

METHODS FOR COLLECTING AND ANALYZING DEMOGRAPHIC DATA ON THE NORTHERN SPOTTED OWL

ALAN B. FRANKLIN, DAVID R. ANDERSON, ERIC D. FORSMAN,
KENNETH P. BURNHAM, AND FRANK W. WAGNER

INTRODUCTION

The collection of demographic data reflecting birth and death rates is important for understanding the life-history characteristics and population trends of the Northern Spotted Owl (*Strix occidentalis caurina*). Demographic parameters generally take the form of age-, sex-, and time-specific survival probabilities and fecundity rates. The first step in assessing the validity of inferences derived from such data is demonstration of an appropriate study design, as well as the field and analytical methods used. In addition, the methods used to collect demographic data should be repeatable, logistically feasible, and support the internal validity of the study design. Both study design and methods used to collect demographic data in the field must support the assumptions of models used to analyze those data for valid inferences to be made.

Demographic studies of Northern Spotted Owls reported in this volume are unique in several ways. First, these studies have been able to incorporate capture-recapture modeling approaches to estimate survival probabilities. These types of models have been rarely used with raptors, primarily because of sample size limitations (see Blondel et al. 1990 for reviews on different avian taxa). Second, standardized methods have been incorporated across all studies reported in this volume. This allowed for consistency in data collection and, hence, consistency in interpretation of results across the range of the owl. Third, the spatial extent and spatial replication of demographic studies allowed for broader inferences across the species' range in a meta-analysis (see Burnham et al. *this volume*).

The purpose of this paper is to present common elements of field and analytical methods used to estimate demographic parameters and population trends in Northern Spotted Owls as reported in this volume. We provide general descriptions of study areas, methods of data collection, and analytical methods used to estimate demographic parameters. Specific methods used in individual studies which depart from this general overview are described in the specific chapters pertaining to those studies. We also address important assumptions pertinent to the analytical models used and the allowable scope of inferences. Although confined to demographic

studies on the Northern Spotted Owl, we feel this paper will also provide a framework useful for similar research with other raptors. Terminology and symbols used throughout this volume are presented in the Appendix.

STUDY AREAS

This volume includes data from 11 study areas in northern California, Oregon and Washington (Fig. 1). Combined area of these study areas was 45,846 km² (Table 1). All of the study areas were primarily located on public lands administered by the U. S. Forest Service, U. S. Bureau of Land Management, and National Park Service. Inclusion of privately-owned lands in most study areas occurred incidentally as "inholdings" within public lands. However, most study areas on Bureau of Land Management districts included nearly equal mixtures of federal and non-federal lands. Inferences concerning Spotted Owl populations were restricted primarily to federally administered lands within the range of the owl except for the Bureau of Land Management studies (Coos Bay, Eugene BLM, Salem BLM, Roseburg BLM, and South Cascades/Siskiyou; see Fig. 1) which contained large amounts of private land. The 11 study areas encompassed about 27% of the 98,967 km² of federally administered land within the range of the Northern Spotted Owl and about 20% of the 230,690 km² range of the Northern Spotted Owl (USDA and USDI 1994).

Study area selection in all the owl demographic studies was based primarily on logistic considerations and objectives of funding agencies. As a result, study areas were not randomly or systematically distributed across the geographic range of the owl. Most studies were concentrated in the coastal mountains of California, Oregon, and Washington with fewer studies in the Cascade Mountains. We do not know if this uneven distribution of study areas caused bias in the overall evaluation of Spotted Owl populations across their range. However, the overall opinion of the research biologists at the Fort Collins workshop (see Gutiérrez et al. *this volume*) was that the broad representation of study areas from different forest types and management regimes was probably reflective of the overall condition of Spotted Owl populations on federal lands.

Of the 11 study areas, eight included intensively surveyed areas referred to as Density Study Areas (DSAs) (Table 1). DSAs were 204–1011 km² in size and established *a priori* with boundaries based on major topographical features and ownership boundaries. All habitats within DSAs were intensively surveyed for Northern Spotted Owls each year (Franklin *et al.* 1990), including at least two replicate surveys of each area each year. Minimum size for DSAs was established based on criteria outlined in Franklin *et al.* (1990) to minimize bias in density estimates due to edge effects. Maximum size for DSAs was dictated by the investigator's ability to survey adequately the entire area given funding and logistical constraints. Outside of the DSAs, no attempt was made to survey entire study areas each year. Rather, surveys focused on specific sites that had a history of occupancy by Spotted Owls. A "site" was defined as an area where Spotted Owls had exhibited territorial behavior in response to surveys on two or more occasions separated by one or more weeks within a given year. Individual sites were surveyed each year regardless of whether they were occupied by Spotted Owls. The use of the two types of survey design (DSAs versus site-specific surveys) reflected a trade-off between gathering additional information on movements and density in the DSAs and increasing sample size and regional scope in the larger study areas.

Two important assumptions regarding study area selection are: (1) study areas are representative of the larger area to which inferences are made, and (2) banded owls within a study area are representative of the population within that area. Whereas the first assumption can be objectively examined by comparison of landscape composition within and outside study area boundaries, the second assumption can not be

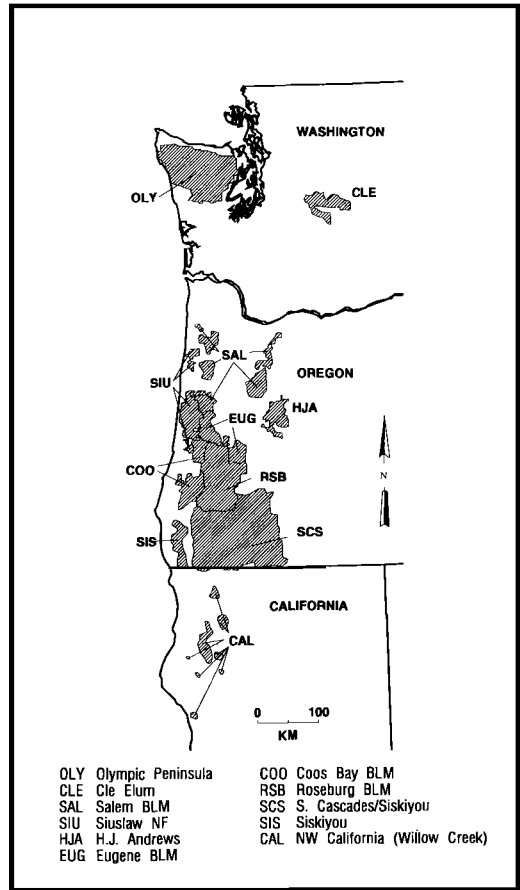


FIGURE 1. Location of 11 study areas used to estimate demographic characteristics of Northern Spotted Owls.

TABLE 1. DESCRIPTION OF THE 11 DEMOGRAPHIC STUDY AREAS FOR THE NORTHERN SPOTTED OWL

Study area (location)	Acronym	Study area size (km ²)	DSA size (km ²)	Years of banding
Willow Creek (NW California)	CAL	1,784	292	1985–1993
Roseburg (Oregon)	RSB	6,044	310	1985–1993
S. Cascades & Siskiyou Mts. (Oregon)	SCS	15,216	326	1985–1993
			1,011	
			300	
Salem BLM (Oregon)	SAL	3,249	491	1986–1993
H. J. Andrews (Oregon)	HJA	1,075	300	1987–1993
Olympic Peninsula (Washington)	OLY	8,145	355	1987–1993
Cle Elum (Washington)	CLE	1,763	196	1989–1993
Eugene BLM (Oregon)	EUG	2,082	273	1989–1993
Coos Bay (Oregon)	COO	2,477	—	1990–1993
Siuslaw NF (Oregon)	SIU	2,749	676	1990–1993
Siskiyou NF (Oregon)	SIS	1,262	—	1990–1993

tested. However, there are three lines of evidence which suggest that assumptions 1 and 2 were probably met. First, the 11 demographic studies encompassed over a quarter of the federal lands within the geographic range of the owl and were reasonably well-spaced throughout that range. This suggests that a large portion of the variability present within the owl's range was probably captured. Second, 3,616 territorial individuals (exclusive of 2,443 juveniles) were marked during these studies (Burnham et al. *this volume*) out of a known population of about 6,000 territorial individuals on federal lands in Washington, Oregon, and California (U.S. Dept. Interior 1992). While not all of these marked individuals were alive at the same time, a large portion of the range-wide population was probably marked, especially considering the high survival rates for ≥ 1 -year old owls (Burnham et al. *this volume*). Third, all research biologists, whose study areas are represented in this volume, agreed that their study areas were not grossly different from habitat amounts and configurations in the matrix surrounding their study areas.

METHODS

FIELD METHODS

The general design of the demographic studies, described in the following chapters, consisted of tracking marked individuals and their associated life history traits over time. Each study area was annually surveyed to locate both marked and unmarked owls. Once owls were located, they were individually marked using unique color bands and numbered U. S. Fish and Wildlife Service (USFWS) bands. Age, sex, and reproductive status of individuals were determined with standardized techniques as detailed below. Thus, for each year, individuals were located, assigned an age-class, identified, and assigned an estimate of their reproductive output.

Surveys

Annual surveys for Spotted Owls were conducted between 1 March and 1 September. Spotted Owls were located using vocal imitations or recorded playback of their calls to elicit responses (Forsman 1983). Both day and night surveys were used to locate owls (Forsman 1983). The primary method for surveying at night was calling for ≥ 10 min from a series of stations spaced 0.3–0.5 km along forest roads or trails. "Leapfrog" surveys were also used where two observers alternated walking along continuous transects. Owls were visually located by conducting calling surveys during the day to identify them and determine their reproductive status. Daytime surveys usually focused on areas where owls had previ-

ously responded during nighttime surveys or where owls had been located in previous years. Most daytime surveys were conducted while hiking cross-country.

Survey effort generally increased in the first few years of each study after which it leveled off. A site was assumed unoccupied if Spotted Owls were not detected after 3–6 night surveys, spaced ≥ 4 days apart, that completely covered 4–16 km² around locations where owls had been previously located during the day. In areas outside of DSAs, the area searched for owls depended on locations of adjacent pairs of owls and topography. Individuals were considered territorial if they exhibited vocal responses to surveys within the same site on ≥ 2 separate occasions within the same sampling period.

Determination of sex and age

With the exception of juveniles, the sex of owls > 1 year old was distinguished by calls and behavior. Males emit lower-pitched calls than females and do not incubate or brood (Forsman et al. 1984). Juveniles could not be accurately sexed until 1992 when some researchers began determining sexes of juveniles through examination of sex chromosomes in blood samples (Dvořák et al. 1992; see chapters on individual studies).

Spotted owls were aged by plumage characteristics (Forsman 1981, Moen et al. 1991) either visually, using binoculars, or when captured. Four age-classes were used: juvenile (J), 1-year old (S1), 2-year old (S2), and ≥ 3 -years old (A). Juveniles were fledged young-of-the-year that were characterized by gray, downy body plumage and retrices with triangular, tufted, white tips through their first summer. One-year old birds possess basic body plumage but are distinguished by tufted white tips on their retrices. Two-year old birds lose the tufts on the tips of the retrices but retain the triangular white tips until the retrices are first molted during the third summer of life. Thereafter, they become indistinguishable from ≥ 3 -year old owls that have retrices with rounded and mottled tips.

Capture and marking

Individuals were identified by initial capture, marking, and subsequent recapture or resighting of colored leg bands. Owls were captured with noose poles (Forsman 1983), snare poles, baited mist nets, or by hand. Handling time of captured owls was typically less than 20 minutes. Each owl was marked with a USFWS 7B numbered lock-on aluminum band placed on the tarso-metatarsus. A colored plastic leg band placed on the opposing tarso-metatarsus was used to identify ≥ 1 -year old birds in subsequent years with-

out recapture. Some researchers modified the color-band by adding a colored vinyl tab to increase the number of color combinations. Protocols for resighting color-marked individuals generally included blind trials where records of color combinations of owls located at a site in previous years were not examined until after a survey for that site was completed. If identification of color-marks was ambiguous, birds were recaptured and the number from the USFWS band recorded. Juveniles were marked with striped color bands indicating the year when they were captured. Cohort bands were replaced with unique color combinations when juveniles were recaptured in later years. The use of both USFWS and color bands allowed us to evaluate band loss. Only two cases of band loss were confirmed in over 6,000 marked individuals indicating the rate of band loss was very nearly zero. In some studies (see Forsman et al. *this volume*, Reid et al. *this volume*, and Wagner et al. *this volume*), radiotransmitters were used on a portion of the birds captured.

Estimation of reproductive output

We used field estimates of reproductive output (the number of young leaving the nest [fledging] per territorial female) as the basis for estimating fecundity. The average date of fledging (1 June) was considered the birth date. Once located during the day, owls were checked for reproductive activity by feeding them live mice (a procedure referred to as mousing) and observing how they behaved after mice were taken (Forsman 1983). Breeding Spotted Owls usually took such offered prey and carried it to the nest or fledged young. Non-reproductive owls either ate or cached the mice. Non-reproduction was inferred if an individual took ≥ 2 offered mice, and cached the last mouse taken, or a female did not have a well-developed brood patch during April–early May (the normal incubation period). In some cases, we also examined brood patches during the incubation period to determine if females were nesting. Territorial individuals were visited at least twice during the sampling period to determine the number of fledged young or to confirm non-reproduction using either the mousing or brood patch criteria on each visit. These techniques enabled us to characterize the reproductive output of territorial individuals as having 0, 1, 2, or 3 fledged young.

ANALYTICAL METHODS

Estimation of survival

Capture-recapture models were used to estimate age- and sex-specific survival for Northern Spotted Owls from the banding data. These models were statistical constructs used to estimate

the parameters of interest from the empirical data. The statistical analysis of capture-recapture or resight sampling data was based on the theory derived by Cormack (1964), Jolly (1965), and Seber (1965) and the simplifications and generalizations published since that time (e.g., Burnham et al. 1987, Clobert et al. 1987, Pollock et al. 1990). Lebreton et al. (1992) provided a comprehensive review of these theories, with examples. The capture history (Burnham et al. 1987: 28) for each owl for each age and sex class provided the basis for parameter estimation and hypothesis testing. The capture history matrix (X , described below in *Parameterization*) is a complete summary of the data. Estimators for all models used various summary statistics from this matrix. Owls were not included in the analysis during the time they carried back-pack transmitters because these types of transmitters may affect survival (Paton et al. 1991, Foster et al. 1992). However, owls fitted with 5-gram tail-mounted transmitters (mostly juveniles) on three study areas were included in the capture-recapture analyses because there was no evidence such small transmitters affected survival (E. Forsman, unpublished data). Owls with tail-mounted transmitters were considered recaptured only if they were located and their identity confirmed during normal calling surveys without the use of radio-telemetry. This ensured that recapture probabilities were not biased by differential detection of radio-marked birds. We assumed a 1:1 sex ratio at fledging for years where juveniles were not sexed. For each cohort of banded juveniles, the individuals subsequently recaptured were sexed and the remaining capture histories (representing individuals never captured) were arbitrarily assigned as males or females such that the total number of males and females was equal (Franklin 1992). The assumption of a 1:1 sex ratio was supported by data on juveniles sexed using chromosomal analysis (see Franklin et al. *this volume*).

Parameterization. The basic model for open mark-recapture populations is the Cormack-Jolly-Seber (CJS) model (Cormack 1964, Jolly 1965, Seber 1965) which considers only time specific survival probabilities (ϕ_i) and recapture probabilities (p_i) for k capture occasions (see Appendix for full summary of notation). These parameters are conditional on an animal being alive at the beginning of occasion i . Survival probabilities are estimated between occasion i and $i+1$ where $i = 1, 2, \dots, k-1$. Recapture probabilities are the probability that an animal alive on occasion i is captured (or recaptured) where $i = 2, 3, \dots, k$ (p_1 is not defined). In the case of the Spotted Owl, "capture" is defined as physical capture of individuals or resighting of their color bands with-

out physical capture. The p_i are nuisance parameters, but must be properly treated or estimators of survival probabilities will be biased. For example, let ϕ_1 be the survival probability between sampling occasions 1 and 2 and ϕ_2 be the survival probability between occasions 2 and 3. Therefore, for $k = 3$ capture occasions, the probability of various capture histories can be parameterized as: $\text{Pr}\{101\} = \phi_1 q_2 \phi_2 p_3$ (where $q_i = 1 - p_i$, the probability of not being captured on occasion i) for individuals captured on the first and third, but not the second, occasion; $\text{Pr}\{111\} = \phi_1 p_2 \phi_2 p_3$, for individuals captured on all three occasions; and $\text{Pr}\{110\} = \phi_1 p_2 \phi_2 q_3$, for individuals captured on the first and second, but not the third, occasion. Assuming the fates of individual animals were independent and that they have the same parameters (ϕ_i and p_i), the data on first recaptures from a single released cohort has a multinomial distribution. Releases from several cohorts are merely a product of these multinomial distributions. The likelihood function follows from this expression and is the basis for statistical inference.

Parameter estimation was based on Fisher's method of maximum likelihood. This method provided estimators of parameters that were asymptotically unbiased, efficient, and normally distributed. Variances and covariances were estimated using quasi-likelihood methods where appropriate (Wedderburn 1974, Cox 1983). These methods allow year- and age-dependent variation to be included in the variance of estimators from models that assume parameters were constant over years or age classes. For example, with the CJS model for a 3-occasion survey, four possible fates were possible for owls marked and released at occasion 1: \mathbf{X}_{111} , \mathbf{X}_{110} , \mathbf{X}_{100} , and \mathbf{X}_{101} . Then the likelihood function of the unknown parameters, given the data (\mathbf{X}) will be:

$$L(\phi_i, p_i | \mathbf{X}) \\ = C(\phi_1 p_2 \phi_2 p_3)^{x_{111}} (\phi_1 p_2 (1 - \phi_3) q_3)^{x_{110}} \\ (\phi_1 q_2 \phi_2 p_3)^{x_{101}} ((1 - \phi_1) q_2 (1 - \phi_2) q_3)^{x_{100}}$$

where C is the multinomial coefficient, involving the data, but not the parameters.

The analysis of multiple data sets provided extensive model building opportunities beyond the CJS model (Lebreton et al. 1992). Relationships of rates to external variables were modeled in this framework using the $\text{logit}(\theta)$ transformation which constrains $0 \leq \theta \leq 1$ as

$$\text{logit}(\theta) = \ln\left(\frac{\theta}{1 - \theta}\right)$$

where θ represents either ϕ or p . Lebreton et al. (1992) and Hosmer and Lemeshow (1989) pre-

sented rationales for use of this logit-link function. Survival probability (ϕ) and recapture probability (p) could then be modeled as a linear logistic function,

$$\text{logit}(\theta_w) = \beta_0 + \beta_1(w)$$

where w is an external or dummy variable. This approach allowed both categorical (e.g., sex, groups) and continuous (e.g., linear time) covariates to be employed in modeling ϕ or p . Lebreton et al. (1992) provided examples of these approaches and more extended theory. We used programs RELEASE (Burnham et al. 1987) and SURGE (Pradel et al. 1990) for analysis of mark-recapture data.

An important consideration with survival probabilities derived from capture-recapture estimators is that $1 - \phi = (\text{mortality rate} + \text{permanent emigration rate})$ whereas with true survival (S), $1 - S = \text{mortality rate only}$. In order for $\phi \cong S$, permanent emigration (E) must be negligible. Therefore, $\hat{\phi}$ must be adjusted when E is substantial to reflect true survival probabilities. Some studies (see Burnham et al. *this volume*, Forsman et al. *this volume*, Reid et al. *this volume*, and Wagner et al. *this volume*) used data from radiomarked owls to adjust some estimates of ϕ for E (see Burnham et al. *this volume* for methodology).

Model notation. Model nomenclature (see Appendix) followed Lebreton et al. (1992) and can be summarized as follows. The basic CJS model has time-specificity only, which can be expressed as $\{\phi_t, p_t\}$. This notation indicates a model whose parameters have unrestricted variation solely over time (occasions). If sex (s) or group (g , e.g., where $g = \text{study area}$) effects are added to the model, it can be written as $\{\phi_{s,t}, p_{s,t}\}$ where parameters exhibit unrestricted variation in time within each sex class, or $\{\phi_{g,t}, p_{g,t}\}$ where there is a group effect other than sex. The asterisk (*) indicates interactions (e.g., $s*t$ indicates interactions of sex with time, as well as both main effects). Therefore, a model examining study area effects, sex effects and unrestricted time variation for 1 age-class would be denoted as $\{\phi_{g*s*t}, p_{g*s*t}\}$. Age (a) can also be added as a factor and combined with sex, time and group effects in the same manner. The pure age model is denoted as $\{\phi_a, p_a\}$ where parameters vary by age only and, for k occasions, $a = 1, 2, \dots, k - 1$ ages. Models that include age restricted to classes are denoted as $a_1, a_2, a_3, \dots, a_n$ where n is the number of age-classes used. In models where p_i were age-specific for birds initially banded as juveniles, parameters are subscripted as a_n' where n' is the number of age-classes over which the restrictions are applied. Additive effects (i.e., no interactions considered) in models are denoted with a '+' instead

of a '*'. For example, the subscript $s + t$ indicates that the subscripted parameter varies over time for both sexes but that the difference between the two sexes is constant over time; plots of logit parameter estimates over time for the two sexes would be parallel. Parameters also can be constrained as linear functions of time, denoted as T . The resulting models are similar to the classical analysis of covariance where (1) parameters subscripted as T represents one intercept and one slope estimated for the parameter over time [$\text{logit}(\phi_i) = \beta_0 + \beta_1$ (time effects)]; (2) $s + T$ represents different intercepts for each sex with a common slope [$\text{logit}(\phi_i) = \beta_0 + \beta_1$ (sex effects) + β_2 (time effects)]; and (3) $s * T$ represents different intercepts and slopes for each sex [$\text{logit}(\phi_i) = \beta_0 + \beta_1$ (sex effects) + β_2 (time effects) + β_3 (sex effects * time effects)]. The $H_0: \beta = 0$ for estimated slope parameters is tested using a Wald test (Carroll and Ruppert 1988, Hosmer and Lemeshow 1989) of the form:

$$\chi^2 = \frac{\hat{\beta}^2}{\widehat{\text{var}}(\hat{\beta})} \quad \text{with 1 df}$$

Tests of assumptions. Goodness of fit tests (Pollock et al. 1985, Burnham et al. 1987) were used to assess the adequacy and utility of the basic CJS model, $\{\phi_i, p_i\}$. Burnham et al. (1987) outlined the requisite assumptions as: (1) capture, handling, and release do not affect survival; (2) the number released on occasion i is known exactly; (3) there is no band loss, and no bands are misread on capture or resighting; (4) all releases and captures of owls occur in relatively brief time intervals, and recaptured birds are released immediately; (5) any unknown emigration out of a study area is permanent (e.g., owls do not become unavailable for recapture by temporarily leaving the study area); (6) the fate of each individual owl, after any known release, is independent of the fate of any other owl; (7) data sets for the various ages, sexes, and areas are statistically independent; (8) statistical analyses of the sample data are based on an appropriate model; and (9) all owls of an identifiable class (e.g., age, sex) have the same survival and capture probabilities, by study area (i.e., parameters are homogenous within subclasses of individuals). Assumption (1) was tested using TEST 3 of program RELEASE which tests whether previously released individuals have the same future fates as newly released individuals. Assumption (2) and (3) probably were met with the Northern Spotted Owl data (see *Capture and marking* section). Assumption (4) was not strictly met in that the sampling period was relatively long (3–4 months). However, $\hat{\phi}_i$ is unbiased given that the shape of the temporal distribution of releases

(TDR) is constant from year to year and bias in ϕ_i is negligible when the medians of TDR are equal even though the distribution shapes may vary (inferred from Smith and Anderson 1987). This can be tested with Kruskal-Wallis tests (Sokal and Rohlf 1981) and multi-response permutation procedures (Mielke et al. 1981). Assumption (5) was untestable although it can be evaluated qualitatively. We tested assumptions (6), (7) and (9) using TEST 2 and 3 in program RELEASE (Burnham et al. 1987). TEST 3 is sensitive to heterogeneity in ϕ_i and p_i (assumption 9), short-term marking effects (e.g., assumption 1), and failure of assumption (6). TEST 2 also tests assumptions (6) and (1) as well as assumption (7) and for temporary emigration where an individual leaves the study area for at least one year and then returns. Assumption (8) can be properly evaluated through appropriate statistical model selection criteria and procedures, as described below.

Model selection. The most critical problem in the comprehensive analysis of capture-recapture data involving several year, age, and sex classes is selecting an appropriate model to describe the data (Burnham and Anderson 1992, Burnham et al. 1995a). A model should have sufficient structure and parameters to account for significant variability in the data or the resulting estimates will likely be biased. However, if the model has too much structure or too many parameters, then precision is lost unnecessarily. Proper model selection seeks a model that is fully supported by the particular data set and, thus, has enough parameters to avoid bias but not so many that precision is lost (Principle of Parsimony; see Burnham and Anderson 1992).

Model building started with a global model of $\{\phi_{s-a-t}, p_{s-a-t}\}$ for each study area (i.e., separate $\{\phi_{a+i}, p_{a+i}\}$ for each sex). We then used Akaike's Information Criterion, AIC (Akaike 1973, Anderson et al. 1994, Burnham et al. 1994, 1995a, 1995b), to objectively select an appropriate "best" model. This criterion was defined as

$$\text{AIC} = -2\ln(L) + 2K$$

where $\ln(L)$ is the natural logarithm of the likelihood function evaluated at the maximum likelihood estimates and K is the number of estimable parameters from that model. After selection of the best model using AIC, neighboring models of interest can be further investigated using likelihood ratio tests (McCullagh and Nelder 1983) as a further aid in selecting the best model for a particular data set. This procedure tests which of two nested model is best supported by the data using H_0 : the model with fewer parameters versus H_A : the model with more par-

rameters. For example, a significant P -value resulting from a test of H_0 : model $\{\phi_T\}$ versus H_A : $\{\phi_s\}$ indicates that $\{\phi_T\}$ should be retained as the best model, whereas a non-significant P -value would support retention of $\{\phi_s\}$. In the same manner, likelihood ratio tests can be used to test for specific effects, such as sex, time, and age, using identical models except that one includes the effect of interest and the other does not.

Estimation of fecundity

Age-specific fecundity (b_x) was defined for Northern Spotted Owls as the average number of female fledglings produced by a territorial female of age x (Caughley 1977). Age-specific fecundity was estimated using analysis of variance (ANOVA). Despite the integer nature of the individual data, sample sizes were sufficiently large to justify the assumptions of ANOVA. Data analysis was performed on reproductive output as the response variable using the general linear models (GLM) procedure in SAS (SAS Institute 1990) to test for significant age and time effects and interactions between effects within each study area. After analyses were performed, age-specific fecundity estimates (b_x) were calculated from estimates of mean reproductive output in each age-class by dividing those estimates by 2 to account for an assumed 1:1 sex ratio. In keeping with the 1:1 sex ratio assumption, standard errors of estimates for mean number of young fledged were divided by 2 (Goodman 1960) to estimate $SE(b_x)$.

In counting number of fledged young, we assumed that detection probabilities (analogous to p_i) of broods, and individual young within broods, after two visits was equal to 1.0. Three additional factors may introduce bias into estimates of fecundity. First, reproductively active individuals may have higher detectability than non-reproductively active individuals (e.g., Lundberg 1980). Therefore, fecundity would be biased high because fewer observations of 0 young would be recorded. Second, some fledged young experience mortality after fledging and before some pairs are checked for reproductive activity. In this case, the number of fledged young would be underestimated and, hence, biased downward. Third, some young are not banded immediately after they are counted. This would introduce a positive bias in the recruitment of first-year birds into the population ($b_x\phi_j$) because fledglings that die between the time they are counted and when the site is revisited again to band young are not included in the releases from which juvenile survival is estimated. It is unknown to what extent these competing biases cancel each other.

A cutoff date of 15 July has been proposed to deal with the second potential source of bias (Max et al. 1990). We examined the utility of this cutoff

date by testing for differences in reproductive output between 1 June–15 July and 16 July–1 September using data for all years and from all of the 11 studies. Prior to 1 June, pairs checked were either nesting or not reproductively active (i.e., had 0 young). Therefore, we compared only time periods of approximately equal lengths where fledged young were present. We found no significant difference (one-way ANOVA $F = 1.18$, $df = 1$, 3247, $P = 0.2778$) between mean reproductive output before ($N = 2824$) and after ($N = 512$) the 15 July cutoff date. In addition, there were no significant interactions between the two groups and years ($F = 1.56$, $df = 8$, 3247, $P = 0.1327$) or the two groups and studies ($F = 1.08$, $df = 9$, 3247, $P = 0.3775$). Therefore, all estimates of reproductive output collected over the sampling period of 1 March through 1 September were used in analyses.

Estimation of population trends

Lambda (λ), the annual rate of population change, was computed from the age-specific survival and fecundity estimates. In general, λ measures both direction in population trend ($\lambda = 1$ indicates a stationary population; $\lambda < 1$, a declining population; and, $\lambda > 1$, an increasing population) and magnitude of population change ($\lambda - 1$) (McDonald and Caswell 1993). For Northern Spotted Owls, we defined the target population to which we made inferences as the territorial, resident females. Although floaters (non-territorial unpaired individuals that do not breed) are known to exist in Spotted Owl populations (Franklin 1992), their influence on the regulation of Spotted Owl populations is unknown. In addition, floaters are undetectable using existing survey methods and, hence, are unmeasurable until they enter the territorial population. Therefore, we restricted our inferences to the territorial portion of the population whose parameters we were able to measure. Thus, λ answers the question, "What is the annual rate of population change for resident, territorial females given that estimated average survival probabilities and fecundity rates stay the same?"

From a management perspective, the research hypothesis of interest is $\lambda < 1$ versus the null that the population is either stationary or increasing ($\lambda \geq 1$), here a 1-tailed test. The form of this test is

$$z = \frac{1 - \hat{\lambda}}{\widehat{SE}(\hat{\lambda})}$$

where $z \approx N(0,1)$.

Leslie (1945, 1948) provides the matrix theory to allow the computation of λ from knowledge of only the age-specific fecundity and survival

probabilities (see Lefkovitch 1965, Usher 1972, Caswell 1989, Noon and Biles 1990). We believe use of a simple Leslie matrix model was an appropriately parsimonious approach because it incorporated only those parameters that we could precisely estimate. We used only the female component of the population to estimate λ . The Leslie-Lefkovitch matrix allows λ to be computed from the characteristic polynomial of this matrix. For the full matrix model which includes all 4 age-classes, this matrix has the form:

$$\begin{bmatrix} \phi_J b_{S1} & \phi_{S1} b_{S2} & \phi_{S2} b_A & \phi_A b_A \\ \phi_J & 0 & 0 & 0 \\ 0 & \phi_{S1} & 0 & 0 \\ 0 & 0 & \phi_{S2} & \phi_A \end{bmatrix} \quad (1)$$

which assumes a birth-pulse population, a post-breeding census, and a time interval of 1 year (Noon and Sauer 1992). The individual studies in this volume included only those age-classes in such matrices for which parameters were estimated. For example, a two age-class matrix was used if parameter modeling procedures indicated the data only supported estimates of survival and fecundity for two age-classes. Lambda can be estimated as the dominant eigenvalue of (1) through matrix eigenanalysis (Caswell 1989) or through numerical search procedures for the unique, positive, real root of the characteristic equation of (1):

$$\lambda^4 - \lambda^3(\phi_A - \phi_J b_{S1}) + \lambda^2 \phi_J (b_{S1} \phi_A - \phi_{S1} b_{S2}) + \lambda \phi_J \phi_{S1} (\phi_A b_{S2} - \phi_{S2} b_A) = 0 \quad (2)$$

Maximum likelihood estimates of the survival and fecundity parameters were used in (1) and (2) to estimate λ . Estimation of survival and fecundity estimates depended on the selected model used in estimating those parameters (see chapters on individual studies). For example, if a model with separate estimates for each year (e.g., ϕ_t or ϕ_T for survival estimates) was selected, an average was estimated as the arithmetic mean (see Jolly 1982) and its standard error computed. If a time invariant model was selected (e.g., model ϕ), the single estimate and its standard error was used. Precision of these estimates included any year-to-year and unaccounted for age-specific variability in the parameters as well as proper estimates of sampling variability. The $SE(\lambda)$ was estimated using the delta method (Seber 1982, Alvarez-Buylla and Slatkin 1994), including the sampling covariance terms for survival estimates. Sampling covariances between fecundity and survival estimates were zero because the two variables were statistically independent. The adequacy of the delta method was verified using a parametric bootstrap method (Efron 1982,

Alvarez-Buylla and Slatkin 1994) assuming a beta distribution for ϕ and a log-normal distribution for \hat{b} .

Four key assumptions are critical to estimating and interpreting λ estimated from the matrix model (Goodman, 1968, Caswell 1989, Noon and Sauer 1992, McDonald and Caswell 1993). First, we assumed that classifying Northern Spotted Owls into four age-classes was more appropriate than other properties relevant to an individual's fate, such as size or developmental stage. Second, we assumed there was no age-dependency in survival or fecundity in birds that were ≥ 3 -year old age-class. Third, use of the matrix model assumes age-specific survival and fecundity rates remain constant over time and are density-independent, and fourth, the population is assumed at a stable age-class distribution where each age-class changes by λ over time. Parsimonious model development dictated the first two assumptions given sample sizes and available data. Concerning the third assumption, there is, in practice, temporal variation in the demographic parameters; our estimates reflect $E(\theta)$ over years for use in the Leslie matrix. Thus, $\hat{\lambda}$ approximates an average estimate over the period of years, even if the estimates of survival and fecundity vary over time. The last assumption becomes largely irrelevant when inferences about λ are limited to projection (what would happen) rather than forecasting (what will happen) (Keyfitz 1972, Caswell 1989:19–20). For the studies in this volume, estimates of λ are properly interpreted as the average annual rate of population change ($E(\lambda)$) for Northern Spotted Owls if conditions during the period of investigation were maintained indefinitely. In other words, the λ estimated from the age-specific survival and fecundity rates would occur if the conditions responsible for shaping the parameter estimates remain unchanged indefinitely. Under this interpretation, the population would eventually reach a stable age distribution. This interpretation differs from one involving forecasting which would state that estimates of λ will apply under future conditions regardless of how they may affect parameter estimates. Alternatively, our estimates of λ can be viewed as integrating environmental effects on survival and fecundity rates into a single index which quantifies the suitability of the environment for a population at a given time and place (McDonald and Caswell 1993).

The estimates of λ referred to the resident population, containing several age classes, and their recruitment. Immigration into the study populations is not estimated by mark-recapture, nor used by the Leslie approach to λ . Estimation of survival probabilities under the mark-recapture framework is conditional on first capture and,

therefore, does not measure immigration. In addition, the parameter "immigration" does not appear in the matrix model; only estimates of survival and fecundity are needed.

Estimates of λ could be biased low if the juvenile survival ($\hat{\phi}_J$) estimate used did not approximate S_J because $1 - \hat{\phi}_J$ includes a significant emigration component (see Raphael et al. *this volume*). To estimate this emigration component, we used the other parameter estimates (assuming that they were unbiased with respect to the true parameters), set $\lambda = 1$ (a stationary population), and computed the juvenile probability required to obtain $\lambda = 1$ (denoted as $S_{J|\lambda=1}$). In the same manner, we computed the emigration rate (E) required to have $\hat{\phi}_J = S_{J|\lambda=1}$ as:

$$E_{|\lambda=1} = 1 - \frac{\hat{\phi}}{S_{J|\lambda=1}}$$

In this way, potential biases due to permanent emigration of juveniles from study areas were addressed in terms of a stationary population.

CONCLUSIONS

In analyzing demographic data for Northern Spotted Owls, we used a direct empirical approach in estimating population parameters and trends. This approach included objective, parsimonious model selection procedures to estimate parameters and their standard errors. A parsimonious approach was extended to the use of the Leslie matrix for estimating population trends.

As a simplification of reality (as all models are), the Leslie matrix approach contained only those parameters that we could precisely estimate and that were supported by the available data. Thus, our approach was driven solely by the available data that could be objectively analyzed in an appropriate statistical framework in contrast to other modeling approaches (e.g., Lamberson et al. 1992, McKelvey et al. 1992)

that necessarily assume mechanisms, such as dispersal behavior and birth and death processes, for which we have little or incomplete information. We acknowledge that our approach is not definitive in describing trends in Spotted Owl populations. However, we view our approach as an initial step in an iterative series of more sophisticated approaches. As understanding of the mechanisms and processes governing Spotted Owl populations increases, more sophisticated, data-based modeling procedures can be supported. However, we believe that appropriate parameter estimation procedures and a parsimonious approach to integrating parameter estimates is essential to any approach attempting to estimate trends in Spotted Owl populations.

SUMMARY

We present field and analytical methods used to estimate life history parameters and population trends for Northern Spotted Owls in the Pacific Northwest. Demographic characteristics were examined on 11 study areas distributed through northern California, Oregon, and Washington. Survival probabilities were estimated using mark-recapture estimators with data from annual surveys of individually color-marked owls. Fecundity rates were estimated using direct counts of fledged young. We discuss model selection procedures, tests of assumptions, and potential sources of bias inherent in the estimation techniques. We outline two approaches to examine trends: (1) testing for time-dependency in life history traits; and (2) estimating the annual rate of population change (λ) from demographic parameters and testing λ against the null hypothesis that the population is stationary ($H_0: \lambda = 1$).

ACKNOWLEDGMENTS

We thank W. R. Clark, M. G. Raphael, T. M. Shenk, J. Reid, G. C. White, and two anonymous reviewers for their thoughtful comments on this manuscript. Their input greatly improved its content.

Key words: capture techniques, demographic techniques, fecundity, Leslie matrices, mark-recapture estimators, Northern Spotted Owl, *Strix occidentalis caurina*, survival estimators.