Monitoring Martes Populations in California: Survey Design and Power Analysis
Author(s): William J. Zielinski and Howard B. Stauffer
Source: *Ecological Applications*, Vol. 6, No. 4 (Nov., 1996), pp. 1254-1267
Published by: Ecological Society of America
Stable URL: http://www.jstor.org/stable/2269605

Ecological Applications, 6(4), 1996, pp. 1254–1267

MONITORING *MARTES* POPULATIONS IN CALIFORNIA: SURVEY DESIGN AND POWER ANALYSIS[1]

WILLIAM J. ZIELINSKI
USDA Forest Service, Pacific Southwest Research Station, Redwood Sciences Laboratory, and Department of Wildlife, Humboldt State University, Arcata, California 95521 USA

HOWARD B. STAUFFER
Department of Mathematics, Humboldt State University, Arcata, California 95521 USA

Abstract. Fishers (*Martes pennanti*) and American martens (*M. americana*) have been protected from trapping in California since the mid-1900s, yet in portions of each of their historic ranges their numbers are extremely low, perhaps due to the effects of timber harvest. We propose a method capable of detecting declines in the occurrence and distribution of fishers or martens using baited track-plate stations. The proposed sampling unit is a small grid of stations that has a high probability of detecting animals when they are present. These multistation units are sufficiently spaced to meet the assumption of independence for a binomial model. We propose a stratified random sampling design with strata sampled for proportions of occurrence at discrete points in time. Stratification is based on variation in occurrence by region and is estimated from preliminary survey data. A previously published bias adjustment is applied to the proportion of units with detections to adjust for possible failure to detect resident individuals at a sampling unit. A Monte Carlo simulation model was developed to determine the sample size necessary to detect 20 and 50% declines, with 80% power, in the proportion of sampling units with occurrence. We assume a 10-yr sampling interval. Sensitivity analysis, using a range of values for means and standard deviations of strata proportions, determined that power was much more sensitive to changes in mean than the standard deviation. When the best current estimates of the fisher strata proportions were input for 10 strata (five regional and two habitat) in California, 115 and 17 sampling units per stratum were necessary to detect 20 and 50% declines, respectively. For some circumstances this sampling effort was also sufficient to achieve strata estimates with 5% error and to detect statistical differences between individual stratum proportions. The steps in the process of implementing a monitoring program for Pacific fishers in California are outlined as an example of the planning and preparation necessary to monitor changes in the distribution of a rare forest carnivore.

Key words: detection; forest carnivores; Martes; *monitoring; occurrence; sign survey; statistical power.*

INTRODUCTION

Background

The fisher (*Martes pennanti*) and American marten (*M. americana*) are protected carnivores in California. Their distributions and abundances have been poorly understood since legal trapping closed in 1945 and 1954, respectively. There has been no concerted effort to determine the status of these species, and the best available information has come from incidental sightings (Schempf and White 1977) and limited use of standardized surveys associated with timber harvest activities (Kucera et al. 1995, Zielinski et al. 1995; W. J. Zielinski, *unpublished manuscript*) or conducted for research purposes (Rosenberg and Raphael 1986, Raphael 1988).

In California, fishers occur most frequently in the mixed-conifer, Douglas-fir (*Pseudotsuga menziesii*), and ponderosa pine (*Pinus ponderosa*) forest types at low-to-mid elevations, whereas martens occur primarily in the true fir (*Abies* spp.) forests that occur at higher elevations, though exceptions occur in the north coast region (Grinnell et al. 1937, Schempf and White 1977). Under the National Forest Management Act of 1976 fishers are entitled to "special management emphasis to ensure their viability and to preclude trends toward endangerment that would result in the need for Federal listing [under the Endangered Species Act]." It is difficult to fulfill this federal mandate, and to understand the effects of land-use changes such as timber harvest on forest carnivores, using a database composed of anecdotal information and haphazard surveys. A specific program is required to detect, inventory, and index the status of each species. We propose a monitoring program that will describe the relative statewide distribution of each species and detect declines in the occurrence of each species throughout its range in California. This is the first proposal for a statewide monitoring scheme for fishers or martens using nonlethal

[1] Manuscript received 18 November 1994; revised 10 November 1995; accepted 29 November 1995; final version received 18 January 1996.

methods and is an example of the type of wildlife population monitoring approach that will become increasingly necessary for mammalian carnivores that are no longer commercially harvested or hunted.

Statistical considerations

Population monitoring requires considerable planning and statistical evaluation before implementation (de la Mare 1984, Peterman and Bradford 1987, Gerrodette 1987, Verner and Kie 1988, Kendall et al. 1992, Taylor and Gerrodette 1993). The null hypothesis that there has been no change in a population index between two time points must be tested against the alternative that the population has simply changed (either increased or decreased: two-tailed test), or has declined (one-tailed test). Sample size, sample variability, and the magnitude of the real difference (effect size) all affect the ability to detect change (Cohen 1988:4). We must therefore ask the important question: if a significant population decline has occurred, what is the probability that we will detect it with our survey? The answer is critical to a monitoring program, yet the probability of detecting a change if it has occurred (i.e., correctly rejecting the null hypothesis when the alternative hypothesis is true), called statistical power ($1-\beta$), is rarely considered. In developing a sampling design to index population change it is essential to determine, a priori, the probability of detecting significant declines and to choose an adequate sample size to detect change with an acceptably high probability. An awareness of the concept of statistical power in the field of ecology has been heightened in recent years (Quinn and Dunham 1983, Toft and Shea 1983, Peterman 1990), but it is critical in the field of conservation biology because failure to detect a significant decline may result in local extirpation or even extinction before rescue efforts can be undertaken (Taylor and Gerrodette 1993). Moreover, the costs of conducting vertebrate monitoring programs can be very high and they should not be undertaken without adequate knowledge of the probability that real changes can be detected (Verner and Kie 1988).

Objectives

Our primary objective is to develop a statistically sound and efficient sampling design for detecting change in an index of the occurrence and distribution of fishers or American martens in California. We also wish to contrast estimates across regional and habitat strata. Surveys using detection devices have been conducted in California for some time (Kucera et al. 1995, Zielinski et al. 1995), but their locations have usually been dictated by the needs of local land managers for information about the presence of a species in areas proposed for timber harvest or recreational development (W. J. Zielinski, *unpublished manuscript*). Consequently, the distribution of surveys throughout the range of each species has been haphazard. We propose a systematic scheme for sampling the distribution of fishers or martens in California and for predicting the sampling effort sufficient to be confident of detecting changes in distribution over time.

Using a Monte Carlo simulation model we mimic the sampling necessary to monitor changes in occurrence and estimate the statistical power of tests for specified declines. To compensate for the possibility that our sign detection stations may not always detect a resident animal at sample sites we also adjust our estimates for bias. Because many of the parameters that influence occurrence and detection are currently unknown for fishers or martens, our initial goal was to simulate a variety of conditions (a sensitivity analysis) to explore the range of sampling effort and power to detect changes in the index. We conclude by using best estimates of the probabilities of occurrence to propose a monitoring scheme for fishers throughout their range in California.

METHODS

The detection method: baited track plates

The sooted track plate has received widespread use in surveys for fisher and marten in the Pacific States (Barrett 1983, Raphael and Marcot 1986, Fowler and Golightly 1994, Raphael 1994, Zielinski and Kucera 1995; W. J. Zielinski, *unpublished manuscript*). Each detection station is composed of a sooted aluminum sheet that, in its most recent rendition, is partially covered with a white imprint surface and enclosed in a plywood box (Fowler and Golightly 1994). Bait is placed at the rear of the box to attract animals across the plate where they leave positive impressions of their tracks on the white surface. Both species are readily attracted to and detected at baited track plates, and their tracks can be easily distinguished (Zielinski and Truex 1995). Ease of use, low cost, reliability of results, and ability to distinguish species easily make the sooted track plate the most efficient and affordable method available. Although the method of detection is not crucial to our analysis, different methods could result in different detection probabilities.

The sampling unit

Most surveys that have been conducted in California have inundated a particular area with track-plate stations (at ≈1.0-km intervals along roads) and checked them every 2 d for 2–3 wk (Zielinski et al., *in press*). This type of survey is conducted to determine presence only. However, similar surveys have been suggested as a means of indexing the abundance of a small, local population. The "detection ratio" (the number of stations visited by a species divided by the total number of stations) has been suggested as a response variable useful for monitoring trends in local populations (Fowler and Golightly 1994). However, the likelihood that individual stations are not independent sampling units (e.g., Roughton and Sweeney 1982, Diefenbach et al.

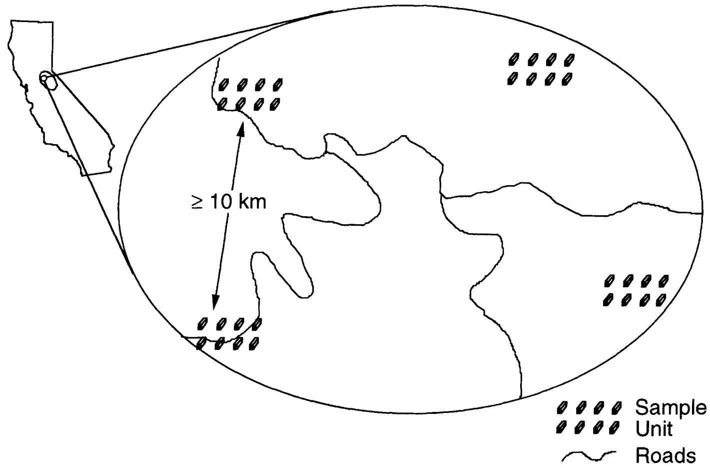

FIG. 1. Schematic example of the distribution of sampling units to survey fishers. Adjacent sampling units are no closer than 10 km apart.

1994), and the small number of individuals that usually occupy any such sampling area (especially for fishers) render the relationship between population size and detection ratio suspect.

Our approach proposes sampling units that are sufficiently distant from one another that we can treat them as independent (Fig. 1). Intersampling unit distances of at least 10 km for fisher and 5 km for marten are probably sufficient as these distances are almost twice the size of the diameter of the mean male home range sizes of each species in the western U.S. (Buck et al. 1983, Buskirk and McDonald 1989, Jones 1991, Buskirk and Ruggiero 1994, Powell and Zielinski 1994) when the ranges are drawn as circles.

Each proposed sampling unit is composed of eight track-plate stations, 0.6 km apart, in an approximate 4 × 2 grid pattern (Fig. 1). The survey objective is to determine whether the area in or nearby each unit is "occupied" or "not occupied," thus a detection need occur at only one of the track-plate stations to register the unit as occupied. When this occurs the effort on that sampling unit is complete because the outcome has been determined. Stations are to be checked every 2 d, and the survey effort at a particular sampling unit will terminate when either the target species is detected at any one of the stations or after eight visits (i.e., 16 d after set up). Multiple stations per sampling unit assure a greater chance of detecting a resident animal in a reasonable period than a single station (Roughton and Sweeney 1982, Diefenbach et al. 1994). Stations within a sampling unit are not considered independent, but this has no effect on the interpretations as only the first detection at any station is recorded. Thus, the multistation sample unit described here is analogous to the "station" described in Raphael (1994). The 0.6 km (0.5 mi) spacing between stations within a sample unit is a default chosen because of its consistency with previous detection protocols (Fowler and Golighty 1994; W. J. Zielinski, *unpublished manuscript*).

The rationale for eight stations and eight visits comes largely from a review of 207 track-plate or line-triggered camera surveys in California (Zielinski et al., *in press*), most conducted in areas proposed for timber harvest. Surveys that had between 6 and 12 stations (n = 50) received the first fisher detection after a mean ± 1 SD of 4.2 ± 2.4 d and the first marten detection after 3.7 ± 2.6 d. Furthermore, 89.7% of the surveys that detected either a fisher or a marten did so in <16 d, the maximum duration proposed here. Although this information indicates that eight stations visited eight times each should be sufficient effort to detect a resident animal, we also include a bias adjustment to compensate for the proportion of sample units that failed to detect a resident.

Many mustelids (including fishers, martens, and weasels [*Mustela* spp.]) are particularly suited for this monitoring approach because they demonstrate intrasexual territoriality (Powell 1994) and have relatively stable home ranges from year to year (Hawley and Newby 1957, Lockie 1966, Arthur et al. 1989; D. Harrison, *unpublished data*). Although the monitoring system we propose will work equally well for fishers and martens, both species cannot be monitored simultaneously because of their largely allopatric distributions in California (Grinnell et al. 1937, Schempf and White 1977). Separate sets of sampling units will be necessary for fishers and martens if information is required on both species.

Correction for failing to detect: a bias adjustment

Because the response at a sampling unit is represented as binary, and multiple visits by one or more individuals are unimportant, variation in individual behavior that affects detection will have less effect on our estimates than on abundance or survival estimates from mark–recapture studies (Pollock et al. 1990). However, the failure to detect at least one resident animal during the survey period introduces bias in the sampling. Despite our current belief in the efficacy of the detection method, too little is known about forest carnivores to assume probabilities of detection are near 100%. When martens or fishers are marked with color ear tags and later resighted at cameras, most individuals are detected at at least one camera station that falls within their home range (Fowler and Golightly 1994; L. Jones and M. Raphael, *unpublished manuscript*; A. Seglund and R. Golightly, *unpublished manuscript*). However, a few individuals are never detected at camera stations within their home range, and detectability could be different at track plate stations than at camera stations (methods for identifying specific individuals of a species from tracks have not been developed). Given these uncertainties we include a bias adjustment in our monitoring protocol until probabilities of detection are more completely understood.

We use a method developed by Azuma et al. (1990) for Northern Spotted Owl (*Strix occidentalis caurina*) populations to estimate the probability of failing to detect an animal on a sampling unit. These authors demonstrate that the average number of visits until first detection, which we redefine as "latency to first detection" (LFD), can be used to develop a function that can correct for bias in estimates of the proportion of occupied sampling units. They used a binomial model with the following assumptions: (1) there is a *constant* conditional probability of determining occupancy for each visit to an occupied unit and, (2) each visit to a unit is an independent and identically distributed Bernoulli trial that can result in one of two possible outcomes, occupancy or not, each with a fixed probability. Thus, if it is determined that visits have varying probabilities of the two outcomes, or that stations within a sampling unit differ in these probabilities, the bias adjustment will be less helpful.

To correct for bias the estimated proportion of occupied units, $\hat{p}_i$, is multiplied by a factor that reflects the magnitude of difference between the maximum number of visits permitted without success and the mean LFD from those sampling units with a detection. LFD is derived from the data after the sampling is complete or from a pilot data set. If first detections tend to occur early relative to the maximum number of visits, the probability of overlooking a resident on any sampling unit is considered to be low and the correction factor is close to 1. However, if the mean LFD is near the maximum number of visits it is likely that new first detections would occur after this time and the correction factor will be substantially greater than 1. These estimates will not only be useful to correct the proportion estimates for surveys conducted according to our recommendations but can also be used to adjust for undetected residents if survey duration varies from our suggestions.

The sampling design

The pattern of occurrence of a species across its historic range is a "frequency index" (Seber 1982, Lancia et al. 1994) of its distribution. Because these indices are sensitive to spatial contagion (Seber 1982) some form of stratification is necessary. To do so we propose regional stratification that is based on the results of previous surveys in California (Kucera et al. 1995, Zielinski et al. 1995). We assume that the probability of presence of the target species at any sampling unit within a stratum is constant. We then simulate stratified random sampling of presence/absence data within strata, using a binomial model.

The design assumes that there are n_i sampled units in each of k strata ($1 \leq i \leq k$) with measurements x_{ij} ($1 \leq j \leq n_i$) equal to 0 or 1 depending on the absence or presence of the species in the j^{th} sampling unit of the i^{th} stratum. For each survey the proportion estimate $\hat{p}_i = \Sigma_{j=1}^{n_i} x_{ij} n_i^{-1}$ can be calculated to estimate the proportion of occupied sampling units in a stratum. A weighted proportion estimate $\hat{p} = \Sigma_{i=1}^{k} w_i \cdot \hat{p}_i$ and confidence interval can also be calculated for the entire population proportion p in all the strata, using the standard formulas for stratified random sampling with weights w_i proportional to the area of the stratum (Cochran 1977, Scheaffer et al. 1990). A number of alternative sampling allocation schemes could be considered (Scheaffer et al. 1990), but we recommend an equal allocation scheme because our proportion estimates are sufficiently tentative that it is best to keep the design simple and flexible. Optimal allocation, with sample sizes proportional to area and variation, and inversely proportional to cost, may ultimately improve the efficiency of sampling, but we presently lack the information to propose such a design.

We explore the detection of declines over two time points. However sampling will presumably continue at specified intervals over time to assess trend in the index of occurrence. Repeated-measures analysis in ANOVA (Kuehl 1994:499–521) can be used to test for significant change over repeated time periods or successive estimates can be compared with one or more previous estimates, adjusting α accordingly. Multiple comparison tests (e.g., Tukey's test, Fisher's test, Duncan's test) can be used to establish groupings over time. Regression analysis can also be applied to a series of proportion estimates over time to estimate trend (Gerrodette 1987).

Statistical power simulations

Power tables have been developed for a variety of sampling designs (Cohen 1988) but not, to our knowledge, for tests we propose here that compare proportions using stratified designs. Sample size can be determined for estimates of proportion for one population using standard error (Cochran 1977, Scheaffer et al. 1990) and for differences between two proportions by using effect size (Cohen 1988:186–197). However, efficient formulas based on standard error or effect size have not been derived for stratified sampling of pairs of proportions. Kendall et al. (1992) developed a power table for a specialized stratified design examining bear sign (scat and tracks) on trail segments, using simulation. We extend their findings more generally, by conducting "prior power analysis" (Fairweather 1991) over a broad range of parameter values, using computer simulation.

The estimates of proportions of occupied sampling units can be based upon either the binomial or hypergeometric model. We adopt the simpler binomial, noting that the two models are essentially equivalent when sampling is with replacement or when sampling units are moderately large in number (Mendenhall et al. 1990). The binomial model assumes that the sampling units are independent (i.e., the same animals are not detected on multiple sampling units) and that the probability of detecting at least one resident animal on a sampling unit is fixed within a stratum. Each sampling unit is, hence, assumed to be an independent Bernoulli trial with one of two possible outcomes, presence or absence, each with a fixed probability of occurrence. It is for this reason that the sampling units have been designed to be more than one home range diameter of the target species apart.

Although we assume that the probability of occupancy at sampling units within each stratum does not vary, T. Matsumoto and H. Stauffer (*unpublished data*) have determined that the binomial power estimates change very little when probabilities of occupancy and detection at sampling units within a stratum vary around a fixed mean. That is, the distribution of the random variable x (the number of occupied sampling units in a survey of sample size n), and hence the distribution of the proportion estimator $\hat{p} = x/n$, are equivalent if probabilities of occurrence at the sampling units within a stratum are allowed to vary. In these simulations, a wide range of beta distributions have been used to generate varying probabilities around fixed means yet the binomial probabilities have proven to be quite robust. Hence, our results should remain valid for uniform, random, and clumped dispersions, as long as the mean averages of the probabilities are used to parameterize the binomial model.

We created a sampling simulator named "POWER" that can mimic the monitoring system and generate power estimates (see Appendix). POWER is designed to simulate stratified random sampling in two separate survey periods to estimate proportions of sampling units with a detection and to test for a significant decline in the proportion of occupied units. POWER accepts a range of varying input parameter specifications, including percent decline in the index, number of strata, number of samples per strata, probability of initial sampling unit occupancy for each stratum, the probability of a Type I error (α), and number of simulations.

We simulated declines of 20 and 50% in the proportion index. It is important to realize that these are declines in the index only; the relationship between this index and the true population size is unknown. However, because positive detections at a sampling unit represent one or more individuals in the area, the change in the population status will most likely be less than the specified rate of decline in the index. For the purpose of this exercise we assume an intersampling interval of 10 yr. Thus, our 20% decline represents a 2.2% annual rate of decline, and the 50% decline a 6.7% annual rate.

We chose 20% as the minimum rate of decline to detect because smaller changes would not only be increasingly difficult and expensive to detect over the specified interval, but because we wished to exclude the "noise" that would result in attempting to detect smaller changes that could be natural variations in equilibrium populations. However, we include a few simulations that estimate the power to detect 5 and 10% declines, merely to demonstrate the additional sampling effort necessary to detect these small effect sizes. A 50% decline was chosen to demonstrate the difference in sampling effort necessary if only catastrophic declines need be detected. We planned to detect declines only, rather than declines or increases in the index, because of the reduced sample sizes for testing one-tailed rather than two-tailed alternatives (Verner 1983, Cohen 1988).

We ran a thorough analysis to check the required normality assumptions of our testing procedures and the accuracy of our results. We examined the differences of the strata proportions, untransformed and transformed, for normality with sample sizes as low as 10 and stratum proportions as low as 5%. Additionally, the sample sizes necessary to ensure adequate power in this design (see *Results*) are also large enough to satisfy the accepted theoretical requirement for normality of the differences (i.e., $n_t \cdot p_t \geq 5$ and $n_t \cdot (1 - p_t) \geq 5$) (Samuels 1989). The Wilcoxon test was used to contrast the results of selected runs but we found it to be less powerful in discriminating differences than the paired t test. Homoscedasticity was stabilized by arcsine transformation. To confirm the accuracy of the Type I error rate (α) we also ran the simulator with zero decline and found that the proportion of rejections of the null hypothesis approximated α. Arcsine-transformed proportion estimates produced power estimates that differed very little (no more than $\approx$3%, well within the bounds of simulation error) from those produced using untransformed estimates.

Type I error rate was set at 20% because we considered the environmental cost of a Type II error to be much larger than that of Type I error; a conclusion consistent with previous authors (e.g., Parkinson et al. 1988, Peterman 1990, Kendall et al. 1992, Mapstone 1995). In doing so we placed priority on increasing the probability of detecting a real decline in occurrence over mistakenly concluding a decline has occurred.

Our initial simulations consisted of a sensitivity analysis where we varied input parameters, such as the mean and standard deviation of strata probabilities, to examine effects on power for different sample sizes. We conclude with an analysis that uses specific input values that are our best estimates of the current strata

FIG. 2. Example of regional strata for a fisher monitoring program. Strata include the historic range of fishers in California and are identified as follows: 1—Klamath, 2—Sierra–Klamath Transition, 3—Northern Sierra Nevada, 4—Central Sierra Nevada, and 5—Southern Sierra Nevada.

proportions for the fisher, the species we use as our specific example. The results are then used to propose a monitoring scheme for this species.

Sensitivity analysis: simulations over a range of parameter values

We estimated the power of detecting 20 and 50% declines, using a broad range of means and standard deviations for the distribution of strata proportions. We were interested in examining how the power values would be affected by the magnitudes of the strata proportion means and their variation, sample sizes per stratum, and number of strata. Mean and standard deviation values for the beta distribution that generated the probabilities for each strata were as follows (standard deviations in parentheses): 5% (1, 3, 5%), 10% (1, 3, 5, 10%), 20% (1, 5, 10, 20%), 30% (1, 5, 10, 20, 30%), 40% (1, 5, 10, 20, 30%), and 50% (1, 5, 10, 20, 30%). For each run a mean and standard deviation were chosen to characterize the distribution of strata proportions across all 5 or 10 strata. Each unique combination of mean, standard deviation, decline, and number of strata was simulated 1000 times.

Simulations with specific parameter values: fisher example

Using the information available on the current (Zielinski et al. 1995) and historic (Grinnell et al. 1937, Schempf and White 1977) distributions of fishers in California we estimated the likely proportions of sample units with detections for each stratum and used this information to calculate the sampling effort necessary to detect 20 and 50% declines. Fixed probabilities of occupancy per sampling unit were assigned to each stratum in two ways: first for five regional strata without habitat substratification and then for two habitat types within each of the five regional strata (10 strata total).

We envisioned the historic range of fishers in California as five equivalent-sized regions: the Southern Sierra Nevada, Central Sierra Nevada, Northern Sierra Nevada, Klamath, and Sierra–Klamath Transition (Fig. 2). Recent evidence suggests that the northwestern part of the state (the Klamath province) and the southern Sierra Nevada have had the highest fisher densities (Zielinski et al. 1995). Preliminary data suggest that $\approx$30% of the surveys in each of these regions were successful so these strata were assigned values of 0.30 as estimates of the expected proportion of occupied units. The northern Sierra Nevada had no surveys with detections, and very few sightings, so was assigned a value of 0.05. The Sierra–Klamath Transition and the Central Sierra Nevada were assigned the values of 0.20 and 0.10, respectively, on the basis of their intermediate survey success (Fig. 2). These estimates are based on the best information available. If they prove to be inaccurate, the sensitivity analysis (described above) provides a range of outcomes that will likely encompass corrected values.

Variation in abundance should also occur among habitats within each regional stratum. Although studies in the western U.S. suggest that fishers and martens prefer late-successional, or old-growth, conifer forests that are in large unfragmented blocks (Rosenberg and Raphael 1986, Buskirk and Ruggiero 1994, Powell and Zielinski 1994) the published and unpublished data from California are somewhat equivocal on the subject. For this reason our second set of simulations includes a sampling design with two different habitat substrata within each regional stratum. These are envisioned as habitats that are largely late-successional (LS) or non-late-successional (non-LS) conifer forest types. We emphasize here the importance of habitat stratification to increase the efficiency of sampling and to compare strata; specific rules for classification of habitat strata are not discussed.

The proposed LS habitats within each region were assigned their regional values as estimates for the expected proportions of occupied sampling units (i.e., 0.05, 0.10, 0.20, 0.30, and 0.30) and the non-LS habitats were assigned values half that amount. The choices of sample size values, alpha level, and percent declines for the 5- and 10-strata simulations were the same as used for the sensitivity analysis. We view the proportion estimates for both simulations as "best guess" starting points in an adaptive process, subject to later refinement.

Precision of estimates of proportions and comparisons of estimates among strata

Sample sizes should be chosen not only to satisfy power requirements for statewide estimates but also to ensure an adequate level of precision for individual

TABLE 1. Bias adjustment [$f(v)$] for estimation of occupancy at a sampling unit (after Azuma et al. 1990).

Maximum visits (s)†	Probability of detection (p)‡	Probability of failing to detect (q)§	LFD (v)‖	Bias adjustment [$f(v)$]¶
4	0.05	0.95	2.44	5.3910
	0.10	0.90	2.37	2.9078
	0.15	0.85	2.30	2.0921
	0.20	0.80	2.22	1.6938
	0.25	0.75	2.15	1.4629
	0.30	0.70	2.07	1.3160
	0.35	0.65	1.99	1.2173
	0.40	0.60	1.90	1.1489
	0.45	0.55	1.82	1.1007
	0.50	0.50	1.73	1.0667
	0.55	0.45	1.65	1.0428
	0.60	0.40	1.56	1.0263
	0.65	0.35	1.48	1.0152
	0.70	0.30	1.40	1.0082
	0.75	0.25	1.32	1.0039
	0.80	0.20	1.24	1.0016
	0.85	0.15	1.17	1.0005
	0.90	0.10	1.11	1.0001
	0.95	0.05	1.05	1.0000
6	0.05	0.95	3.35	3.7749
	0.10	0.90	3.19	2.1342
	0.15	0.85	3.03	1.6055
	0.20	0.80	2.87	1.3553
	0.25	0.75	2.70	1.2165
	0.30	0.70	2.53	1.1333
	0.35	0.65	2.37	1.0816
	0.40	0.60	2.21	1.0489
	0.45	0.55	2.05	1.0285
	0.50	0.50	1.90	1.0159
	0.55	0.45	1.77	1.0084
	0.60	0.40	1.64	1.0041
	0.65	0.35	1.53	1.0018
	0.70	0.30	1.42	1.0007
	0.75	0.25	1.33	1.0002
	0.80	0.20	1.25	1.0001
	0.85	0.15	1.18	1.0000
	0.90	0.10	1.11	1.0000
	0.95	0.05	1.05	1.0000
8	0.05	0.95	4.23	2.9711
	0.10	0.90	3.95	1.7558
	0.15	0.85	3.67	1.3746
	0.20	0.80	3.39	1.2016
	0.25	0.75	3.11	1.1113
	0.30	0.70	2.84	1.0612
	0.35	0.65	2.59	1.0329
	0.40	0.60	2.36	1.0171
	0.45	0.55	2.15	1.0084
	0.50	0.50	1.97	1.0039
	0.55	0.45	1.80	1.0017
	0.60	0.40	1.66	1.0007
	0.65	0.35	1.54	1.0002
	0.70	0.30	1.43	1.0001
	0.75	0.25	1.33	1.0000
	0.80	0.20	1.25	1.0000
	0.85	0.15	1.18	1.0000
	0.90	0.10	1.11	1.0000
	0.95	0.05	1.05	1.0000
12	0.05	0.95	5.89	2.1756
	0.10	0.90	5.28	1.3936
	0.15	0.85	4.68	1.1658
	0.20	0.80	4.11	1.0738
	0.25	0.75	3.61	1.0327
	0.30	0.70	3.16	1.0140
	0.35	0.65	2.79	1.0057
	0.40	0.60	2.47	1.0022
12	0.45	0.55	2.21	1.0008
	0.50	0.50	2.00	1.0002
	0.55	0.45	1.82	1.0001
	0.60	0.40	1.67	1.0000
	0.65	0.35	1.54	1.0000
	0.70	0.30	1.43	1.0000
	0.75	0.25	1.33	1.0000
	0.80	0.20	1.25	1.0000
	0.85	0.15	1.18	1.0000
	0.90	0.10	1.11	1.0000
	0.95	0.05	1.05	1.0000

† s = maximum number of visits to the sampling unit.
‡ p = probability of detecting the target species at the sampling unit in one visit.
§ q = probability of failing to detect the target species at the sampling unit in one visit $(1 - p)$.
‖ v = average latency to first detection (LFD)
$$= \left[\frac{1}{p} - \frac{sq^s}{(1 - q^s)}\right].$$
¶ $f(v)$ = bias adjustment of occupancy estimate,
$$= \frac{1}{(1 - q^s)}.$$

strata proportion estimates. Precision will need to be controlled and between-strata comparisons will be of interest. For example, managers may be interested in whether the proportion of occupied sampling units in the Klamath region is different from the proportion in the southern Sierra Nevada, or whether the LS habitats have higher occupancies than the non-LS habitat strata. Estimates for sampling error of individual stratum proportion estimates were obtained using the following formula:

$$E = z\sqrt{\frac{\hat{p}\cdot(1 - \hat{p})}{n}},$$

with $\hat{p}$ = proportion estimate, n = sample size, z = 1.28 for 80% confidence, and z = 1.96 for 95% confidence.

RESULTS

Correction for failing to detect

An unbiased estimate of the true proportion of occupied sampling units can be calculated by multiplying the actual proportion of occupied sampling units/strata by bias adjustments provided in Table 1. To illustrate the use of this table, using 8 as the maximum number of visits, a mean LFD of 3.4 would require a correction factor of 1.2016. This must be multiplied by the proportion of occupied units in that stratum to correct for bias (i.e., the corrected unbiased estimator for the proportion would be 1.2016 × $\hat{p}$). Alternatively, the conditional probability of detection may be used to correct for bias; for example, if the conditional probability of detection is 10%, a correction factor of 1.7558 should

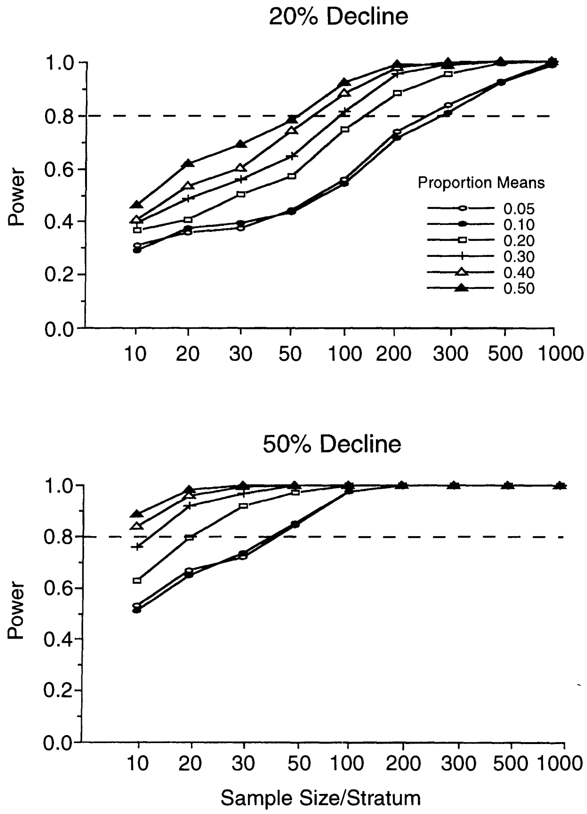

FIG. 3. Relationship between the number of sampling units for each of five strata and the power to detect 20 and 50% declines for various proportion means from the beta distribution when the standard deviation from the beta distribution is low (0.01).

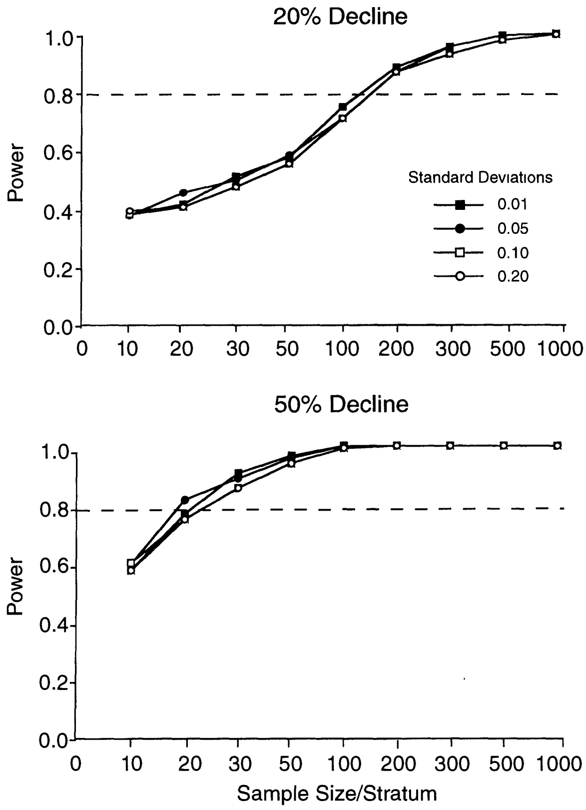

FIG. 4. Relationship between the number of sampling units for each of five strata and the power to detect 20 and 50% declines for various values of the standard deviation from the beta distribution when the beta mean proportion is held constant at 0.20.

be applied to the proportion estimate when the maximum number of visits is 8.

Simulations over a range of parameter values

Varying proportion means.—As the beta proportion mean increases (i.e., mean proportion of sampling units with a detection), but standard deviation is kept low and constant, the number of samples necessary to achieve 80% power decreases (Fig. 3). Thus, when populations of the target species are uniformly high among all strata, the number of sampling units needed to detect a decline will be relatively low. For example, if the mean proportion of sample units with a detection is 0.50 (individual stratum values for the proportion of sampling units with a detection of, say, 0.25, 0.40, 0.50, 0.60, 0.75) only ≈55 samples per strata would be necessary to detect a 20% decline in the index at a power of 80%. Conversely, a mean proportion of 0.05 (only 5 of 100 sampling units, on average, with a detection) would require ≈300 samples per strata.

Varying proportion standard deviation.—Variation around the mean proportion of occupied sampling units has substantially less effect on sample size, and power, than the magnitude of the mean itself (Fig. 4). Higher standard deviations, across strata, decrease the power to detect declines only slightly. For example, for the five-strata condition with a mean proportion of 0.20 and a sample size of 100, the power to detect a 20% decline is 0.74 when the standard deviation equals 0.01 and drops only to 0.70 when the standard deviation is increased to 0.20. Thus, when the proportion of occupied sampling units differs considerably among strata the effort required to detect a specific decline is not much greater than when animals are more uniformly dispersed across their range.

Varying percent decline.—With other factors held constant, detecting a 50% decline requires considerably less sampling effort to achieve 80% power than detecting a 20% decline (Figs. 3 and 4). In some cases less than half the number of sampling units are required to detect the 50% than the 20% decline.

Simulations with specific parameter values: fisher example

Using our specific estimates of fisher detection rates for different regions and habitats, the sampling effort necessary to detect 20 and 50% declines is substantial. In the five-strata condition (regional stratification only)

≈170 and 25 samples per stratum are required to detect 20 and 50% declines, respectively, at 80% power (Fig. 5A). With 10 strata (regional and habitat stratification), ≈115 and 17 samples per stratum would be necessary to detect 20 and 50% declines with 80% power (Fig. 5B). To detect a significant decline (20%) in the index, given the proportion estimates we used, 850 total sampling units would need to be surveyed at each sampling period if there are five strata and 1150 would be needed if there were 10 strata. The statewide number of sampling units is higher for the 10- compared to the 5-strata design because the mean estimated proportion of occupied units is lower in the former (0.14) than the latter (0.19). Not surprisingly, fewer sampling units are required to detect the 50% decline compared to the 20% decline.

The detection of more modest declines, 5 and 10%, for the 10-strata condition using the same strata values, would require over 1000 and 500 sampling units per strata, respectively. Thus, detecting these small declines would require the effort of sampling over 10 000 (for 5% decline) and over 5000 (for 10% decline) sampling units statewide to achieve 80% power.

Precision of estimates of proportions within and among strata

Sampling errors (E) for a range of proportion estimates ($\hat{p}$) and sample sizes (n), at 95 and 80% confidence, permit the choice of sample size adequate to bound the sampling error for a given stratum proportion estimate (Table 2). For example, if we anticipate a stratum estimate of $\hat{p}_i$ = 30% and wish to limit E to 5%, at 95% confidence (i.e., the 95% confidence interval would be 25–35%), ≈300 samples would be required for that stratum. If the goal is to distinguish between two strata anticipated to have proportion estimates of $\hat{p}_i$ = 20% and $\hat{p}_j$ = 30%, for example, then we could choose E = 5% for both estimates to distinguish the confidence intervals and would need sample sizes of 110 and 141 (interpolating in Table 2), respectively, at 80% confidence. If k multiple comparisons are planned and a Type I error rate of α is desired ($\alpha = 1 - P$ where P is the confidence level), each individual pairwise comparison should conservatively use an approximate Type I error of α/k to compensate for a compounding of error (cf. the Bonferroni method, Sokal and Rohlf 1987).

Note that the amount of sampling necessary to achieve a precise stratum estimate (and to conduct cross-strata comparisons) can exceed that required for sampling the strata only to achieve a statewide estimate of p. If precise strata estimates are necessary and between-stratum comparisons are planned, either additional samples/stratum may be needed or a more permissive confidence level (or higher sample error) will need to be tolerated.

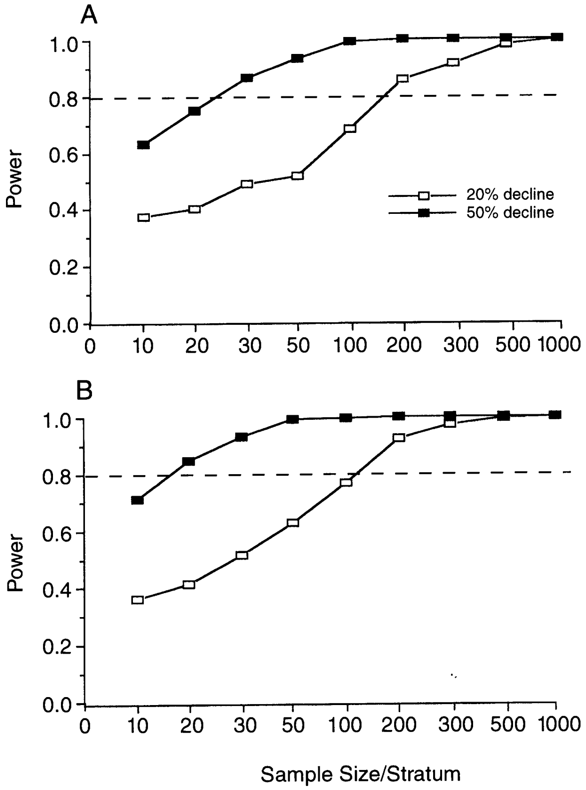

FIG. 5. Power and sample size relationships when the proportion of occupied sampling units is estimated for each of either 5 (A) or 10 (B) strata for 20 and 50% declines. Proportion estimates for the five-strata design are 0.05 (Northern Sierra Nevada), 0.10 (Central Sierra Nevada), 0.20 (Sierra–Klamath Transition), 0.30 (Southern Sierra Nevada), and 0.30 (Klamath). Estimates for the 10-strata design are 0.05 (Northern Sierra Nevada LS), 0.025 (Northern Sierra Nevada non-LS), 0.10 (Central Sierra Nevada LS), 0.05 (Central Sierra Nevada non-LS), 0.20 (Sierra–Klamath Transition LS), 0.10 (Sierra–Klamath Transition non-LS), 0.30 (Southern Sierra Nevada LS), 0.15 (Southern Sierra Nevada non-LS), 0.30 (Klamath LS), and 0.15 (Klamath non-LS). (LS denotes late-successional.)

The monitoring process: fisher example

Here we describe the steps in the process to monitor abundance of fishers in California using information gained from the statistical analyses conducted above and preliminary survey information. A 10-strata design is appropriate because of the expected variation in detection probabilities across regions and between LS and non-LS habitat. As an initial allocation we propose equal number of samples per stratum in equivalent size strata, though ultimately this will depend on the distribution of LS and non-LS habitat in each region. We assume that cost-conscious managers will be satisfied with being able to detect a decline only (vs. an increase or decrease in the index) and that 20% is the largest decline that should elapse before it is detected.

Using the estimates of fisher detection probabilities for the 10-strata design (Fig. 5), ≈115 samples per

TABLE 2. Percent sampling error† (%) as a function of proportion estimate and number of sampling units (n) for 80 and 95% confidence.

	Proportion estimates												
n	0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99
80% confidence													
10	4.4	9.5	13.1	17.5	20.0	21.4	21.9	21.4	20.0	17.5	13.1	9.5	4.4
20	3.0	6.5	8.9	11.9	13.6	14.5	14.8	14.5	13.6	11.9	8.9	6.5	3.0
30	2.4	5.2	7.2	9.6	11.0	11.7	12.0	11.7	11.0	9.6	7.2	5.2	2.4
40	2.0	4.5	6.2	8.2	9.4	10.1	10.3	10.1	9.4	8.2	6.2	4.5	2.0
50	1.8	4.0	5.5	7.4	8.4	9.0	9.2	9.0	8.4	7.4	5.5	4.0	1.8
100	1.3	2.8	3.9	5.2	5.9	6.3	6.5	6.3	5.9	5.2	3.9	2.8	1.3
150	1.0	2.3	3.2	4.2	4.8	5.2	5.3	5.2	4.8	4.2	3.2	2.3	1.0
200	0.9	2.0	2.7	3.6	4.2	4.4	4.5	4.4	4.2	3.6	2.7	2.0	0.9
250	0.8	1.8	2.4	3.2	3.7	4.0	4.1	4.0	3.7	3.2	2.4	1.8	0.8
300	0.7	1.6	2.2	3.0	3.4	3.6	3.7	3.6	3.4	3.0	2.2	1.6	0.7
500	0.6	1.2	1.7	2.3	2.6	2.8	2.9	2.8	2.6	2.3	1.7	1.2	0.6
1000	0.4	0.9	1.2	1.6	1.9	2.0	2.0	2.0	1.9	1.6	1.2	0.9	0.4
95% confidence													
10	7.1	15.6	21.5	28.6	32.8	35.0	35.8	35.0	32.8	28.6	21.5	15.6	7.1
20	4.7	10.2	14.0	18.7	21.4	22.9	23.4	22.9	21.4	18.7	14.0	10.2	4.7
30	3.7	8.1	11.2	14.9	17.1	18.3	18.7	18.3	17.1	14.9	11.2	8.1	3.7
40	3.2	7.0	9.6	12.8	14.6	15.7	16.0	15.7	14.6	12.8	9.6	7.0	3.2
50	2.8	6.2	8.5	11.4	13.0	13.9	14.2	13.9	13.0	11.4	8.5	6.2	2.8
100	2.0	4.3	6.0	7.9	9.1	9.7	9.9	9.7	9.1	7.9	6.0	4.3	2.0
150	1.6	3.5	4.8	6.5	7.4	7.9	8.1	7.9	7.4	6.5	4.8	3.5	1.6
200	1.4	3.0	4.2	5.5	6.4	6.8	6.9	6.8	6.4	5.5	4.2	3.0	1.4
250	1.2	2.7	3.7	5.0	5.7	6.1	6.2	6.1	5.7	5.0	3.7	2.7	1.2
300	1.1	2.5	3.4	4.5	5.2	5.5	5.7	5.5	5.2	4.5	3.4	2.5	1.1
500	0.9	1.9	2.6	3.5	4.0	4.3	4.4	4.3	4.0	3.5	2.6	1.9	0.9
1000	0.6	1.4	1.9	2.5	2.8	3.0	3.1	3.0	2.8	2.5	1.9	1.4	0.6

† Sampling error derived from the following formula: $E = z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ with $\hat{p}$ = proportion estimate, n = sample size, z = 1.28 for 80% confidence, and z = 1.96 for 95% confidence.

stratum (1150 statewide) are necessary to be 80% confident that a 20% decline will be detected. This translates to an average of 82 for each of the 14 national forests and national parks in California, fewer if private lands are also included. Land ownership would have to be considered an additional level of stratification, with effects on total sampling effort, unless habitat stratification accounted for most of the variation in forest condition among ownership types. If sampling were to be restricted to federal forest lands the scope of inference would be similarly restricted.

The 115 samples per stratum necessary to achieve an estimate of statewide occurrence would also be sufficient to distinguish at least a 13% difference (each estimate bound by a 6.5% error; Table 2) between the proportions of any two strata, with 80% confidence. For example, this level of sampling would be sufficient to distinguish statistically two strata with 15 and 30% of the sampling units with a detection, respectively. Furthermore, 115 samples would be sufficient to exceed 80% power in a one-tailed test (α = 0.05) comparing two strata (Cohen 1988:188–194). More samples may be necessary to distinguish two estimates that differ by <15%.

After collecting data on the proportion of occupied units in each stratum these proportions would be adjusted for the bias due to failure to detect residents. For example, if 20 of 100 sampling units in a particular stratum had at least one detection, and the Latency to First Detection was 3.4 visits, then that strata proportion of 0.20 would be multiplied by 1.20 (Table 1) to produce a new, unbiased estimate of the proportion of occupied units of 0.24. Thus, we would estimate that 24 of the 100 sampling units had at least one resident animal in the vicinity, 20 of which had actual detections. This process would be conducted for all strata to generate the final estimates. The final analysis for a particular sampling period would consist of weighing estimates of proportion occupancy for the state, with confidence intervals, and estimating confidence limits around the proportion of occupied sampling units for each stratum. The sampling units should remain at the same locations for each subsequent survey.

Assuming the preliminary estimates we used for the proportion of occupied units for each stratum are reasonable, the sample sizes chosen above should ensure at least 80% power for testing the hypothesis that the index has declined by 20%. Additional data should be collected so that the proportion of occupied units can be estimated with greater precision before embarking on statewide sampling. If new data differ significantly from the estimates that we use here, Figs. 3 and 4 can be consulted (or the simulator can be run with other initial values) to determine new sample sizes necessary to achieve a desired power. If proportion means are lower or variance greater than we estimated here, a larger sample will be necessary to achieve equivalent power.

Discussion

Our simulations, using estimated values for strata proportions, suggest that a monitoring system capable of detecting even rather large declines (20%) in a track-plate index will require ≈1100 sampling units distributed across the range of fishers in California. And, because our parameter estimates are based on relatively few data it is possible that pilot implementation will produce results that suggest even more sampling units per stratum. We are encouraged, however, that our initial sample size estimates do not appear to be prohibitive, a result we attribute to our use of a realistic minimum power value (80%), a relatively large Type I error rate, and a one-sided alternative that precludes the detection of an increase in the index. If managers either cannot tolerate a 20% chance that a decline will be undetected or demand that population increases be as detectable as decreases, then monitoring effort and costs will increase. For example, if 90% power was required the number of samples would at least double in most cases (Figs. 3–5).

We began with the assumption that detecting declines only (one-tailed tests) would be sufficient information, given the anticipated increase in sampling effort to detect an increase or decrease (two-tailed test). Using one-tailed tests is a satisfactory way to minimize sampling effort, while assuring that declines in the populations of uncommon animals will be discovered (e.g., Kendall et al. 1992, Taylor and Gerrodette 1993). Required sample sizes to test a two-tailed alternative hypothesis were often 20–50% higher than the one-tailed alternative used in our simulations (H. Stauffer, *unpublished data*). Similarly, the sample size necessary to test the alternative that any two strata proportions (e.g., the proportions for any two regional or any two habitat strata) are different (two-tailed test) was from one to two times that necessary to distinguish a one-sided alternative, at 80% power (Cohen 1988:179–203). For one-sided tests, the strata proportion with the smaller value needs to be specified a priori, and a result in the unintended direction of the one-tailed test must be judged as statistically insignificant (Rice and Gaines 1994). Even if the index appears to increase significantly, failure to reject the null hypothesis can result in only one decision: that there was no change in the index.

If knowledge of declines as great as 50% were all that were required, the sample effort would be reduced greatly but at significant risk to losing the population before it could recover. Although this is not recommended, effort at this level would be better than no effort at all. For example, this may be a starting point for the development of a monitoring program for martens. Currently, marten populations in California appear well distributed throughout most of their historic range (Kucera et al. 1995), and occur at greater densities than fisher populations. A system designed to detect only catastrophic declines (such as 50%) may be appropriate until resources are available to increase the effort. With even a modest mean proportion of occupied units, say 20%, only ≈20 sampling units per stratum (200 statewide) would be necessary to detect a 50% decrease in the index. Fishers, on the other hand, require urgent conservation measures. They are much less common than martens, and we recommend that pilot implementation of a monitoring system capable of detecting 20% declines begin immediately. A system that detects a 20% decline with 80% confidence is probably the most liberal that should be established for fishers.

It is reassuring to realize that the variation around the mean proportion of occupied units across strata affects the power far less than the magnitude of the mean itself. Thus, we suffer little from the fact that fisher populations are nonuniformly distributed in California, being much more abundant in the Klamath and Southern Sierra Nevada regions than elsewhere. We are also confident that should our initial estimates of strata proportions prove to be inaccurate, the corrected estimates are likely to fall within the range of values used in the sensitivity analysis and we can easily recalculate the sampling effort required.

The samples necessary to achieve 80% power to detect a 20% decline in the statewide index may also be sufficient, in some cases, to compare pairs of regional and habitat strata proportion estimates (e.g., the Klamath-Sierra transition vs. the Southern Sierra Nevada). For example, the 115 sampling units that need to be sampled to detect a 20% decline in the statewide index are also sufficient to distinguish two strata proportions that differ by at least 13% (if only one, one-tailed comparison is planned). However, if more than one pairwise comparison is planned (necessitating a reduction in α), a two-tailed test is desired, or the proportions differ by <13% the number of samples per stratum needed to generate a statewide index value will be insufficient also to distinguish statistically individual strata proportion estimates.

If we could ignore the question of habitat differences and were satisfied with sampling to detect a statewide decline in occurrence only, the sampling effort per stratum could be reduced considerably. If this were the case the most economical method would be to sample only in habitats where the detection probabilities are highest (perhaps the LS habitat only) because higher proportion means result in higher power estimates for lower sample sizes (Fig. 3). The five-strata condition we simulated using specified values comes closest to this case; ≈850 sampling units per stratum would be necessary if we sample only where fisher densities are thought to be highest (compared to ≈1150 when habitat stratification is included). However, estimates would apply only to LS habitat and we would risk overlooking declines in marginal habitats that may, in fact, occupy most of the species range. For this reason we recommend against sampling only in areas with a high probability of occupancy.

Alternative methods to index fisher or marten populations, and to independently validate the index, have been considered and dismissed for a variety of reasons. Mark–recapture methods, particularly using photographic recapture (Mace et al. 1994), may be possible for individual populations but even at this level would probably produce highly variable estimates for fishers because of their extremely low densities. More importantly, the work would be prohibitive at the statewide scale we believe is the most relevant. Information from demographic studies provides better information on population status than simple surveys of presence/absence. However, these studies are impractical, especially for fishers, due to the difficulty of locating and monitoring natal dens and marking and tracking the fate of juveniles. Taylor and Gerrodette (1993) compared the conditions where survey methods were superior to demographic studies as a means of detecting declines in abundance. They found that power increased with decreasing population size when the demographic method is used; the opposite occurs with the survey method. In large study areas that include >100 individuals (the case for our proposal), they recommend the survey method be used. Of the survey methods available, track plates are easier to use and more cost effective at a large scale than cameras, and in California snow is too unpredictable, especially at elevations where fishers occur, for snowtracking to be a viable survey option. Finally, because both species are protected in California an experimental approach to validate an index using kill-trapping is out of the question.

We lack information to determine whether the monitoring approach we describe for indexing change in occurrence might also be useful for indexing changes in population abundance. Although others have used the pattern of occurrence of a species across its historic range as an index of its abundance and viability (Geissler and Fuller 1986, Lomolino and Channell 1995), the relationship between distribution and abundance is unknown for fishers or martens. A number of factors will confound the relationship between our index and population size. We would have to assume that individuals occupy relatively fixed home ranges during the sampling period, that transients are rare, and that population density or food abundance have relatively small changes on home range size. Accounting for the effect of transient individuals (a "floater" population) on our estimates (e.g., Bart 1995) is probably the biggest concern. Where fisher and marten spacing systems have been studied in detail, transient adults are uncommon (Arthur 1987, Paragi 1990; D. Harrison, *personal communication,* W. Krohn, *personal communication*). Female fishers in Maine have stable home range sizes from year to year and male home ranges expand briefly during the breeding season (Arthur et al. 1989). Adult male martens, also studied in Maine, were never observed to abandon home ranges, but some adult females did (D. Harrison, *unpublished data*). However, home range abandonment typically occurred in winter, or very early spring, so that by May virtually every animal in the study population was resident. If it could be demonstrated that transient individuals (dispersing juveniles and adults in the process of relocation) occur primarily during the fall and winter this bias could be reduced by restricting surveys to the early spring and summer, before either species exhibits significant extraterritorial movements or dispersal.

Data from fishers or martens are lacking, but the home ranges of other terrestrial mustelids have been reported to be both sensitive (e.g., Thompson and Colgan 1987) and insensitive (Lockie 1966) to changes in food availability. Lockie (1966) found that as a weasel (*Mustela nivalis*) population declined the remaining animals did not share the remaining area. He suspected that animals can hold only so much ground before suitable habitat becomes unoccupied. Moreover, the only territory-holding fishers that make forays outside their ranges are males and only during the short breeding season (Arthur et al. 1989). Evidence from short-tailed weasels (*Mustela erminea*) indicates that only the few dominant males exercise this option (Sandell 1986). Clearly, the effect of density, season, and food availability on fisher and marten home range size will need to be better understood before we can predict effects on our estimates. Although it appears that some actions can be taken to minimize the effects of confounding variables on the relationship between our index and population size, much more information will be necessary before we can be comfortable with the index as anything but a measure of change in distribution and occurrence.

Our approach is the first to estimate the sampling effort necessary to monitor change in an index of statewide population status of martens or fishers using nonlethal methods. It is an example of the type of exercise that will become increasingly necessary where commercial trapping is an unreliable, or socially unacceptable, means for indexing populations. Importantly, the method can simultaneously index and monitor as many sympatric carnivore species as are reliably attracted to meat baits and scent lures and whose spatial and social systems qualify them for consideration. A number of other carnivorous mammal species (e.g., bobcat, *Felis rufus*; gray fox, *Urocyon cinereoargenteus*; ringtail, *Bassariscus astutus*; spotted skunk, *Spilogale putorius*) are readily detected at track-plate stations (Raphael 1988; W. J. Zielinski, *unpublished data*). In locations where these, or other species, are sympatric with martens or fishers important information on their occurrence could be collected as well.

ACKNOWLEDGMENTS

We are grateful to the numerous government and consulting biologists who have conducted surveys that contributed to our recent knowledge of the distribution of martens and fishers in California. The work was supported by the Pacific Southwest Experiment Station and the Department of Mathematics, Humboldt State University. We thank M. Raphael, B. Noon, J. Bart, C. Hargis, and several anonymous reviewers for comments that significantly improved the manuscript. POWER is available on request from H. Stauffer.

LITERATURE CITED

Arthur, S. M. 1987. Ecology of fishers in southcentral Maine. Dissertation. University of Maine, Orono, Maine, USA.

Arthur, S. M., W. B. Krohn, and J. R. Gilbert. 1989. Home range characteristics of adult fishers. Journal of Wildlife Management **53**:674–679.

Azuma, D. L., J. A. Baldwin, and B. R. Noon. 1990. Estimating the occupancy of spotted owl habitat areas by sampling and adjusting for bias. USDA Forest Service Pacific Southwest Research Station General Technical Report **PSW-124**.

Barrett, R. H. 1983. Smoked aluminum track plots for determining furbearer distribution and relative abundance. California Fish and Game **69**:188–190.

Bart, J. 1995. Evaluation of population trend estimates calculated using capture–recapture and population projection methods. Ecological Applications **5**:661–671.

Buck, S., C. Mullis, and A. Mossman. 1983. Corral Bottom-Hayfork Bally fisher study. Final report to USDA Forest Service. Humboldt State University, Arcata, California, USA.

Buskirk, S. A., and L. L. McDonald. 1989. Analysis of variability in home-range size of the American marten. Journal of Wildlife Management **53**:997–1004.

Buskirk, S. A., and L. F. Ruggiero. 1994. The American marten. Pages 7–37 *in* L. F. Ruggiero, K. B. Aubry, S. W. Buskirk, L. Jack Lyon, and W. J. Zielinski, editors. The scientific basis for conserving forest carnivores: American marten, fisher, lynx, and wolverine in the western United States. USDA Forest Service General Technical Report **RM-254**.

Cochran, W. G. 1977. Sampling techniques. Third edition. John Wiley & Sons, New York, New York, USA.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences. Second edition. Lawrence Erlbaum, Hillsdale, New Jersey, USA.

de la Mare, W. K. 1984. On the power of catch per unit effort series to detect declines in whale stocks. Report of the International Whaling Commission **34**:657–661.

Diefenbach, D. R., M. J. Conroy, R. J. Warren, W. E. James, L. A. Baker, and T. Hon. 1994. A test of the scent-station survey technique for bobcats. Journal of Wildlife Management **58**:10–17.

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

Fowler, C. H., and R. T. Golightly. 1994. Fisher and marten survey techniques on the Tahoe National Forest. California Fish and Game Nongame Bird and Mammal Section Report **94–9**.

Geissler, P. H., and M. R. Fuller. 1986. Estimation of the proportion of an area occupied by an animal species. Pages 533–538 *in* Proceedings of the Section on Survey Research Methods of the American Statistical Association. American Statistical Association, Washington, D.C., USA.

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

Grinnell, J., J. S. Dixon, and J. M. Linsdale. 1937. Fur-bearing mammals of California. University of California Press, Berkeley, California, USA.

Hawley, V. D., and F. E. Newby. 1957. Marten home ranges and population fluctuations. Journal of Mammalogy **38**: 174–184.

Jones, J. L. 1991. Habitat use of fisher in northcentral Idaho. Master's thesis. University of Idaho, Moscow, Idaho, USA.

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

Kucera, T. E., W. J. Zielinski, and R. H. Barrett. 1995. The current distribution of American martens in California. California Fish and Game **81**:96–103.

Kuehl, R. O. 1994. Statistical principles of research design and analysis. Wadsworth, Belmont, California, USA.

Lancia, R. A., J. D. Nichols, and K. H. Pollock. 1994. Estimating the number of animals in wildlife populations. Pages 215–253 *in* T. A. Bookhout, editor. Research and management techniques for wildlife and habitats. Fifth edition. Wildlife Society, Bethesda, Maryland, USA.

Lockie, J. D. 1966. Territory in small carnivores. Symposium of the Zoological Society London **18**:143–165.

Lomolino, M. V., and R. Channell. 1995. Splendid isolation: patterns of geographic range collapse in endangered mammals. Journal of Mammalogy **76**:335–347.

Mace, R. D., S. C. Minta, T. L. Manley, and K. E. Aune. 1994. Estimating grizzly bear population size using camera sightings. Wildlife Society Bulletin **22**:74–83.

Mapstone, B. D. 1995. Scalable decision rules for environmental impact studies: effect size, type I, and type II errors. Ecological Applications **5**:401–410.

Mendenhall, W., D. D. Wackerly, and R. L. Scheaffer. 1990. Mathematical statistics with applications. Fourth edition. Duxbury, Belmont, California, USA.

Paragi, T. F. 1990. Reproductive biology of female fishers in southcentral Maine. Master's thesis. University of Maine, Orono, Maine, USA.

Parkinson, E. A., J. Berkowitz, and C. J. Bull. 1988. Sample size requirements for detecting changes in some fisheries statistics from small trout lakes. North American Journal of Fisheries Management **8**:181–190.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences **47**:2–15.

Peterman, R. M., and M. J. Bradford. 1987. Statistical power of trend in fish abundance. Canadian Journal of Fisheries and Aquatic Science **44**:1879–1889.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture–recapture experiments. Wildlife Monograph **107**.

Powell, R. A. 1994. Structure and spacing of *Martes* populations. Pages 101–121 *in* S. A. Buskirk, A. S. Harestad, M. G. Raphael, and R. A. Powell, editors. Martens, sables, and fishers: biology and conservation. Cornell University Press, Ithaca, New York, USA.

Powell, R. A., and W. J. Zielinski. 1994. The fisher. Pages 38–73 *in* L. F. Ruggiero, K. B. Aubry, S. W. Buskirk, L. Jack Lyon, and W. J. Zielinski, editors. The scientific basis for conserving forest carnivores: American marten, fisher, lynx, and wolverine in the western United States. USDA Forest Service General Technical Report **RM-254**.

Quinn, J. F., and A. E. Dunham. 1983. On hypothesis testing in ecology and evolution. American Naturalist **122**:602–617.

Raphael, M. G. 1988. Long-term trends in abundance of amphibians, reptiles, and mammals in Douglas-fir forests of northwestern California. Pages 23–31 *in* R. C. Szaro, K. E. Severson, and D. Patton, technical coordinators. Management of amphibians, reptiles and small mammals in

North America. USDA Forest Service General Technical **RM-166**.

Raphael, M. G. 1994. Techniques for monitoring populations of fishers and American martens. Pages 224–240 *in* S. A. Buskirk, A. S. Harestad, M. G. Raphael, and R. A. Powell. Martens, sables, and fishers: biology and conservation. Cornell University Press, Ithaca, New York, USA.

Raphael, M. G., and B. G. Marcot. 1986. Validation of a wildlife-habitat-relationships model: vertebrates in a Douglas-fir sere. Pages 129–138 *in* J. Verner, M. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison, Wisconsin, USA.

Rice, W. R., and S. D. Gaines. 1994. 'Heads I win, tails you lose': testing directional alternative hypotheses in ecological and evolutionary research. Trends in Ecology and Evolution **9**:235–237.

Rosenberg, K. V., and M. G. Raphael. 1986. Effects of forest fragmentation on vertebrates in Douglas-fir forests. Pages 263–272 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison, Wisconsin, USA.

Roughton, R. D., and M. W. Sweeney. 1982. Refinements in scent-station methodology for assessing trends in carnivore populations. Journal of Wildlife Management **46**: 217–229.

Samuels, M. L. 1989. Statistics for the life sciences. Dellen, San Francisco, California, USA.

Sandell, M. 1986. Movement patterns of male stoats *Mustela erminea* during the mating season: differences in relation to social status. Oikos **47**:63–70.

Schempf, P. F., and M. White. 1977. Status of six furbearer populations in the mountains of northern California. USDA Forest Service, Pacific Southwest Region, San Francisco, California, USA.

Scheaffer, R. L., W. Mendenhall, and L. Ott. 1990. Elementary survey sampling. Fourth edition. PWS-Kent, Boston, Massachusetts, USA.

Seber, G. A. F. 1982. The estimation of animal abundance. Second edition. Griffin, London, England.

Sokal, R. R., and F. J. Rohlf. 1987. Introduction to biostatistics. Second edition. W. H. Freeman, New York, New York, USA.

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.

Toft, C. A., and P. J. Shea. 1983. Detecting community-wide patterns: estimating power strengthens statistical inference. American Naturalist **122**:618–625.

Thompson, I. D., and P. W. Colgan. 1987. Numerical responses of marten to a food shortage in northcentral Ontario. Journal of Wildlife Management **51**:824–835.

Verner, J. 1983. An integrated system for monitoring wildlife on the Sierra National Forest. Transactions of the North American Wildlife and Natural Resources Conference **48**: 355–366.

Verner, J., and J. G. Kie. 1988. Population monitoring: an essential link between theoretical and applied conservation biology. Transactions of the Western Section of the Wildlife Society **24**:18–25.

Zielinski, W. J., and T. E. Kucera. 1995. Survey methods for the detection of wolverines, lynx, fishers and martens. USDA Forest Service General Technical Report **PSW-157**.

Zielinski, W. J., T. E. Kucera, and R. H. Barrett. 1995. The current distribution of fishers in California. California Fish and Game **81**:104–112.

Zielinski, W. J., and R. L. Truex. 1995. Distinguishing tracks of marten and fisher at track-plate stations. Journal of Wildlife Management **59**:571–579.

Zielinski, W. J., R. L. Truex, C. V. Ogan, and K. Busse. *In press*. Detection surveys for fishers and martens in California, 1989–1994: summary and interpretations. Proceedings of the Second International *Martes* Symposium, Edmonton, Alberta. Provincial Museum of Canada, Edmonton, Alberta, Canada.

APPENDIX

POWER (see *Acknowledgments*) was developed using the programming language available in Minitab (Minitab Statistical Software, Version 10.1, State College, Pennsylvania). Statistical functions are already available in Minitab for generation of random variable values from varying distributions and for testing procedures.

POWER executes as follows:

(1) input is read from a data file containing the input parameter values for the simulation run;
(2) strata probabilities p_i are assigned, either from a beta distribution (a flexible, two-parameter continuous distribution from 0 to 1) using a specified mean and standard deviation, or specified from input;
(3) binomial random variable values x_i (= the number of occupied sampling units in the ith stratum) are generated by the program, simulating the sampling of n_i sampling units for presence/absence, based upon the strata probabilities p_i and the binomial model $B(x_i;n_i,p_i)$;
(4) strata proportion estimates $\hat{p}_i = x_i/n_i$ (the binomial random variable value over the sample size) are calculated;
(5) the strata probabilities p_i are reduced by 20 and 50% decline; if p_i is the probability before decline, then $p_i' = 1 - (1 - p_i)^{(1-r)}$ is the probability after a decline r (Kendall et al. 1992);
(6) steps 3–4 are repeated, using the reduced strata probabilities;
(7) the paired untransformed and transformed (using the 2× arcsine $(\sqrt{\hat{p}_i})$ transformation; see Cohen 1988) strata proportion estimates are compared over the two survey years, before and after decline, using a one-tailed, paired t test;
(8) steps 2–7 are repeated for as many simulations as requested;
(9) the proportion of simulation test decisions that rejected the null hypothesis is calculated, for both untransformed and transformed strata proportions. These proportions are the estimated power of the test under these sampling conditions.