimates of rate of population change (λ ; random effect) with fixed effects for time-specific apparent survival (ϕ) and resight probability (\hat{p}), and Markov chain Monte Carlo (MCMC) estimates are based on a Bayesian approach to generate posterior distributions of annual rate of population change (λ_t) that is implemented in Program MARK.

Study area	Estimator ^a	$\bar{\lambda}$ ^b	SE	95% CI		$\hat{\sigma}$	95% CI	
LAS	FE	0.985	0.017	0.950	1.022			
	RE	0.987	0.011	0.967	1.008	0.018	0.013	0.026
	MCMC ^c	0.986		0.880	1.072	0.052	0.012	0.212
SIE	FE	0.993	0.017	0.959	1.027			
	RE	0.990	0.009	0.973	1.008	0.000	0.000	0.052
	MCMC	0.993		0.925	1.058	0.037	0.011	0.134
SKC	FE	1.016	0.021	0.976	1.057			
	RE	1.010	0.014	0.982	1.038	0.018	0.000	0.101
	MCMC	1.014		0.907	1.170	0.058	0.015	0.217

^a FE are fixed effects estimates, RE are random effects estimates from a mixed effects model, and MCMC are estimates from 20,000 simulations.

^b This estimate does not include the first 2 estimates and last estimate of λ_t because they were confounded or potentially biased.

^c MCMC estimates are the median and 2.5 and 97.5 percentiles from the posterior distribution, based on 20,000 simulations.

For LAS and SKC, the annual pattern of λ_t was the same between the fixed effects, random effects, and MCMC estimates, although values of $\tilde{\lambda}_t$ and $\lambda_{t\text{MCMC}}$ were closer to their estimated means compared to $\hat{\lambda}_t$ (Fig. 2). On average, the variance was greater for fixed effects estimates (Fig. 2a) than for random effects estimates from the mixed model (Fig. 2b). However, because we used RMSE for the random effects estimates, the variance for several estimates of $\tilde{\lambda}_t$ was greater than for the fixed effects $\hat{\lambda}_t$ (most noticeably, SIE 1992 and 1997 and SKC 1992, 2001, and 2002). The MCMC estimates of median $\lambda_{t\text{MCMC}}$ had smaller credible intervals than fixed effects CIs (Fig. 2a,c), and although some credible intervals were larger than the random effects CIs, they were more consistent (Fig. 2b,c).

The consecutive estimates of Δ_t reflected the temporal variation in λ_t , with $\tilde{\Delta}_t$ (Fig. 3b) not necessarily having less variance than $\hat{\Delta}_t$ (Fig. 3a and Table 3). Credible intervals of estimates of median $\Delta_{t\text{MCMC}}$ (Fig. 3c) started out smaller than CIs of $\hat{\Delta}_t$ (Fig. 3a), but grew to be approximately the same size by the end of the monitoring period. Over the 18-year study period, the different estimators of Δ_t indicated population declines of 21–22% for LAS and 11–16% for SIE,

and an increase of 16–27% for SKC, although CIs overlapped 1.0 for all estimators and study areas (Fig. 3 and Table 3). We found small but variable differences in magnitude between the estimators of Δ_t (1–9% absolute difference; Table 3). The size of standard errors for Δ_t were also variable with no particular pattern between the estimators (Table 3).

From the Bayesian approach, posterior distributions of $\Delta_{t\text{MCMC}}$ showed the 3 populations had different population trajectories (Fig. 4). The probability that any population declined by $\geq 30\%$ was low, varying from 0.24 for LAS to $< 0.01\%$ for SKC (Table 4). The probability of a minor decline was more substantial; we calculated a 0.69 and 0.40 probability LAS and SIE declined by $\geq 15\%$ (Table 4). The populations had variable probabilities of remaining stationary or increasing as well; SKC had a relatively high probability (0.82), whereas LAS and SIE had lower probabilities (0.07 and 0.22; Table 4).

DISCUSSION

For long-term monitoring data, Δ_t is an insightful metric for portraying the cumulative dynamics of a population across

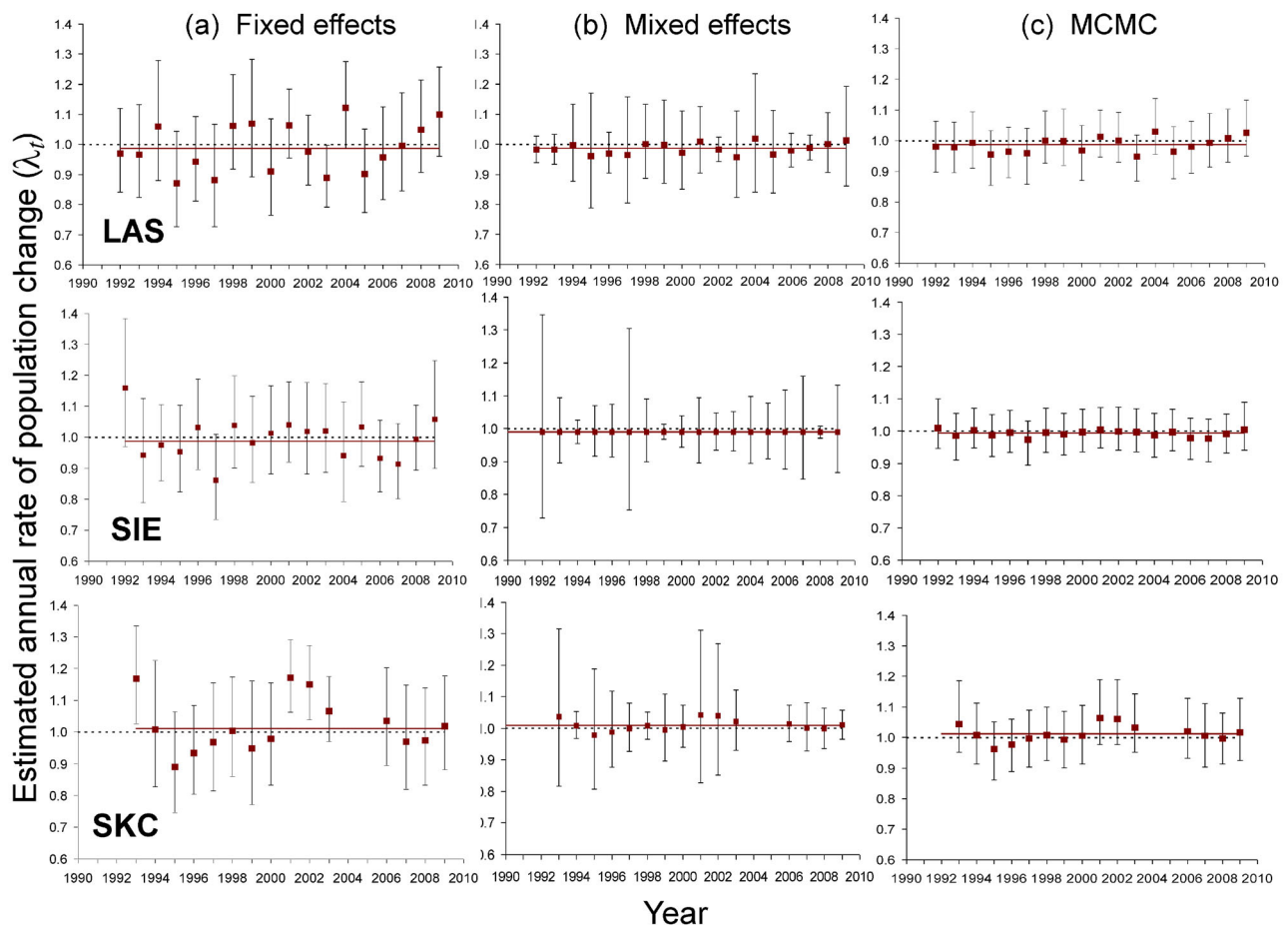


Figure 2. Estimates of annual rate of population change (λ_t) and 95% confidence intervals based on (a) fixed effects estimates, (b) random effects estimates from a mixed effects model, and (c) a Bayesian Markov chain Monte Carlo (MCMC) approach for California spotted owls on 3 study areas (Lassen [LAS], Sierra [SIE], and Sequoia Kings Canyon [SKC]) in the southern Cascades and Sierra Nevada, California, 1992–2010. We excluded the first 2 and last estimates because of confounding or potential bias. Solid lines indicate the best estimates of mean λ from a mixed effects model for each study area. We used the root means squared error for 95% confidence intervals for random effects estimates, and show the median and 2.5 and 97.5 percentiles for MCMC methods.

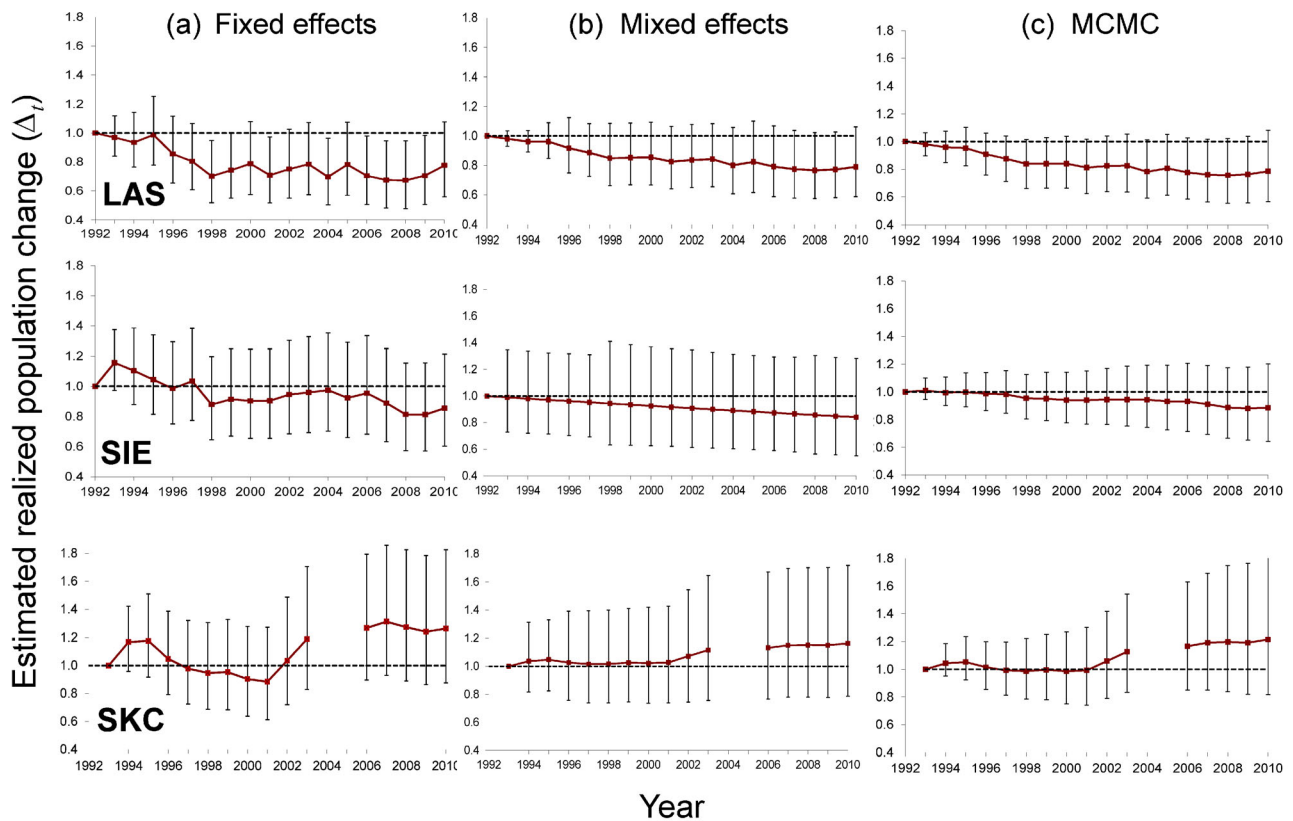


Figure 3. Estimates of realized population change (Δ_t) and 95% confidence intervals based on (a) fixed effects estimates, (b) random effects estimates from a mixed effects model, and (c) a Bayesian Markov chain Monte Carlo (MCMC) approach. Data are from California spotted owls on 3 study areas (Lassen [LAS], Sierra [SIE], and Sequoia Kings Canyon [SKC]) in the southern Cascades and Sierra Nevada, California, 1992–2010. We excluded the first 2 and last estimates because of confounding or potential bias. We used the root means squared error for 95% confidence intervals for random effects estimates, and show the median and 2.5 and 97.5 percentiles for MCMC methods. Realized population change is the proportion of the initial population size remaining each year.

medium and long time scales, such as 5–20 years. Further, if a Bayesian approach is used to estimate Δ_t , uncertainty is portrayed more completely through the posterior distribution, and the probability of decline can be calculated. For the California spotted owl, values of Δ_t are more immediately interpretable than $\bar{\lambda}$ because interpreting differences in $\bar{\lambda}$ from 1.0 (stationary) depends on the time period, and summarizing how much a population will change over a given time period can be difficult. For example, estimated median Δ_t MCMC was 0.78 for LAS, indicating that the population likely declined by 22% over the 18-year monitoring period. This is more directly interpretable than stating that estimated median λ MCMC was 0.986 over the 18-year period. Although $\bar{\lambda}$ is an important metric of population performance, we conclude that Δ_t is generally more interpretable, particularly for depicting population changes over longer time periods.

Because estimates of $\bar{\lambda}$ and Δ_t can be biased by sampling variation, we recommend using a hierarchical Bayesian MCMC or a random effects approach to estimate these parameters. Both these approaches remove sampling variation and generally have improved precision as a result of reduced mean square error, and equal or improved coverage when compared to fixed effects estimators (Burnham et al. 1987, Link 1999, Burnham and White 2002). However, in this study, we found random effects estimates

for Δ did not necessarily have improved precision. Although the average variance of the random effects point estimates of λ_t was less, 17–25% of the random effects estimates had variances that were 1.3–3.2 \times larger than their fixed effects counterparts. These large variances occurred when fixed effects estimates of λ_t were far from the random effects estimate; for example see SIE in 1992 and 1997 (Fig. 2). Because we added variances and covariances to estimates of Δ_t on the natural log scale, the few large variances propagated larger variances for subsequent estimates of Δ_t and, compared to their fixed effects counterparts, resulted in larger CIs for estimates of Δ_t for SIE and little reduction of the CIs for SKC.

We also found that the credible intervals on estimated median Δ_t MCMC increased through the monitoring period, such that by the end of the 18 years they were similar to fixed effects CIs. This was not initially intuitive because the credible intervals for estimated median λ_t MCMC were smaller than the CIs of the fixed effects $\hat{\lambda}_t$. However, for these data we found relatively high proportions of negative covariances between fixed effects estimates that reduced the overall additive variance of subsequent estimates of Δ_t . Thus, the data and type of estimator (e.g., when the estimates of interest are the result of additive or multiplicative combinations) will determine whether random effects estimates or estimates from Bayesian MCMC methods

Table 3. Estimates of overall realized population change (Δ_t) for 3 estimators, which are based on estimates of annual rate of population change (λ_t). Data are from California spotted owls on 3 study areas (Lassen [LAS], Sierra [SIE], and Sequoia Kings Canyon [SKC]) in the southern Cascades and Sierra Nevada, California, for 1990–2011. Overall realized population change is the proportion of the initial population size remaining at the end of the time period.

Study area	Estimator ^a	Δ_t ^b	SE	95% CI		CV (%)
LAS	FE	0.78	0.13	0.56	1.08	17
	RE	0.79	0.12	0.59	1.06	15
	MCMC ^c	0.78		0.57	1.08	
SIE	FE	0.85	0.15	0.60	1.21	18
	RE	0.84	0.18	0.55	1.28	22
	MCMC	0.89		0.65	1.20	
SKC	FE	1.27	0.24	0.88	1.83	19
	RE	1.16	0.23	0.79	1.72	20
	MCMC	1.22		0.82	1.84	

^a FE are fixed effects estimates, RE are random effects estimates from a mixed effects model, and MCMC are estimates from 20,000 simulations.

^b This estimate does not include the first 2 estimates and last estimate of λ_t because they were confounded or potentially biased.

^c MCMC estimates are the median and 2.5 and 97.5 percentiles from the posterior distribution, based on 20,000 simulations.

will show improved precision compared to fixed effects estimates.

However, for estimating and removing process variation, a Bayesian hierarchical MCMC approach is generally preferable to the random effects estimator for a more practical reason. For the random effects estimator, a minimum of 9,

and preferably 15, estimates are required (Efron and Morris 1977, Burnham and White 2002). Thus, the random effects method may be daunting in terms of required effort, or even impossible for many studies that lack sufficient spatial replication or temporal longevity. In addition, the variation in length of CI for random effects estimates (Fig. 2b compared to c) may not be readily interpretable and, therefore, may be less desirable.

Finally, we found that estimates of σ for mean λ were $\geq 2.9 \times$ larger from the Bayesian MCMC approach compared to random effects estimates (Table 2). This is presumably because inference is exact and MCMC incorporates all uncertainty in the modeled system (Royle and Dorazio 2008, Higgins et al. 2009, Kéry 2010), thus providing a more realistic estimate of the variance of mean λ and Δ . In contrast, random effects estimates are based on maximum likelihood and are approximations with asymptotic properties, which means $\hat{\sigma}^2$ may be biased low for small sample sizes (Gelman and Hill 2007). In addition, unbiased estimates of σ^2 may be harder to achieve when σ^2 is small or sampling variation is relatively large (Burnham and White 2002). For the California spotted owl case study, the sampling variation was likely large relative to σ^2 , resulting in estimates that were biased low or estimated as 0.

The biggest advantage of using a Bayesian approach and the posterior distribution of Δ_{MCMC} is that probability of decline can be estimated for a retrospective analysis, similar to how projection models or PVAs are used for a prospective analysis. In the past, evaluating whether the population declined typically relied on hypothesis tests with a null hypothesis of no decline. If the null hypothesis was not rejected, a retrospective power analysis was recommended to address the probability of accepting a false null or to determine the smallest detectable effect size, given study sample sizes (Taylor and Gerrodette 1993, Steidl et al. 1997, Loughheed et al. 1999). More recently, research has shown that retrospective power analyses are not appropriate to make inference about a non-significant result (Hoenig and Heisey 2001, Lenth 2001, Dixon and Pechmann 2005). Even if retrospective power was a valid approach, it still does not answer the real question of interest, what is the probability the population declined?

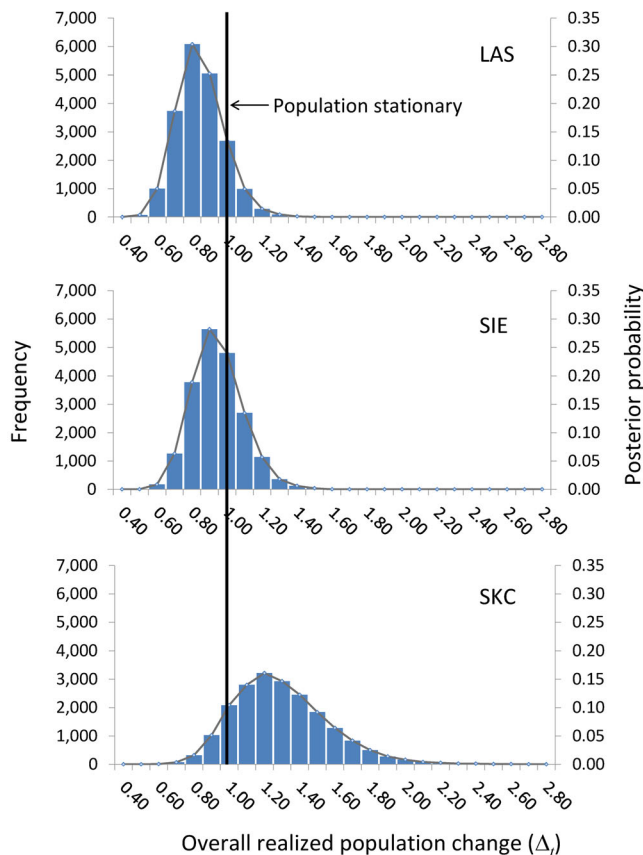


Figure 4. Estimated posterior distributions of overall realized population change (Δ_t) based on posterior distributions of λ_t from 20,000 Markov chain Monte Carlo (MCMC) simulations. Data are from California spotted owls on 3 study areas (Lassen [LAS], Sierra [SIE], and Sequoia Kings Canyon [SKC]) in the southern Cascades and Sierra Nevada, California, 1992–2010. We excluded the first 2 and last estimates because of confounding or potential bias. Overall realized population change is the proportion of the initial population size remaining at the end of the monitoring time period.

Table 4. Estimates of the probability of a population declining or increasing a given percentage or greater over a 17-year (SKC) or 18-year (LAS and SIE) monitoring period. Data are from California spotted owls on 3 study areas (Lassen [LAS], Sierra [SIE], and Sequoia Kings Canyon [SKC]) in the southern Cascades and Sierra Nevada, California, for 1990–2011. Probabilities are based on a posterior distribution of overall realized population change (Δ_t MCMC), which is the proportion of the initial population size remaining at the end of the monitoring time period.

Study area	Overall decline in population							
	$\geq 50\%$	$\geq 30\%$	$\geq 25\%$	$\geq 20\%$	$\geq 15\%$	$\geq 10\%$	$\geq 5\%$	$> 0\%$
LAS	0.00	0.24	0.39	0.54	0.69	0.80	0.88	0.93
SIE	0.00	0.07	0.15	0.26	0.40	0.54	0.67	0.78
SKC	0.00	0.00	0.01	0.02	0.04	0.07	0.12	0.18

Study area	Overall increase in population							
	$\geq 0\%$	$\geq 5\%$	$\geq 10\%$	$\geq 15\%$	$\geq 20\%$	$\geq 25\%$	$\geq 30\%$	$\geq 50\%$
LAS	0.07	0.04	0.02	0.01	0.01	0.00	0.00	0.00
SIE	0.22	0.14	0.08	0.05	0.03	0.01	0.01	0.00
SKC	0.82	0.76	0.68	0.60	0.52	0.45	0.38	0.10

Hypothesis testing is a less informative evaluation compared to using a Bayesian approach to generate a posterior distribution of Δ_t . If the hypothesis is rejected, or if the 95% CI of estimated $\bar{\lambda}$ or Δ_t covers 1.0, even if barely, we have no way of evaluating the probability of a decline or lack thereof. The posterior distribution of Δ_t has broader application beyond use with Pradel model estimates of λ_t ; it can be used for changes in density, abundance, or any population parameter of interest. For example, Gerrodette (2011) used data on the vaquita to examine the usefulness of frequentist versus other methods of hypothesis support for determining the change in abundance over an 11-year period. In a frequentist test for difference in abundance (i.e., overall change), $P = 0.38$, which did not mean abundance was equal, but rather that the data were not inconsistent with this hypothesis. However, although Gerrodette (2011) found a non-significant P -value when testing whether abundance had changed, Bayesian methods suggested a high probability of decline (0.88).

In our study, all CIs of estimates of $\bar{\lambda}$ and Δ_t covered 1. However, the probability of a decline varied widely between the 3 study sites, with LAS having a substantial probability of $\geq 10\%$ decline (0.80), whereas SKC only had a 0.07 probability of the same decline. The problem of having a large probability of decline, even when the null hypothesis of no decline is not rejected, is especially important for small populations. Smaller populations can drop, undetected by a null hypothesis approach, to low levels where the probability of extinction increases to unacceptably high levels due to demographic variability and stochastic environmental events (White 2000, Lande 2001, Morris and Doak 2002). The undetected decline is a problem faced by ecologists studying sensitive populations. For example, even with 12–19 years of northern spotted owl (*Strix occidentalis caurina*) monitoring data collected on 13 study areas, Anthony et al. (2006) found that although 12 of 13 areas had estimates of $\bar{\lambda}$ that were < 1 , they could only conclude that populations were declining on 4 areas based on whether confidence intervals overlapped 1. Further, using Δ_t Anthony et al. (2006) were only able to detect a change in population for 6 of the 12 areas with estimated $\bar{\lambda} < 1$. Clearly, using a Bayesian approach to

generate the posterior distribution of Δ_t MCMC is a powerful tool for estimating probability of change and provides a more refined measure of risk to monitored populations, which is particularly significant when monitoring results affect the legal status of a species and have broad management ramifications.

Although measures of central tendency, such as mean or median $\bar{\lambda}$ and Δ_t , have the discussed advantages, they also have limitations. If a population is growing ($\bar{\lambda} > 1$), managers cannot tell whether the growth is from internal recruitment or immigration. Likewise, if a population is declining, managers cannot determine whether the declines are due to deaths within the population or emigration. Thus, additional information on specific vital rates is necessary to understand what is driving λ and ultimately, the mechanisms driving population dynamics. Although $\bar{\lambda}$ and Δ_t are important metrics, they may not suffice for a full assessment of a population's health. Ultimately, the choice of appropriate monitoring parameters should be based on management and conservation objectives and information needs, given available resources.

MANAGEMENT IMPLICATIONS

Managers and regulators charged with formulating conservation strategies for focal wildlife species require robust information to support decisions that are often ecologically complex and socially and politically controversial. Realized population change is a useful metric for assessing population trends of focal species. Further, using Bayesian MCMC methods to generate a posterior distribution of Δ_t is a valuable tool for robustly estimating risk, or probability of declines, for retrospective analyses of monitored populations. In addition, compared to a random effects estimator, Bayesian MCMC methods also provide a more broadly applicable approach for estimating mean or median λ and σ because they do not have the restriction of requiring ≥ 10 estimates. Finally, Bayesian MCMC methods are also preferable for point estimates of λ_t because they are exact for the data and, with sampling variation removed, their credible intervals are smaller than CIs for likelihood estimates. The Bayesian approach to generate a posterior distribution of Δ_t ,

in particular, provides biologists, managers, and regulators with more insightful information on the probability of population changes that can better inform conservation assessments and strategies.

ACKNOWLEDGMENTS

We thank the many field technicians who carefully collected data over the 20 years of the study. In particular, we thank those who ran field projects, were involved for the long haul (>7 years), or helped with data analyses, including J. Blakesley, D. Shaw, G. Steger, R. Scherer, S. Sutton, K. Mazzocco, R. Kussow, M. Warner, and A. Peet, for their work in the field and out. Their efforts resulted in a high quality and long-term data set. We thank A. White, S. Bigelow, M. Bokach, K. Hubbert, and T. Biasioli, for helpful comments on drafts of the manuscript. We acknowledge the United States Department of Agriculture (USDA) Forest Service Region 5 for their long-term support of spotted owl population monitoring. Funding for our analysis was provided by the USDA Forest Service Region 5 and the Pacific Southwest Research Station.

LITERATURE CITED

- Anthony, R. G., E. D. Forsman, A. B. Franklin, D. R. Anderson, K. P. Burnham, G. C. White, C. J. Schwarz, J. D. Nichols, J. E. Hines, G. S. Olson, S. H. Ackers, L. S. Andrews, B. L. Biswell, P. C. Carlson, L. V. Diller, K. M. Dugger, K. E. Fehring, T. L. Fleming, R. P. Gerhardt, S. A. Gremel, R. J. Gutierrez, P. J. Happe, D. R. Herter, J. M. Higley, R. B. Horn, L. L. Irwin, P. J. Loschl, J. A. Reid, and S. G. Sovern. 2006. Status and trends in demography of northern spotted owls, 1985–2003. *Wildlife Monographs* 163.
- Blakesley, J. A., M. E. Seamans, M. M. Conner, A. B. Franklin, G. C. White, R. J. Gutiérrez, J. E. Hines, J. D. Nichols, T. E. Munton, D. W. H. Shaw, J. J. Keane, G. N. Steger, and T. L. McDonald. 2010. Population dynamics of spotted owls in the Sierra Nevada, California. *Wildlife Monographs* 174.
- Burnham, K. P., and D. R. Anderson 2002. *Model selection and multimodel inference*. Second edition. Springer-Verlag, New York, New York, USA.
- Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. *Design and analysis methods for fish survival experiments based on release-recapture*. American Fisheries Society, Monograph 5, Bethesda, Maryland, USA.
- Burnham, K. P., and G. C. White 2002. Evaluation of some random effects methodology applicable to bird ringing data. *Journal of Applied Statistics* 29:245–264.
- Dixon, P. M., and J. H. K. Pechmann 2005. A statistical test to show negligible trends. *Ecology* 86:1751–1756.
- Efron, B., and C. Morris 1977. Stein's paradox in statistics. *Scientific American* 236:119–127.
- Forsman, E. D., editor. 1983. *Methods and materials for locating and studying spotted owls*. U.S. Forest Service, General Technical Report, PNW-GTR-162, Portland, Oregon, USA.
- Franklin, A. B., D. R. Anderson, E. E. Forsman, K. P. Burnham, and F. F. Wagner. 1996. Methods for collecting and analyzing demographic data on the northern spotted owl. *Studies in Avian Biology* 17:12–20.
- Franklin, A. B., R. J. Gutiérrez, J. D. Nichols, M. E. Seamans, G. C. White, G. S. Zimmerman, J. E. Hines, T. E. Munton, W. S. LaHaye, J. A. Blakesley, G. N. Steger, B. R. Noon, D. W. H. Shaw, J. J. Keane, T. L. McDonald, and S. Britting. 2004. Population dynamics of the California spotted owl (*Strix occidentalis*): a meta-analysis. *Ornithological Monograph* 54.
- Gelman, A., J. A. Carlin, H. S. Stern, and D. B. Rubin. 2004. *Bayesian data analysis*. Second edition. Chapman & Hall/CRC, Boca Raton, Florida, USA.
- Gelman, A., and J. Hill 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, Cambridge, United Kingdom.
- Gerrodette, T. 2011. Inference without significance: measuring support for hypotheses rather than rejecting them. *Marine Ecology* 32:404–418.
- Gerrodette, T., and L. Rojas-Bracho 2011. Estimating success of protected areas for the vaquita, *Phocena sinus*. *Marine Mammal Science* 27:E101–E125.
- Higgins, J. P. T., S. G. Thompson, and D. J. Spiegelhalter. 2009. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society* 172:137–159.
- Hines, J. E., and J. D. Nichols 2002. Investigations of potential bias in the estimation of λ using Pradel's (1996) model for capture–recapture data. *Journal of Applied Statistics* 29:573–587.
- Hoienig, J. M., and D. M. Heisey 2001. The abuse of power: the pervasive fallacy of power calculations for data analysis. *American Statistician* 55: 19–24.
- Kéry, M. 2010. *Introduction to WinBUGS for ecologists*. Academic Press, Burlington, Massachusetts, USA.
- Lancia, R. A., W. L. Kendall, K. H. Pollock, and J. D. Nichols. 2005. Estimating the number of animals in wildlife populations. Pages 106–153 in C. E. Braun, editor. *Techniques for wildlife investigations and management*. The Wildlife Society, Bethesda, Maryland, USA.
- Lande, R. 2001. Incorporating stochasticity in population viability analysis. Pages 18–41 in S. R. Beissinger, and D. R. McCullough, editors. *Population viability analysis*. The University of Chicago Press, Chicago, Illinois, USA.
- Lenth, R. V. 2001. Some practical guidelines for effective sample size determination. *American Statistician* 55:187–193.
- Link, W. A. 1999. Modeling pattern in collections of parameters. *Journal of Wildlife Management* 63:1017–1027.
- Lougheed, L. W., A. Breault, and D. B. Lank. 1999. Estimating statistical power to evaluate ongoing waterfowl population monitoring. *Journal of Wildlife Management* 63:1359–1369.
- Moen, C. A., A. B. Franklin, and R. J. Gutiérrez. 1991. Age determination of subadult northern spotted owls in northwest California. *Wildlife Society Bulletin* 19:489–493.
- Morris, W. F., and D. F. Doak 2002. *Quantitative conservation biology: theory and practice of population viability analysis*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- Nichols, J. D., and J. E. Hines. 2002. Approaches for the direct estimation of λ and demographic contributions to λ using capture–recapture data. *Journal of Applied Statistics* 29:539–568.
- Pradel, R. 1996. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biometrics* 52:703–709.
- 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.
- Soulé, M. E. 1980. Thresholds for survival: maintaining fitness and evolutionary potential. Pages 151–169 in M. E. Soulé, and B. A. Wilcox, editors. *Conservation biology: an evolutionary-ecological perspective*. Sinauer Associates, Inc., Sunderland, Massachusetts, USA.
- Steidl, R. J., J. P. Hayes, and E. Schaubert. 1997. Statistical power analysis in wildlife research. *Journal of Wildlife Management* 61:270–279.
- 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.
- Thompson, W. L., G. C. White, and C. Gowen. 1998. *Monitoring vertebrate populations*. Academic Press, San Diego, California, USA.
- U.S. Forest Service. 2004. *Sierra Nevada forest plan amendment: final supplemental environmental impact statement*. Region 5. USDA Forest Service, Vallejo, California, USA.
- Ver Hoef, J. M. 1996. Parametric empirical Bayes methods for ecological applications. *Ecological Applications* 6:1047–1055.
- White, G. C. 2000. Population viability analysis: data requirements and essential analyses. Pages 288–331 in L. Boitani, and T. K. Fuller, editors. *Research techniques in animal ecology: controversies and consequences*. Columbia University Press, New York, New York, USA.
- White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study* 46 (Supplement):120–139.

White, G. C., K. P. Burnham, and D. R. Anderson. 2001. Advanced features of Program MARK. Pages 368–377 in R. Field, R. J. Warren, H. Okarma, and P. R. Sievert, editors. Land and people: priorities for the 21st century: proceedings from the 2nd international wildlife management congress. The Wildlife Society, Bethesda, Maryland, USA.

White, G. C., A. B. Franklin, and T. M. Shenk. 2002. Estimating parameters of PVA models from data on marked animals. Pages 169–190

in S. R. Beissinger, and D. R. McCullough, editors. Population viability analysis. University of Chicago Press, Chicago, Illinois, USA.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press, San Diego, California, USA.

Associate Editor: Andrew Kroll.