United States
Department of
Agriculture

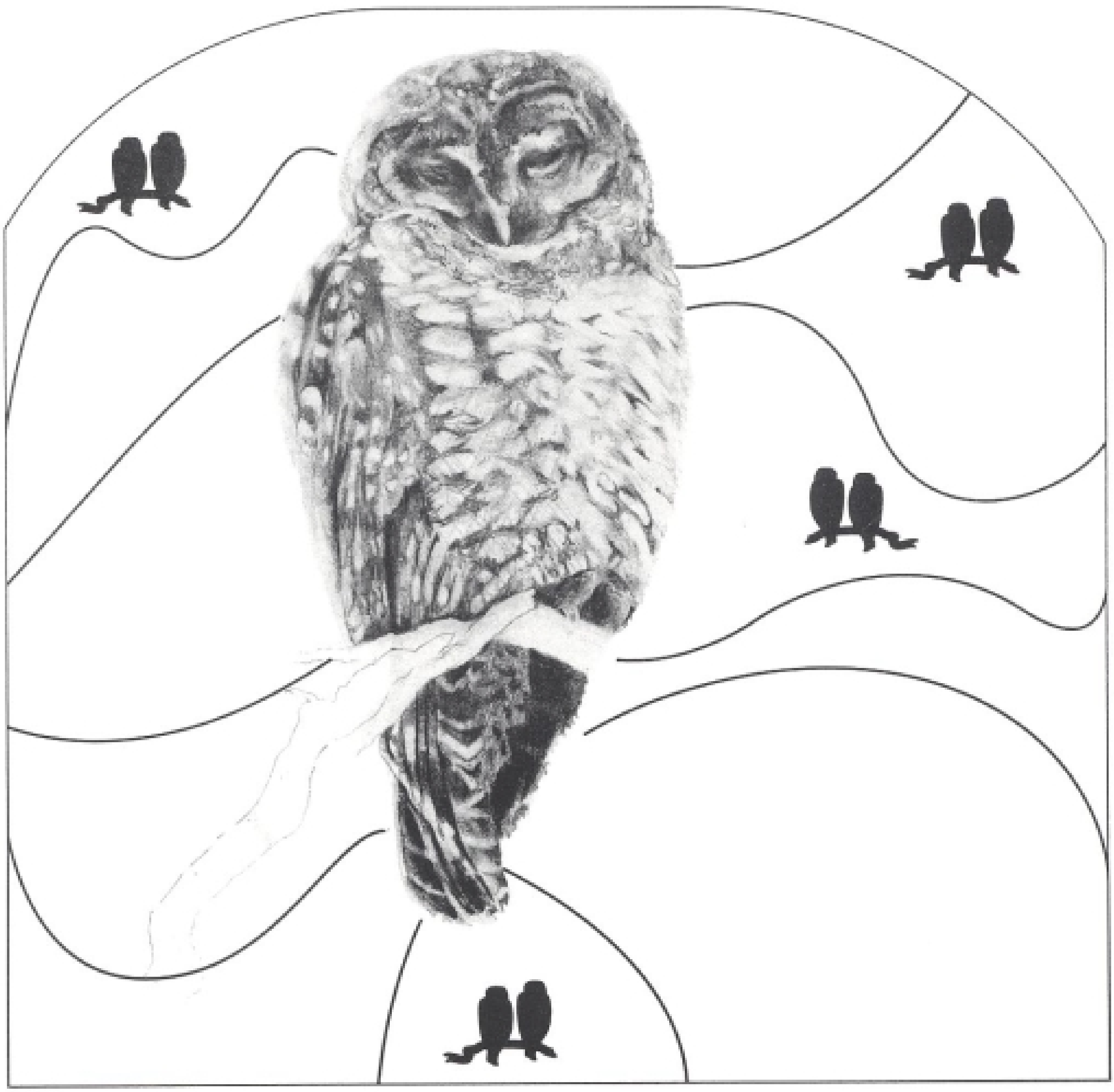
Forest Service

**Pacific Southwest
Research Station**

General Technical
Report PSW-124

Estimating the Occupancy of Spotted Owl Habitat Areas by Sampling and Adjusting for Bias

David L. Azuma James A. Baldwin Barry R. Noon



Azuma, David L.; Baldwin, James A.; Noon, Barry R. 1990. **Estimating the occupancy of spotted owl habitat areas by sampling and adjusting for bias**. Gen. Tech. Rep. PSW-124. Berkeley, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; 9 p.

A basic sampling scheme is proposed to estimate the proportion of sampled units (Spotted Owl Habitat Areas (SOHAs) or randomly sampled 1000-acre polygon areas (RSAs)) occupied by spotted owl pairs. A bias adjustment for the possibility of missing a pair given its presence on a SOHA or RSA is suggested. The sampling scheme is based on a fixed number of visits to a sample unit (a SOHA or RSA) in which the occupancy is to be determined. Once occupancy is determined, or the maximum number of visits is reached, the sampling is completed for that unit. The resulting data are summarized as a set of independent Bernoulli trials; a zero (no occupancy) or one (occupancy) is recorded for each unit. The occupancy proportion is the sum of these Bernoulli trials divided by the sample size. The bias adjustment estimates this occupancy proportion for the estimated number of units on which a pair of owls was present but not detected. The bias adjustment requires the recording of the number of the visit during which occupancy was first detected. The distributional assumptions are checked with five different sets of data.

Retrieval Terms: spotted owl, SOHA, sampling scheme, bias adjustment, occupancy

The Authors:

DAVID L. AZUMA is a mathematical statistician assigned to the Station's Cumulative Watershed Effects/Inland Fishery Research Unit in Berkeley. **JAMES A. BALDWIN** is supervisory mathematical statistician in the Statistics and Computer Services Group in Berkeley. **BARRY R. NOON** is supervisory research wildlife biologist and project leader of the Timber Management/Wildlife Interactions in Northern California Forest Types Research Unit, with headquarters in Arcata, Calif.

Acknowledgments:

We acknowledge the work of Florence N. David, Forest Service, who originated the bias adjustment. We are grateful to Cordon Could, California Department of Fish and Game, for providing the survey data of Yosemite, Sequoia, and Kings Canyon National Parks. And we thank the field crews in California, Oregon, and Washington for their hard work in collecting the data on Spotted Owl Habitat Areas and Randomly Sampled Areas.

Publisher:

Pacific Southwest Research Station
P.O. Box 245, Berkeley, California 94701

September 1990

Estimating the Occupancy of Spotted Owl Habitat Areas by Sampling and Adjusting for Bias

David L. Azuma James A. Baldwin Barry R. Noon

CONTENTS

INTRODUCTION

The spotted owl (*Strix occidentalis*) in Washington, Oregon, and California is represented by two subspecies: the northern spotted owl (*S. o. caurina*), which occupies Washington, Oregon, and the Klamath physiographic province in California, and the California spotted owl (*S. o. occidentalis*), which occupies the Sierra Nevada physiographic province and areas of Southern California as well. Both subspecies are believed to occur most frequently in large areas of mature or old-growth coniferous forests. Most suitable spotted owl habitat in California, Oregon, and Washington is found on Federal lands with 67 percent of the habitat located on National Forests (USDA 1988a). There is concern that loss and fragmentation of suitable spotted owl habitat because of timber harvest activities threaten the continued existence of the species. The Forest Service must respond to this potential threat because the Forest Service is mandated by the National Forest Management Act (1976) to maintain viable populations of all vertebrate species on National Forest lands. The requirement to "Manage fish and wildlife habitats to maintain viable populations of native and desired nonnative vertebrates" (National Forest Management Act 1976, 36 CFR 219.19) translates into the management goal of maintaining the population number and distribution needed to ensure continued existence of the species.

Population viability requires an adequate amount and distribution of suitable habitat for pairs of owls to ensure a specified level of population persistence over some period of time. Suitable habitat must be distributed so as to provide for the levels of genetic and demographic interchange needed to assure adequate numbers to minimize the risk of extinction. Suitable spotted owl habitat on National Forest lands in California, Oregon, and Washington is found on lands both reserved from timber harvest and open to harvest (nonreserved). To meet the distributional requirements of the species, it is necessary to protect, in nonreserved lands, areas of suitable habitat that provide for both the needs of reproductive pairs and a link between the suitable owl habitat found in reserved lands. Each of the designated "islands" of suitable owl habitat is referred to as a Spotted Owl Habitat Area (SOHA). Together with SOHA units on reserved lands, these units form the SOHA network. By addressing issues of distribution and number of owls, the Forest Service has proposed the SOHA network as its solution to maintaining the viability of spotted owls. The network sets distances among SOHAs so as to provide the owls a high likelihood of successful dispersion between SOHAs.

An important assumption of the management plan for the spotted owl is that the proportion of network SOHAs occupied by pairs of owls will remain constant over time. Another assumption is that owl populations outside the network in nonreserved lands will decline because of habitat loss, but owl populations will continue to provide future alternatives for extending the network if the occupancy rate within the network declines. The Forest Service has begun a monitoring program to check the validity of these assumptions.

It is not feasible to census all owls on Forest Service lands, so a subset of the network SOHAs and a number of randomly sampled areas (RSAs) are monitored to estimate their occupancy by owls. RSAs are 1000-acre polygons located at random within National Forest lands conditioned on the polygon area having at least 25 percent canopy cover by trees. The size of SOHAs varies geographically to reflect variation in spotted owl home range sizes. SOHA sizes range from 1000 acres in the Klamath and Sierra Nevada Provinces to 3000 acres on the Olympic Peninsula. The basic assumption underlying monitoring is that occupancy rates of the two populations of sample units (network SOHAs and RSAs) are a valid index for the trends in population size and reproductive rate. The trends in occupancy rates of the SOHAs and RSAs will be compared in an effort to evaluate the efficiency of the network. This paper outlines the sampling design for estimating occupancy rates and monitoring the SOHA network and RSAs.

The Forest Service may change its policy on Spotted Owls and the SOHAs may become obsolete; however, this sampling scheme can be applied to other situations in which occupancy of specific units is a reasonable surrogate variable for trends in population size and reproductive rates.

POPULATION

The estimation methods described below can be applied separately to three distinct subpopulations of land area within California, Oregon, and Washington (Max and others). The subpopulations are defined by joint consideration of two factors: SOHAs and timber harvest land classification (reserved or non-reserved). Subpopulations are defined as: (1) all SOHAs on reserved lands, (2) all SOHAs on non-reserved land, and (3) all non-reserved land outside of SOHAs. The sampling units for the first two subpopulations are the SOHAs, although SOHAs on reserved land are not sampled. The 1000-acre RSA polygons encompass the third and portions of the others. The trends in occupancy rates of the RSAs should be indicative of owl populations in general, whereas the trends in occupancy rates in the SOHAs represent the owl populations in the network.

Within the three subpopulations, land area can be further classed by physiographic province (Franklin and Dyrness 1973). This classification is usually necessary because SOHA sizes, and thus detection probabilities, vary by physiographic province (USDA 1988a).

PARAMETERS

The basic variable to be observed for each sample unit is "occupancy." Occupancy is defined to be a pair of owls, as described in the Spotted Owl Inventory and Monitoring Handbook (USDA 1988b). The following sampling scheme and bias adjustment could also be used to estimate the proportion of units with "presence" (single owl) or "reproductive pairs of owls" (pair with young confirmed). The parameter likely to be estimated most frequently is the proportion of occupied sample units. Changes in occupancy estimates of the subpopulations can be monitored through time as an index of population trend within a specific subpopulation.

BASIC SAMPLING SCHEME

The basic scheme for sampling the SOHA network assumes there are N units of which N_1 have owls and N_2 do not ($N_1 + N_2 = N$). We will sample n of these N units at random. A unit will be visited until occupancy is established (Spotted Owl Inventory and Monitoring Handbook, USDA 1988b) or until six visits are completed. For each year, a binomial proportion, or the proportion of units occupied, is estimated.

We define the following terms and notation:

- N — Number of potential units (population size).
- N_1 — Actual number of occupied units.
- N_2 — Actual number of unoccupied units, $N = N_1 + N_2$.
- P — The true proportion of occupied units.
- $\hat{P}$ — Estimate of proportion of occupied units.
- s — Maximum number of visits to one sample unit ($s = 6$, in this study).
- d — Maximum distance we want the estimate to be away from P, the true parameter.
- $1 - \alpha$ — Confidence coefficient (proportion of times the estimate will be no farther than d away from P).
- v — True mean number of visits required to determine occupancy at an occupied unit given a maximum of s visits.
- $\hat{v}$ — Estimated mean number of visits to determine occupancy at an occupied unit given a maximum of s visits.
- P — The probability of determining occupancy at a unit during a single visit, given that the unit is occupied.
- n — The number of units sampled out of N.
- n_1 — Number of sampled units that were occupied.

The number of network SOHAs in both California and Forest Service Region 6 (Oregon and Washington) is approximately 450. Because n units are randomly selected, the number of SOHAs with the attribute "occupied" has a hypergeometric probability distribution. In contrast, there are approximately 5000 RSAs in California and approximately 5000 for Oregon and Washington. Given that n is small relative to N for this population, the number of occupied sample RSA units can be approximated with a binomial probability distribution. Both the RSA and the SOHA populations are assumed to be fixed over the monitoring period.

If we assume that the number of visits is large enough to ensure determination of occupancy on units where pairs occur, then the simple ratio of the number of units occupied to the number sampled ($\hat{P}$) is an unbiased estimate of the proportion of units with pairs of spotted owls. The variance of this proportion is estimated by

$$\widehat{Var}(\hat{P}) = \hat{P}(1-\hat{P})\frac{(N-n)}{N}\cdot\frac{1}{(n-1)}.$$

The problem with this method of estimation is that the assumption that s visits are sufficient to ensure detection of a pair of owls (given they are present) on a unit may not be valid. The estimate would be biased if there were some probability of failing to determine occupancy on an occupied unit within s visits. As a consequence, the simple ratio of the number of occupied units to the number of units sampled would tend to underestimate the true proportion (P).

A THEORETICAL BIAS ADJUSTMENT

The estimate of the occupancy proportion can be adjusted to account for the failure to detect owls on occupied units. To adjust for bias we assume: (1) that there is a constant probability p of determining occupancy for each visit to an occupied unit; and (2) that each visit to a unit is an independent and identical Bernoulli trial that can result in one of two possible outcomes: occupancy or not; and (3) that we can have an unlimited number of total visits. Under these assumptions, the probability that X, the number of visits required to one unit to determine occupancy, is equal to x is given by the geometric distribution:

$$Pr(X = x) = p(1-p)^{x-1}, \quad x = 1, 2, 3, etc.$$
$$= 0, \quad \text{otherwise.}$$

One of the assumptions for this probability distribution is that there are an unlimited number of visits. If the number of visits is limited to a maximum of s visits, then the probability that X, given that $0 < X \le s$, is equal to x is given by:

$$Pr(X = x | 0 < X \le s) = \frac{p(1-p)^{x-1}}{1-(1-p)^s}, \; x = 1, 2, 3, \ldots, s$$
$$= 0, \quad \text{otherwise.}$$

This truncated geometric distribution is shown in *figure* 1 for various values of p.

Given the above assumptions, both the sample size estimates and the proportion of occupied units can be corrected for this bias. Given that there exists a known probability $q(q = 1 - p)$ that a pair occupies a sample unit (SOHA or RSA) but is missed during a single visit, the expected value of our estimate n_1/n is (N_1/N) $(1 - q^s)$. In this instance n_1/n will tend to underestimate $P = N_1/N$ by a factor of $(1 - q^s)$. If p is close to 1 or s is large, then this factor is near 1 and no adjustment is necessary. Otherwise, we can adjust for the negative bias by dividing n_1/n by $(1 - q^s)$.

Our estimate of the number of occupied units is as follows: $\hat{P} = (n_1/n)/(1 - q^s)$ with a variance of

$$P(1-P)\frac{N-n}{N-1}\cdot\frac{1}{n}+P\cdot\frac{q^s}{1-q^s}\cdot\frac{1}{n}\ (appendix).$$

In practice, the value of q (or p) will not be known and must be estimated from the data. However, if occupancy is undetermined after s visits, then whether a pair was present and missed or whether the unit was unoccupied is unknown. Therefore, we use just those observations from units for which occupancy was determined (where $0 < X \leq s$), to estimate q.

The expected value of X given that $0 < X \leq s$ is:

$E(X|0 < X \leq s) = 1/p - sq^s/(1 - q^s)$

We substitute $\hat{v}$, the estimated mean of the positive X's from the sample, on the left side of the equation and solve for $\hat{p}$. Then we can use the estimate of p to adjust n_1 / n to

$$\tilde{P} = (n_1/n)/\left[1-(1-\hat{p})^s\right] = (n_1/n)/(1-\hat{q}^s) = (n_1/n)f(\hat{v}),$$

where $f(v)$ is a function of the average number of visits to determine occupancy, and end up with a reasonable estimator for $P = N_1 / N$. *To* solve for p requires an iterative process. An alternative is to create a table (see *table 1*) for different values of p and v, and interpolate between values of v to determine $\hat{p}$.

A limitation of this bias adjustment is that it depends critically on the assumptions for the geometric random variable. Any deviation from these assumptions can necessitate a different bias correction. Failure of the observed frequency distribution of initial occupancy determinations to be closely aligned with the theoretical, expected distributions (*fig. 1*) suggests that the geometric model is inappropriate; the empirical data cannot rigorously verify our assumptions. However, it is possible to get an acceptable fit between the observed and expected distribution of first determinations even if the assumptions of the geometric model are invalid.

Table 1--*The multiplicative adjustment factor, f(v), for $n_1$ln as a function of v, the average number of visits to first detection. The associated value of q=(1-p), the probability of not determining occupancy during a single visit, given that the unit was occupied*

v	$f(v)$	q
1.90	1.1059	.50
2.05	1.0285	.55
2.21	1.0489	.60
2.24	1.0543	.61
2.27	1.0602	.62
2.30	1.0567	.63
2.33	1.0738	.64
2.37	1.0816	.65
2.40	1.0901	.66
2.47	1.1097	.68
2.50	1.1210	.69
2.53	1.1333	.70
2.57	1.1469	.71
2.60	1.1619	.72
2.63	1.1783	.73
2.67	1.1965	.74
2.70	1.2165	.75
2.73	1.2387	.76
2.77	1.2633	.77
2.80	1.2907	.78
2.83	1.3212	.79
2.87	1.3553	.80
2.90	1.3936	.81
2.93	1.4368	.82
2.97	1.4858	.83
3.00	1.5415	.84
3.03	1.6055	.85
3.10	1.7656	.86
3.13	1.8671	.87
3.16	1.9880	.88
3.19	2.1342	.90
3.23	2.3141	.91
3.26	2.5404	.92
3.29	2.8328	.93
3.32	3.2245	.94
3.35	3.7749	.95

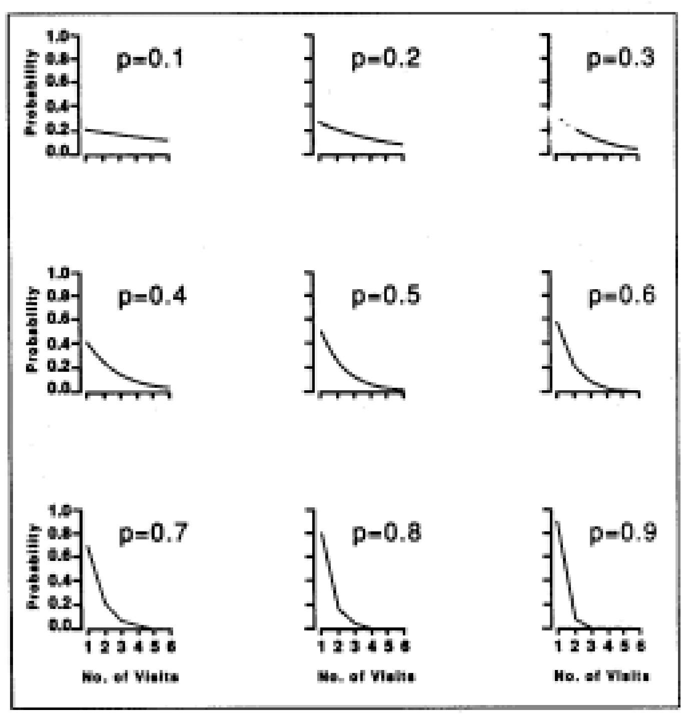


Figure 1-Probability functions of the number of visits required to first see an owl, given an owl was observed within 6 visits for various values of p (p is the probability of observing an owl on an occupied unit on any single visit).

ESTIMATING PROPORTION OF OCCUPIED UNITS AND ITS VARIANCE

Suppose n units out of N were sampled up to s times, occupancy was determined on n_1 units, and the mean number of visits required of those n_1 units was v. Then an unbiased estimate of

P (the proportion of occupied units) is given by $\hat{P} = \frac{n_1}{n} \cdot f(v)$,

where an approximate value for $f(v)$ (a multiplicative adjustment for n_1/n) is obtained from *table 1*:

The estimate of the variance of the estimate is found by

$$\widehat{Var}(\hat{P}) = \hat{P}(1-\hat{P}) \cdot \frac{N-n}{N} \cdot \frac{1}{n-1} + \hat{P}\frac{\hat{q}^s}{1-\hat{q}^s} \cdot \frac{1}{n},$$

where $\hat{q}$ is found in *table 1*.

The calculation of $Var(\hat{P})$ is shown in the *appendix*. The *appendix* also proposes an estimated variance for $\tilde{P}$.

The approximate 95 percent confidence interval for the proportion of occupied units is $\hat{P} \pm 1.96\sqrt{\widehat{Var}(\hat{P})}$.

***Example*:** Suppose we sampled 60 units out of 150 and found 45 occupied units that required an average of 2.43 visits. If the maximum number of visits, *s*, equals 6, then the estimate of the occupied units is $\hat{P} = \frac{45}{60} \cdot 1.0995 = 0.825$.

The estimated variance is

$$\widehat{Var}(\hat{P}) = 0.825\,(1-0.825)\frac{150-60}{150-1} \cdot \frac{1}{60} + 0.825\frac{.67^6}{1-.67^6} \cdot \frac{1}{60}$$

$= 0.002821$.

Approximate 95 percent confidence intervals for $\hat{P}$ are $0.825 \pm 1.96(0.002821)^{1/2} = 0.825 \pm 0.104$ (0.721, 0.929).

ESTIMATING SAMPLE SIZE

If we assume that the number of visits is large enough to ensure that if a pair is in the area it will be detected, the following procedure can be used to obtain a sample size. To set sample size, we want our estimate, n_1/n, to be no farther away than d from $P = N_1/N$, $100(1 - \alpha)$ percent of the time. To select a sample size we will need to specify values for d and α. The sample size is calculated by the following equation:

$n = n_0/[1 + (n_0 - 1)/N]$ (correction for sampling from a finite population), where $n_0 = Z^2(\tilde{N}_1/N)(\tilde{N}_2)/d^2$, $\tilde{N}_1$ and $\tilde{N}_2$ are initial guesses for N_1 and N_2, and Z is the value of the standard normal curve that cuts off 100α percent in the tails.

However, if there is a concern that not all pairs are detected at occupied units and we are willing to assume that the probability of detecting a pair on any particular visit is a constant p (Note: $q = 1 - p$), then the following becomes the sample size formula:

$$n = (1 + c)\, n_0 / [1 + (n_0 - 1)/N]$$

n_0 is given by: $n_0 = Z^2\hat{P}(1-\hat{P})/d^2$

and c is given by: $c = \frac{1}{1-P} \cdot \frac{q^s}{1-q^s}$

Notice that this estimate of sample size is bigger than the previous estimate by a factor of $1 + c$.

If we estimate $\hat{P}$, the sample sizes can be approximated from *tables 2* and *3*. The determination of sample size depends not only on the proportion of occupied units but on the average number of visits required to determine occupancy.

Calculating sample size involves obtaining numbers from two different tables. *Table 3* lists the required sample sizes with $1 - \alpha = 0.95$, $s = 6$, $v = 1$, and $d = 0.01, 0.05, 0.10$, and 0.20. (Additional tables with other settings can be constructed). Setting $v = 1$ means that if a pair exists on a unit, then it is certain to be observed on the first visit. Therefore, *table 2* shows the *minimum* required sample sizes. *Table 3* shows the percentage increase in sample size for several values of v (average number of visits required to determine occupancy at an occupied unit).

Example: Suppose we have 150 potential units, the occupied proportion of units is 0.50, the average number of visits required to provide this estimate was 2.5 out of a maximum 6 visits, and 95 percent of the time we want the estimate to be no farther than 0.10 from 0.50. First, the value from *table 2* is 59. The percentage increase from *table 3* with $v = 2.5$ is 24 percent. The required sample size is therefore $59 \cdot 1.24 = 73$ units out of 150 units.

ANALYSIS OF DISTRIBUTIONAL FIT

We need to check whether the number of visits to determine occupancy fits a truncated geometric distribution. Five sets of data are available to check the distributional assumptions: the California SOHAs and RSAs, the Oregon and Washington SOHAs and RSAs, and survey data from Yosemite, Sequoia, and Kings Canyon National Parks. All of these data sets have $s = 6$. Failure of the empirical frequency distribution of first determinations to align with the theoretical distributions suggests that the geometric model is inappropriate. A failure would be most likely to occur because of a varying probability of detection from one visit to the next at a given unit. Another possibility might be varying detection probabilities across units or among pairs of owls.

We used the average number of visits to determine occupancy to estimate the value p. Once p is estimated, it is possible to calculate the probability of determining occupancy on any visit, 1 through s. The theoretical and the empirical frequency distributions can then be compared. *Table 4* gives the observed values for each of the five data sets.

The usual X^2 statistic is calculated for each of the five sets of data by:

$$\sum_i^6 \frac{(E_i - O_i)^2}{E_i},$$

where E_i = expected quantity and O_i = observed quantity (Mosteller and Rourke 1973). Expected, theoretical probabilities are given by:

$Pr\ (X = x) = pq^{s-1}/(1-q^s)$, the truncated geometric distribution.

The expected observations are then computed as the product of the sample size and the theoretical probability for a given visit number. *Table 4* shows the contribution to the X^2 statistic from each visit over the five sets of data. The statistic is compared to a critical value of 11.1, with 5 degrees of freedom, at the 5 percent level of significance.

Table 2--*Sample sizes with v = 1 (certain to observe an owl in an occupied unit)*

	Proportion of occupied units (P)									
N[1]	.50	.45 .55	.40 .60	.35 .65	.30 .70	.25 .75	.20 .80	.15 .85	.10 .90	.05 .95
d[2] = 0.01										
40	40	40	40	40	40	40	40	40	40	39
50	50	50	50	50	50	50	50	50	49	49
60	60	60	60	60	60	60	59	59	59	58
70	70	69	69	69	69	69	69	69	69	67
80	79	79	79	79	79	79	79	79	78	77
90	89	89	89	89	89	89	89	88	88	86
100	99	99	99	99	99	99	98	98	97	95
110	109	109	109	109	109	108	108	108	107	104
120	1I9	119	118	118	118	118	118	117	116	113
130	128	128	128	128	128	128	127	127	125	113
140	138	138	138	138	138	137	137	136	135	130
150	148	148	148	147	147	147	146	146	144	139
160	157	157	157	157	157	157	156	155	153	147
170	167	167	167	167	167	166	165	164	162	156
180	177	177	177	176	176	176	175	174	171	164
190	186	186	186	186	186	185	184	183	180	172
200	196	196	196	196	195	195	194	192	189	180
d = 0.05										
40	36	36	36	36	36	35	35	33	31	26
50	44	44	44	44	43	43	42	40	37	30
60	52	52	52	51	51	50	48	46	42	33
70	59	59	59	58	58	56	55	52	47	36
80	66	66	66	65	64	63	61	57	51	38
90	73	73	73	72	71	69	66	62	55	41
100	80	79	79	78	77	74	71	66	58	42
110	86	85	85	84	82	80	76	71	62	44
120	92	91	91	90	88	85	81	75	65	46
130	97	97	96	95	93	90	85	78	67	47
140	103	103	102	100	98	94	89	82	70	48
150	108	108	107	105	103	99	93	85	72	49
160	113	113	112	110	107	103	97	88	74	50
170	1I8	118	117	115	112	107	101	91	77	51
180	123	122	121	119	116	111	104	94	78	52
190	127	127	126	123	120	115	107	97	80	53
200	132	131	130	127	124	118	111	99	82	54

[1]N = number of potential units (population size).
[2]d = maximum distance from true proportion 95 percent of the time.

(Continued)

Table 2--*Sample sizes with v = 1 (certain to observe an owl in an occupied unit) (continued)*

	Proportion of occupied units (*P*)									
		.45	.40	.35	.30	.25	.20	.15	.10	.05
N[1]	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
$d^2 = 0.10$										
40	28	28	28	28	27	26	24	22	19	13
50	33	33	33	32	31	30	28	25	21	14
60	37	37	37	36	35	33	31	27	22	14
70	41	41	40	39	38	36	33	29	23	15
80	44	44	43	42	40	38	35	31	24	15
90	47	46	46	45	43	40	37	32	25	15
100	49	49	48	47	45	42	38	33	26	16
110	52	51	50	49	47	44	40	34	26	16
120	54	53	52	51	48	45	41	35	27	16
130	55	55	54	53	50	47	42	36	27	16
140	57	57	56	54	51	48	43	36	28	16
150	59	58	57	55	53	49	44	37	28	16
160	60	60	59	57	54	50	45	38	29	16
170	62	61	60	58	55	51	45	38	29	17
180	63	62	61	59	56	52	46	39	29	17
190	64	64	62	60	57	52	47	39	29	17
200	65	65	63	61	58	53	47	40	30	17
d =.015										
40	21	21	20	20	19	18	16	14	11	7
50	23	23	23	22	21	20	18	15	12	7
60	25	25	25	24	23	21	19	16	12	7
70	27	27	26	25	24	22	20	17	13	7
80	28	28	27	26	25	23	21	17	13	7
90	29	29	28	27	26	24	21	18	13	8
100	30	30	29	28	27	24	22	18	13	8
110	31	31	30	29	27	25	22	18	14	8
120	32	31	31	30	28	25	22	19	14	8
130	32	32	31	30	28	26	23	19	14	8
140	33	33	32	31	29	26	23	19	14	8
150	33	33	32	31	29	27	23	19	14	8
160	34	34	33	31	29	27	23	19	14	8
170	34	34	33	32	30	27	24	19	14	8
180	35	34	34	32	30	27	24	20	14	8
190	35	35	34	32	30	28	24	20	14	8
200	35	35	34	33	31	28	24	20	14	8

[1]*N*= number of potential units (population size).
[2]*d*= maximum distance from true proportion 95 percent of the time.

That there was no significant difference between the theoretic-cal and empirical frequency distributions does not necessarily mean that the assumptions of independence of visits and a fixed probability of detection are valid. Small to moderate variations in detection probabilities among units and among visits at a unit could result in empirical data that closely fit a theoretical distribution based on the bias adjustment. Unfortunately, the empirical data do not allow a rigorous test of assumptions. However, the fit between the empirical and theoretical distributions gives no indication to discredit the assumptions.

DISCUSSION

The basic assumption underlying monitoring is that the occupancy status of owl pairs in sample units (SOHAs or RSAs) within a subpopulation is strongly correlated with population levels in that subpopulation. Given this assumption, trends in occupancy status of owl pairs through time are indicative of trends in the population. SOHA units were chosen on the basis of known or historical owl presence and spacing of available

Table 3—*Percentage increase in number of sampled units*

	Proportion of occupied units																		
v[1]	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
3.1	81	85	90	96	102	109	118	128	139	153	170	191	219	255	306	383	510	766	1531
3.0	57	60	64	68	72	77	83	90	98	108	120	135	155	181	217	271	361	542	1083
2.9	41	44	46	49	52	56	61	66	72	79	87	98	112	131	157	197	262	394	787
2.8	31	32	34	36	39	42	45	48	53	58	65	73	83	97	116	145	194	291	581
2.7	23	24	25	27	29	31	33	36	39	43	48	54	62	72	87	108	144	217	433
2.6	17	18	19	20	22	23	25	27	29	32	36	40	46	54	65	81	108	162	324
2.5	13	13	14	15	16	17	19	20	22	24	27	30	35	40	48	60	81	121	242
2.4	9	10	11	11	12	13	14	15	16	18	20	23	26	30	36	45	60	90	180
2.3	7	7	8	8	9	10	10	11	12	13	15	17	19	22	27	33	44	67	133
2.2	5	5	6	6	7	7	8	8	9	10	11	12	14	16	20	24	33	49	98
2.1	4	4	4	4	5	5	5	6	6	7	8	9	10	12	14	18	24	36	71
2.0	2	3	3	3	3	3	3	4	4	5	5	6	6	8	9	11	15	23	45
1.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

[1]v= Mean number of visits required to first observe an owl at an occupied site. If the increase in sample size makes $n>N$, use $n = N$.

Table 4-*The number of units where occupancy was determined on visit number i (i = 1 to 6) for each set of data and the contribution to the X^2 statistic for each visit number*

Visit	R-6[1] SOHA[3]	R-5[2] SOHA	Survey Data	R-6 RSA[4]	R-5 RSA
1	36	33	8	10	7
2	19	17	8	9	5
3	16	12	3	3	5
4	6	7	1	1	3
5	0	7	1	4	2
6	2	5	2	2	2
Total	79	81	23	29	24
Average visits	2.00	2.42	2.35	2.52	2.75
Estimated p	.46	.37	.39	.35	.29
	Contribution to x^2				
1	0.04	0.03	0.22	0.09	0.12
2	0.06	0.48	0.86	0.49	0.08
3	2.42	0.03	0.07	0.57	0.24
4	0.01	0.12	0.61	1.34	0.01
5	3.17	0.77	0.07	2.12	0.01
6	0.05	1.05	1.81	0.41	0.22
Total	5.75	2.48	3.64	5.02	0.68

[1]R-6 refers to Forest Service Region 6 (Oregon and Washington).
[2]R-5 refers to Forest Service Region 5 (owls in California).
[3]SOHA = Spotted Owl Habitat Area.
[4]RSA = randomly sampled area.

habitat. Therefore, we would expect, a *priori*, that SOHA units should initially have higher occupancy rates than RSAs. Comparison of trends between subpopulations, particularly SOHA units versus RSAs, should provide insights into the effectiveness of the SOHA network to maintain viable populations of spotted owls.

Caution is advised in the interpretation of any observed differences in trend between network SOHAs and RSAs. Higher occupancy rates in the network are not necessarily indicative of a correct SOHA size or spacing. Many factors, in addition to size and spacing, affect occupancy rate. These include availability of suitable habitat outside of the SOHA, prey availability, local predator populations, and juxtaposition to other, non-reserved units with owls. The significance of these additional factors will not be revealed by the monitoring program described here.

An important aspect of monitoring is that the population to be sampled is clearly defined and does not change over the course of the monitoring period. If the population of units changes, for example as a consequence of the addition or deletion of SOHAs, then it is unclear about what population inferences are being made. Annual estimates of occupancy could still be made, but it would make little sense to compute a trend. For example, if there were significant changes in the network among years, then any change in occupancy proportion may reflect changes in the sampled units, and may be unrelated to an actual change in owl population size.

Ideally, we would like to know the occupancy status of each SOHA and RSA. However, because of costs, only a random sample of SOHAs and RSAs has been selected for monitoring, and this sample will remain constant for the first 5 years of monitoring. The accuracy of the annual estimate of occupancy in the second and subsequent years will be less than if a new random sample were drawn each year, but retaining the same sample will provide the most reliable estimate of trend in occupancy.

The size of sample units (SOHAs) is not constant across all physiographic provinces (USDA 1988a). Therefore, direct comparisons of occupancy proportions across provinces that differ in size are precluded. However, it is possible to compare trend in occupancy estimates across provinces.

In order to estimate sample size we had to assume a distributional model for the number of occupied units. The population of SOHAs was assumed to be fixed through time and finite. The RSAs were assumed to be drawn from an essentially infinite population. Given these assumptions, the number of occupied SOHA units should follow a hypergeometric distribution, and the number of occupied RSAs a binomial distribution.

The existence of the bias adjustment is not a sufficient reason to reduce the number of visits to a unit. If the probability of determining occupancy is 0.2 for any visit, then truncation of the geometric distribution distributes 27 percent of the probability mass among the first six visits. If the number of visits is further reduced, the amount of redistributed probability mass is increased. Deviations from the assumptions become increasingly critical with an increase in the amount of redistributed probability mass.

CONCLUSIONS

There is some likelihood that a sample unit may be misclassified as unoccupied. To the extent that this occurs, estimates of the occupancy proportion would be negatively biased, and calculations of sample size would need to be adjusted. In this paper we proposed a bias adjustment based on the average number of visits to detect an owl pair, a variable that, we assumed, followed a truncated geometric distribution. The assumptions of independence between visits and the constant probability of detection were critical for employing the bias adjustment. These assumptions made it possible to think, collectively, about the six visits to a sample unit as a simple Bernoulli trial. A strict set of sampling protocols has been implemented (Spotted Owl Inventory and Monitoring Handbook, USDA 1988b) to make these assumptions as reasonable as possible. These protocols were used to estimate the average number of visits to first detection and an adjusted occupancy proportion for each of five independent sets of data. Using these data, no significant differences were found between the theoretical and empirical frequency distributions of first detections.

REFERENCES

Franklin, Jerry F.; Dyrness, C.T. 1973. Natural vegetation of Oregon and Washington. Gen. Tech. Rep. PNW-8. Portland, OR: Pacific Northwest Forest and Range Experiment Station, Forest Service, U.S. Department of Agriculture; 417 p.

Max, Timothy A.; Souter, Ray A.; O'Halloran, Kathy A. Statistical estimators for monitoring spotted owls in Oregon and Washington in 1987. Unpublished draft supplied by author.

Mosteller, Frederick; Rourke, Robert E.K. 1973. Sturdy statistics. Memo Park, CA:Addison-Wesley; 395 p.

National Forest Management Act 1976. 36 CFR 219.19.

USDA Forest Service. 1988a. Final supplement to the Environmental Impact Statement for an Amendment to the Pacific Northwest Regional Guide. Volumes 1 and 2. Washington, DC: US Government Printing Office; 1988-0-591-057/80700.

USDA Forest Service. 1988b. Spotted owl inventory and monitoring handbook. Pac. Northwest Reg., Portland, Oreg., and Pac. Southwest Reg., San Francisco, Calif. February 1988. Addendum of May 1988. 18 p.

APPENDIX

The algebra used to construct the occupancy estimator and an estimator of its variance are described below. Of N sites, N_1 have spotted owls and N_2 do not. A simple random sample of size n is taken without replacement. For convenience, let the N_1 occupied sites be indexed 1,2,. . . , N_1 and the remaining sites be indexed N_1 + 1,. . . , N.

Let α_i = 1 if site i is selected and α_i = 0 otherwise. The probability that site i is selected is $\Pr(\alpha_i = 1) = n/N$. And because we are sampling with replacement, the probability that sites i and j (where $i \neq j$) are both selected is

$$\Pr(\alpha_i = \alpha_j = 1) = n(n-1)/N(N-1).$$

Also let (β_i = 1 if an owl is observed on site i and (β_i = 0 otherwise. The probability that an owl is observed on any of the N_1 occupied sites is Pr(($\beta_i = 1) = 1 - q^s$ where q is the probability of not detecting an owl on a single visit to an occupied site and s is the maximum number of visits.

Because the visits to each site are independent, the probability that owls are observed at any two occupied sites i and j (where $i \neq j$) is just the square of $1 - q^s$.

The number of observed owls is

$$n_1 = \sum_{i=1}^{N} \alpha_i \beta_i = \sum_{i=1}^{N_1} \alpha_i \beta_i$$

The expectation of n_1 is

$$E(n_1) = E\left(\sum_{i=1}^{N_1} \alpha_i \beta_i \right)$$

$$= N_1 \cdot \Pr(\alpha_i = 1)\Pr \cdot (\beta_i = 1)$$
$$= N_1 (n/N)(1 - q^s)$$
$$= nP(1 - q^s)$$

Our unbiased estimator of P can be $\hat{P} = (n_1/n)/(1 - q^s)$.

To find the variance we first find the expectation of the square of n_1.

$$E(n_1^2) = E\left(\sum_{i=1}^{n_1} \alpha_i \beta_i\right)^2$$

$$= E\left(\sum_{i=1}^{n_1} \alpha_i^2 \beta_i^2 + 2\sum_{i=j+1}^{N_1-1} \sum_{j=i+1}^{N_1} \alpha_i \beta_i \alpha_j \beta_j\right)$$

$$= E\left(\sum_{i=1}^{N_1} \alpha_i \beta_i + 2\sum_{i=1}^{N_1-1} \sum_{j=i+1}^{N_1} \alpha_i \alpha_j \beta_i \beta_j\right)$$

$$= N_1 \cdot \Pr(\alpha_i = 1) \cdot \Pr(\beta_i = 1)$$
$$+ 2\sum_{i=1}^{N_1-1} \sum_{j=i+1}^{N_1} \Pr(\alpha_i = \alpha_j = 1) \cdot \Pr(\beta_i = 1) \cdot \Pr(\beta_j = 1)$$
$$= N_1 (n/N)(1 - q^s) + N_1^2 (n^2/N^2)(1 - q^s)^2$$
$$= nP(1 - q^s) + nP(NP - 1)(n - 1)(1 - q^s)^2 / (N - 1)$$

The variance of n_1 is

$$Var(n_1) = E(n_1^2) - (E(n_1))^2$$
$$= nP(1 - q^s)$$
$$+ nP(NP - 1)(n - 1)(1 - q^s)^2/(N - 1) - (nP(1 - q^s))^2$$

$$= nP(1 - q^s) + \frac{nP(1 - q^s)^2}{N - 1}((NP - 1)(n - 1) - (N - 1)nP)$$

$$= nP(1 - q^s) + \frac{nP(1 - q^s)^2}{N - 1}((1 - P)(N - n) - (N - 1))$$

$$= nP(1 - P)\frac{N - n}{N - 1}(1 - q^s)^2 + nP(1 - q^s) - nP(1 - q^s)^2$$

$$= nP(1 - P)\frac{N - n}{N - 1}(1 - q^s)^2 + nP(1 - q^s)q^s$$

Therefore, the variance of $\hat{P}$ is

$$Var(\hat{P}) = Var(n_1)/n^2(1 - q^s)^2 = P(1 - P)\frac{N - n}{N - 1}\frac{1}{n} + P\frac{1}{n}\frac{q^s}{1 - q^s}$$

In practice we will not know q, but we have an estimate, $\hat{q}$, based on the average number of visits required to observe an owl. In addition, P is also unknown, and an estimate must be used. We propose the following estimator for P:

$$\tilde{P} = (n_1/n)/(1 - \hat{q}^s)$$

with estimated variance

$$Var(\tilde{P}) = \tilde{P}(1 - \tilde{P})(1 - n/N)\frac{1}{n - 1} + \tilde{P}\frac{1}{n - 1}\frac{\hat{q}^2}{1 - \hat{q}^2}$$

When $q = 0$ this function provides an unbiased estimate of the variance. In effect, we have simply plugged in estimates of P and q and have not investigated the consequences.

The Forest Service, U. S. Department of Agriculture, is responsible for Federal leadership in forestry. It carries out this role through four main activities:

- Protection and management of resources on 191 million acres of National Forest System lands
- Cooperation with State and local governments, forest industries, and private landowners to help protect and manage non-Federal forest and associated range and watershed lands
- Participation with other agencies in human resource and community assistance programs to improve living conditions in rural areas
- Research on all aspects of forestry, rangeland management, and forest resources utilization.

The Pacific Southwest Research Station

- Represents the research branch of the Forest Service in California, Hawaii, American Samoa and the western Pacific.