Special Section: The Value and Utility of Presence–Absence Data to Wildlife Monitoring and Research

A FIELD-BASED EVALUATION OF A PRESENCE–ABSENCE PROTOCOL FOR MONITORING ECOREGIONAL-SCALE BIODIVERSITY

PATRICIA N. MANLEY,[1] U.S. Forest Service, Pacific Southwest Research Station, 2121 Second Street, Suite A-101, Davis, CA 95616, USA
MATTHEW D. SCHLESINGER, U.S. Forest Service, Pacific Southwest Research Station, 2121 Second Street, Suite A-101, Davis, CA 95616, USA, and University of California, Department of Environmental Science and Policy, One Shields Avenue, Davis, CA 95616, USA
JULIE K. ROTH, U.S. Forest Service, Lake Tahoe Basin Management Unit, 35 College Drive, South Lake Tahoe, CA 96150, USA
BEATRICE VAN HORNE, U.S. Forest Service, National Office, Rosslyn Plaza C - Fourth Floor, 1601 North Kent Street, Arlington, VA 22209, USA

Abstract: Declines in populations of many species across the United States are prompting the need for monitoring programs that can detect population change effectively for large suites of species. The proliferation of reliable standardized monitoring techniques enhances the potential that we may meet these burgeoning information needs. However, multiple-species monitoring approaches are still a rarity. In an attempt to meet monitoring information needs on National Forest System (NFS) lands, the U.S. Forest Service developed the Multiple Species Inventory and Monitoring (MSIM) protocol, an approach that employs a variety of vertebrate and habitat survey methods to obtain presence–absence data for a broad spectrum of species at a systematic array of sample points throughout ecoregions. A recent evaluation of the MSIM protocol predicted that its implementation across the Sierra Nevada ecoregion would detect a 20% change in proportion of points occupied for 76% of the species with 80% confidence and power. This evaluation was based on a qualitative estimate of probability of detection and basic habitat association information to estimate probability of presence. We conducted a field test to evaluate the ability of a multi-taxonomic, presence–absence monitoring approach to detect and estimate population change at multiple scales and to meet the monitoring needs of land management agencies. Our field data consisted of bird, mammal, amphibian, and reptile surveys conducted at terrestrial and aquatic sites in the Lake Tahoe basin, located in the central Sierra Nevada. We used a maximum likelihood function to quantitatively estimate the probability of detection and sampling adequacy of species and species groups. We compared these observed estimates to the previously derived expected estimates. We detected 185 species by field testing, including 89% of the expected species and several species that were not expected. Observed probabilities of detection were generally lower than we expected. Sampling adequacy was also lower than we expected but still relatively high, with 66% of the Lake Tahoe species and 47% of all Sierra Nevada species adequately sampled. Species with small geographic ranges, narrow habitat specificity, or small population sizes were considered rare, and they were not sampled as adequately as other species groups. Still, 69 of the rarest species, including 12 species with state or federal designation as endangered, threatened, or sensitive, were estimated to be adequately sampled. Our analysis highlighted the ability of a broad-scale, multi-taxonomic presence–absence monitoring approach, such as the MSIM protocol, to detect population change for a large number and wide variety of vertebrate species. The outcome of our test also suggests that many data yields would be garnered at the scale of National Forests and National Parks, such as distribution and occurrence data for most species, including little-known species, population trend data for many species, habitat relationships and habitat condition data, and data on association strengths of indicator species.

JOURNAL OF WILDLIFE MANAGEMENT 69(3):950–966; 2005

Key words: biodiversity, detection, ecoregion, National Forest System, monitor, presence–absence data, vertebrates, Sierra Nevada.

Concerns for the fate of biological diversity in the United States and around the world are growing (Flather et al. 1999, United Nations 2002). Ecoregional assessments conducted over the past 10 years in the United States suggest that many ecosystems may be on the brink of substantial losses in biological diversity and degradation of allied ecosystem functions and services. Four ecoregional assessments conducted around the country found that 29 to 46% of all vertebrate species, ranging from 90 to over 200 species in a given ecoregion, were considered at risk of population decline (Southern Appalachian Man and the Biosphere 1996, Stephenson and Calcorone

[1] E-mail: pmanley@fs.fed.us

1999, Wisdom et al. 2000, U.S. Forest Service 2001). Single-species approaches to conservation, management, and monitoring are insufficient to combat the threat of such substantial losses in biological diversity. Multiple-species, ecosystem-based monitoring approaches are needed to provide reliable, timely, and informative measures of change in the status of populations, communities, and biological diversity.

In the face of ongoing population growth in the United States (~10% per decade over the next 50 years; U.S. Census Bureau 2004), public lands will play an increasingly pivotal role in conserving biological diversity at ecoregional and national scales. Federal land management agencies are increasing their investments in the development of regional and national monitoring approaches to more closely inventory and monitor their biological resources (e.g., Landres et al. 1994, Gillespie 1999, Hall and Langtimm 2001, Kingsbury and Gibson 2002, McDonald et al. 2002).

The building blocks needed to construct successful large-scale multi-taxonomic monitoring programs are maturing, enhancing the prospect of meeting biodiversity monitoring needs in a timely manner. For example, standardized survey protocols are being developed for a growing array of taxa (e.g., Ralph et al. 1993, Heyer et al. 1994, Wilson et al. 1996), and the accuracy and strength of inference garnered from more affordable presence–absence monitoring efforts are enhanced by advances in sampling design (e.g., Thompson et al. 1998) and statistical analysis techniques, particularly pertaining to imperfect detection probabilities and presence–absence data (Azuma et al. 1990, MacKenzie et al. 2002, MacKenzie et al. 2003). However, fully developed, large-scale, multi-taxonomic protocols for vertebrate monitoring are few.

Through a collaborative research–management partnership, the U.S. Forest Service developed the MSIM protocol, a nationally standardized protocol for monitoring large numbers of plant and animal species (Manley and Van Horne 2005). The primary objective of the MSIM protocol is to estimate status and change in populations and habitat conditions of a broad suite of species throughout large geographic areas. The MSIM protocol was developed to meet biodiversity monitoring needs across the 76 million ha of NFS lands in the United States, most of which are located in the western states. Population status and change are represented by presence–absence data and expressed as the proportion of occupied sample points per sample period at a variety of scales.

Manley et al. (2004) conducted a quantitative evaluation of the potential information yields of the national MSIM protocol using a simulated implementation scenario for the Sierra Nevada ecoregion in California. They predicted the number and characteristics of species that would be adequately sampled by MSIM's primary survey methods conducted at systematically established grid points throughout federal lands (6.5 million ha) in the Sierra Nevada. Adequate sampling was defined as the minimum sample size needed to detect a ≥20% relative change in the proportion of sample points occupied between 2 points in time with 80% confidence and power. Their evaluation concluded that >75% of all vertebrate species in the Sierra Nevada would be adequately sampled, and the species adequately sampled represented a wide array of ecological characteristics (Manley et al. 2004). This evaluation generated promising evidence that multi-taxonomic presence–absence-based approaches such as the MSIM protocol could yield reliable population data for a large number and wide variety of species over broad spatial scales. Field testing the MSIM protocol was the next logical step to generate empirically based results reflecting its ability to effectively monitor populations and biological diversity.

We implemented components of the MSIM protocol at a grid of points that mimicked implementation of the MSIM protocol on a systematic grid at the scale of a National Forest from 2002 to 2004. Our objectives were to derive empirically based estimates of probability of detection for all species in our study area and then make inferences about the potential success of a multiple species monitoring approach to meet land management monitoring objectives throughout an ecoregion. Based our field test of the MSIM protocol, we evaluated its potential success in terms of implementation on federal lands in the Sierra Nevada, using the same criteria as Manley et al. (2004) to define adequate sampling (minimum sample size needed to detect a ≥20% relative change in occupancy of sample points between 2 points in time with 80% confidence and power). Specifically, we wanted to address the following questions: (1) How representative are species that are estimated to be adequately sampled in the study area in terms of number, taxonomic diversity, and vulnerability to population decline? (2) What do the results suggest about how well the MSIM protocol would monitor vertebrate species throughout the Sierra Nevada? (3) How

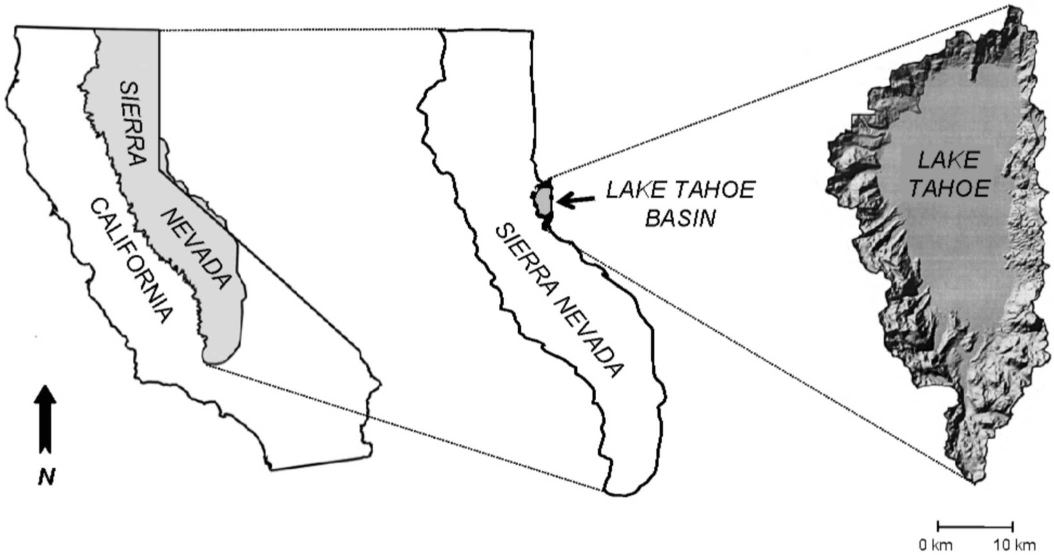

Fig. 1. Field data were collected in the Lake Tahoe basin of the Sierra Nevada, 2002–2004, and the results were applied across the Sierra Nevada ecoregion.

did the observed number and representativeness of species predicted to be adequately sampled throughout the Sierra Nevada compare to that expected based on Manley et al. (2004)?

STUDY AREA

We conducted our field test of the MSIM protocol in the central Sierra Nevada in the 131,000-ha Lake Tahoe basin. A large proportion of the basin (38%) is occupied by Lake Tahoe, leaving 81,000 ha of upland area remaining in the watershed (Fig. 1). The Lake Tahoe basin had many characteristics that made it suitable for a Sierra Nevada field test. The Lake Tahoe basin spans a broad range of elevations from 2,000 m to over 3,500 m and includes all life zones in the greater Sierra Nevada except those at the lowest elevations, including lower and upper montane, subalpine, and alpine zones (Manley et al. 2000). Similar to the Sierra Nevada, it is predominantly public land and consists of almost 80% NFS lands (approx 65,000 ha). Of the 465 vertebrate species in the Sierra Nevada, 209 were known to have breeding populations in the Lake Tahoe basin. The vertebrate fauna of Lake Tahoe fairly represented the diversity of vertebrates throughout the Sierra Nevada both taxonomically (60% of the Sierran aquatic reptiles, 58% of the birds, 38% of the small mammals, and 25% of the aquatic amphibians) and across ecological characteristics (e.g., habitat specificity).

METHODS

Sampling Design

Data used in our analysis were garnered from a monitoring program designed and implemented for the U.S. Forest Service Lake Tahoe Basin Management Unit (see Roth, J. K., P. N. Manley, M. M. McKenzie, and M. D. Schlesinger, Multiple Species Inventory and Monitoring 2002 monitoring report, Lake Tahoe Basin Management Unit, South Lake Tahoe, California, USA). We created a hexagonal grid across the Lake Tahoe basin using the spacing parameters of the systematic grid of the Forest Inventory and Analysis (FIA) program, a nationwide program that monitors the composition and structure of forested ecosystems (Roesch and Reams 1999). The current FIA design consists of a 2.7-km radius hexagonal grid, with 1 FIA sample point randomly located within each hexagon. The resulting density of points amounts to approximately 250 points per 600,000 ha, an area representative of a typical National Forest or National Park.

Given the small area of the Lake Tahoe basin, we randomly established 4 sample points in each of our hexagons, and we prorated the number of points in each hexagon based on the proportion of NFS lands in the hexagon. The result was 46 hexagons with 1 or more points on NFS lands in the basin, for a total of 100 sample points. In addition to the hex-based sample points, we randomly selected 158 aquatic sample sites consisting of lakes, ponds, and wet meadows, stratified by elevation and size, to sample aquatic-associated vertebrates.

Field Data Collection

We implemented 6 of the primary survey methods specified in the MSIM protocol: terrestrial point counts, Sherman live trapping, trackplate and camera surveys, bat mistnetting, aquatic visual encounter surveys, and aquatic point counts. The 3 MSIM survey methods we did not implement were terrestrial visual encounter surveys, nocturnal broadcast calling, and Tomahawk live trapping (Manley et al. 2004). Primary survey methods in the MSIM protocol are commonly employed, standardized survey methods that detect a large number and variety of species per unit effort. The 6 methods we employed provided presence–absence data on landbirds, aquatic birds, small mammals, midsized and large mammals, bats, amphibians, and reptiles. Most data collected at each point occurred within a 200-m radius sample hexagon centered on each sample point.

Terrestrial Point Counts.—We conducted point counts to detect landbirds (Ralph et al. 1993, 1995). We surveyed 84 sample points with terrestrial point counts, 44 in 2002 and 40 in 2004. We located 7 count stations at the center and vertices of the sample hexagon. We recorded the number of individuals of each species of vertebrate seen or heard during a 10-min count period, and we repeated the survey 3 times during the breeding season from late May through July.

Sherman Live Trapping.—We used Sherman live trapping to detect small mammals (Jones et al. 1996). We conducted trapping from early June to early September at 80 sample points, 40 in 2002 and 40 in 2004. We placed extra-long traps (7.5 × 9.4 × 37.5 cm) 15 m apart around the perimeter and down the center of the sample hexagon. In 2004, we alternated extra-large traps (10 × 11 × 37.5 cm) to increase capture rates for larger-bodied squirrels (e.g., Slade et al. 1993). Traps were set, opened, and baited with grains and seeds in the afternoon of the first day, checked twice daily, and collected on the afternoon of the fourth day for a total of 3 trap days. We treated each trap day as a visit.

Baited Trackplates and Cameras.—We used baited trackplates and cameras to detect mid-sized to large mammals, primarily carnivores (Zielinski and Kucera 1995). We surveyed 22 sample points in 2002. Trackplate stations consisted of sooted aluminum plates baited with meat, and we identified to species the tracks created by animals walking on the sooted surface. We baited camera stations with meat and vegetables; as animals visited the bait, they triggered the camera. We located 10 devices (6 trackplates and 4 cameras) in a star-shaped array centered on each sample point. We located 1 trackplate at the center sample point, and we located 5 others 500 m away and equidistant from one another. We located 1 camera 100 m away from the center trackplate, and we placed 3 others 100 m away from a random selection of 3 of the remote trackplates. We checked all 10 devices every other day over a 10-day period; we considered each of these 5 checks a visit.

Bat Mistnetting.—We used mistnets to detect bat species. Mistnets are commonly used to detect bats in a variety of situations (Jones et al. 1996, Kunz et al. 1996). We surveyed 3 sample sites (bodies of water or forest openings) within a 1-km radius sample unit centered on each of 30 sample points. We set 3 to 5 nets, each 4 m tall and extending 6 to 18 m across, at each site and opened them at dusk for 3 to 4 hrs. We surveyed each of the 90 sample sites 3 to 6 times per year between June and September 2002, and we also surveyed 12 of these sites in 2001, for a total of 102 site/year combinations that we treated as independent samples.

Aquatic Visual Encounter Surveys.—We conducted aquatic visual encounter surveys to detect aquatic and riparian-associated herpetofauna (Crump and Scott 1994, Fellers and Freel 1995). We surveyed each aquatic sample site twice during the spring/summer season, with surveys separated by at least 2 weeks. At lakes and ponds, we walked 100% of the perimeter at a pace of approximately 100 m per 15 min, and we visually scanned all substrates, searched through emergent vegetation with a long-handled dip-net, and overturned rocks, logs, and debris to reveal amphibians and reptiles (Fellers and Freel 1995). At each wet meadow, we meandered from side to side covering its entire length and width. We surveyed the 158 aquatic sample sites twice per year between June and early September in 2002 and/or 2004 for 211 site/year combinations that we treated as independent samples.

Aquatic Point Counts.—We conducted point counts to detect waterbirds at aquatic sample sites. We conducted ≥2 point counts spaced 500 m apart around the perimeter of each aquatic sample site, with each count lasting 20 min. We surveyed each site twice during the breeding season in the same year as the aquatic visual encounter survey.

Estimation of Probability of Detection and Sampling Adequacy

We based all estimates of detectabilty and sampling adequacy on evaluations of the 6 MSIM survey methods to detect target species in the Lake Tahoe basin. Schlesinger and Romsos (2000) identified 212 species suspected to have established populations in the Lake Tahoe basin. Based on local knowledge (W. Richardson, University of Nevada, personal communication), we eliminated 14 and added 11 species to derive a list of 209 species with established populations in the basin. Consistent with the selection of primary survey methods for each species by Manley et al. (2004), we expected the 6 survey methods we tested to detect 183 of the 209 species, including 105 species by terrestrial point counts, 21 species by Sherman live traps, 37 species by aquatic surveys (visual encounters and point counts), 10 species by trackplates and cameras, and 10 species by mistnets. Hereafter, our analysis is based on the 183 expected Lake Tahoe species because we only implemented a subset of the MSIM primary survey methods.

We used a maximum likelihood function to estimate the per-visit probability of nondetection (q) of each species for each survey method (MacKenzie et al. 2002) using PROC NLMIXED (SAS version 8.3). We assumed that presence and detection probabilities were constant across monitoring sites and the sample period. The resulting model for calculating likelihood for detection and presence, as per MacKenzie et al. (2002), is as follows:

$$L(\psi,q)=\left[\psi^{n.}\prod_{t=1}^{T}(1-q_t)^{n_t}q_t^{n.-n_t}\right]\times\left[\psi\prod_{t=1}^{T}q_t+(1-\psi)\right]^{N-n.} \tag{1}$$

where ψ is the probability that a species is present, $q_t = 1 - p_t$ where p_t is the probability that a species will be detected when present at time t, N is the total number of surveyed sites, T is the number of sampling occasions (i.e., visits), n_t is the number of sites with detections at time t, and $n.$ is the total number of sites at which the species was detected at least once.

We calculated detection probabilities for each species per survey method based on species detected per site visit per survey method. We estimated standard errors for q estimates ($\hat{q}$) as part of the iterative process to determine the maximum likelihood estimates for q. The large number of surveys associated with each survey method ($N * T$ = 66 to 306) enabled us to use the approximate standard error estimate consisting of the negative of the inverse of the observed Fisher information statistic determined by the second partial derivative of the log-likelihood function with respect to each of the parameters (Lebreton et al. 1992, Pawitan 2001).

The number of sites sampled per survey method in the field test (n = 22 to 158 sites) was 1 to 2 orders of magnitude smaller than the number of FIA points on NFS lands in the Sierra Nevada (n = 2,760). Given that standard errors are affected by sample size (Zar 1998), we generated estimates of q and its standard error, $\sigma_{\hat{q}}$, using a nonparametric Monte Carlo approach consisting of random selection with replacement of sample sites to build sample sizes commensurate with the number of FIA points within the species' ranges throughout the Sierra Nevada ecoregion (J. Baldwin, Pacific Southwest Research Station, personal communication). We simulated 3 sample sizes (500, 1,500, and 2,500 sample sites) corresponding to restricted, moderate, and large geographic ranges of associated species. We assigned species to range size classes as follows: 44 species had ≤1,000 FIA points in their range, and $\hat{q}$ was based on 500 points; 92 species had 1,001 to 2,000 FIA points in their range, and $\hat{q}$ was based on 1,500 points; and 47 species had >2,000 FIA points in their range, and $\hat{q}$ was based on 2,500 points.

We then followed the approach of Manley et al. (2004) to assess sampling adequacy. The proportion of monitoring sites at which a species is observed (i.e., the probability of observation) during a sample period t (P_t), is a function of the species' (1) probability of presence (p_p), and (2) probability of detection per survey method if present (p_d). The relationship between the 3 probabilities is:

$$P_t = p_p * p_d \tag{2}$$

We estimated the probability of presence (p_p) for each species based on the proportion of the species' range in the Sierra Nevada that consisted of suitable habitat, based on species range maps for the Sierra Nevada obtained from the California Wildlife Habitat Relationships Program (CDFG 2000; see Manley et al. 2004). As per Manley et al. (2004), we assigned each species 1 of 3 levels of

p_p: those with <30% of their range in suitable habitat were assigned a p_p of 0.1, those with 30–69% of their range in suitable habitat were assigned a p_p of 0.5, and those with ≥70% of their range in suitable habitat were assigned a p_p of 0.8. We used the 3 levels of p_p as opposed to the actual values because we wanted to limit the differences between our estimates and those of Manley et al. (2004) to factors associated with detectability. However, it is unlikely that using the 3 representative levels of p_p, as opposed to the estimated values, affected the outcome of the analysis (P. Manley, U.S. Forest Service, Pacific Southwest Research Station, unpublished data).

We calculated the probability of detection across the multiple visits associated with each survey method (p_d) as:

$$p_d = 1 - \hat{q}^T \quad (3)$$

where $\hat{q}$ is the per-visit estimate of probability of nondetection, and T is the number of visits conducted as part of the survey method. We calculated the 90% confidence interval (CI) for p_d for each species and used the lower bound as the minimum p_d to estimate N^{min}, the minimum number of FIA points needed in each species' range to meet the criteria of an adequate sample. The lower bound of the 90% CI for p_d provided a conservative estimate of sample size requirements for each species. We derived the 90% CI for each species using the approximate standard deviation of $p_d = 1 - \hat{q}^T$ calculated with the Delta method (Bishop et al. 1975) as follows:

$$\mathrm{SD}(p_d) = T_i * \hat{q}^{T_i - 1} * \sigma_{\hat{q}} \quad (4)$$

where T_i is the number of visits for survey method i, and $\sigma_{\hat{q}}$ is the standard error of $\hat{q}$. Therefore, the 90% confidence lower bound is given by:

$$\begin{aligned}\mathrm{Min}(\hat{p}_d) &= \mathrm{Lower}\ (\mathrm{CI}_{90\%}, p_d)\\ &= \hat{p}_d - 1.644 * \mathrm{SD}\ (p_d)\end{aligned} \quad (5)$$

We determined the minimum sample size, N^{min}, requirements based on the proportion (P_1) of sample sites estimated to have observations during the first sample period, the effect size ($\delta = 0.2 * P_1$), the prescribed error rates (α and $\beta = 0.2$), the direction of change desired to be detected (2-tailed), and site correlation ($\rho = 0.9$) between sample periods (Hoel et al. 1971*a*, Sokal and Rohlf 1995). We used a 2-tailed test because it is a more rigorous test of the detection adequacy, and we always selected the larger of the 2 sample size estimates to increase the rigor of our evaluation. As in Manley et al. (2004), we assumed that sites would be remeasured with minimal error, so we modeled site correlation between sample periods as high (0.90) but not perfect (1.0). Thus, we estimated the N^{min} necessary to detect a ≥20% change between 2 sample periods for a given species using the normal approximation (Fleiss 1981):

$$N^{\min} > \frac{\left[z_\alpha \times \sigma_o + z_\beta \times \sigma_a\right]^2}{(\delta \times P_1)^2} \quad (6)$$

where z_α and z_β represent the 2-tailed critical values from a normal distribution, and σ_o and σ_a represent standard deviations of the difference between $\hat{P}_1$ and $\hat{P}_2$ under the null and alternative hypothesis, respectively (Fleiss 1981).

We calculated the variance based on a binomial distribution, and we assumed that P_t was approximately normally distributed (Hoel et al. 1971*a,b*). Variance associated with binomial distributions is greatest at the midpoint (0.50) and tapers toward 0 and 1 from the midpoint (Zar 1998); therefore, associated sample size requirements were asymmetrical, with the larger value associated with increases when $\hat{P}_1 \le 0.5$ and declines when $\hat{P}_1 > 0.5$. We calculated the standard deviation of the difference between $\hat{P}_1$ and $\hat{P}_2$ using the standard formula:

$$\sigma_i / \sqrt{N} = \sqrt{\sigma^2_{1i} + \sigma^2_{2i} - 2\rho\sigma_{1i}\sigma_{2i}} / \sqrt{N} \quad (7)$$

where, for the null hypothesis of no change ($i = o$), σ^2_{1o} and $\sigma^2_{2o} = \hat{P}_1(1 - \hat{P}_1)$, and for the alternative hypothesis of ≥20% change ($i = a$), $\sigma^2_{1a} = \hat{P}_1(1 - \hat{P}_1)$, $\sigma^2_{2o} = \hat{P}_2(1 - \hat{P}_2)$, $\hat{P}_2 = (1 - 0.20)\ \hat{P}_1$, and N = sample size.

Evaluation of Sampling Adequacy Across Taxa

We evaluated the ability of the MSIM protocol to represent an array of taxa and to effectively monitor biological diversity. Most natural communities consist of few common and many rare species (Preston 1948, 1962), and rare species are more vulnerable than common species to population decline and extinction through a variety of processes (Terborgh 1974, Simberloff 1986, Lawton 1994, Davies et al. 2004, Henle et al. 2004). Thus, the ability to effectively monitor rare species is essential to meet the goals of providing a sensitive measure of biological diversity and providing information critical to the conservation of species at greatest risk. We summarized our results across all species, by taxonomic group and by 3 factors commonly associated with rarity and vulnerability to population decline (Rabinowitz 1981, Rabinowitz et al. 1986, Davies et al. 2004, Henle et al. 2004): geographic range, habitat specificity, and population size. First we addressed the number and characteristics of species occurring in the Lake Tahoe basin that were estimated to be adequately sampled throughout out the Sierra Nevada, and then we expanded these estimates to all species in the Sierra Nevada.

We based all classifications used in the analysis of sampling adequacy on the same data and derived them in the same manner as Manley et al. (2004) described for their simulations. We evaluated taxonomic representation by vertebrate class; birds and mammals were further subdivided into species groups targeted by the primary survey methods. We identified 2 conditions of geographic range based on the number of FIA points within each species' range in the Sierra Nevada; small distributions were those with 1 to 1,000 FIA points, and large distributions were those with >1,000 FIA points. We derived 2 conditions of habitat specificity (narrow and broad) based on the number of combinations of associated vegetation type, seral stage, and canopy cover class that CDFG (2000) ranked as suitable for each species; narrow habitat specificity was equivalent to Manley et al.'s (2004) low, and broad was equivalent to the combination of their moderate and high categories. We used 2 population size conditions: small was ≤1,000 individuals, and large was >1,000. Population size values were available for all Lake Tahoe species, but they were not available for 56 of the remaining 282 species in the Sierra Nevada, most of which had a range that overlapped to only a small extent with the Sierra Nevada. For the purposes of our analysis, we assumed that population sizes were small for these 56 species. We also evaluated sampling adequacy for species of concern that were defined as those with federal or state designation as threatened, endangered, or sensitive.

The 2 conditions of each rarity factor represented values associated with lesser and greater vulnerability to population decline, and in all cases, the more constrained condition (i.e., small geographic range, narrow habitat specificity, small population size) conferred greater vulnerability. All possible combinations of the 2 conditions for each of the 3 rarity factors resulted in 8 categories that represented forms and degrees of rarity (as per Rabinowitz et al. 1986). In general, the degree of rarity and vulnerability to population decline and extinction was greater as the number of constrained factors increased (Reed 1992). Therefore, we grouped the 8 rarity categories to represent 4 degrees of rarity: (1) common species, those for which no rarity factors exhibited a constrained condition; (2) moderately rare species, those for which 1 of the 3 factors was constrained; (3) highly rare species, those for which 2 of the 3 factors were constrained; and (4) extremely rare species, those for which all 3 factors were constrained.

We determined sampling adequacy for each of the 183 Lake Tahoe species by comparing N^{min} to the number of FIA points in each species' range; if N^{min} for a species was less than or equal to the number of FIA points in the species' range, we concluded that the survey methods and FIA grid density were adequate for sampling to detect the desired effect size. We then used the sampling adequacy of the 183 Lake Tahoe species to make inferences about sampling adequacy for all species in the Sierra Nevada. We determined the proportion of Lake Tahoe species adequately sampled in each of the 8 rarity categories, and we applied a rounded proportion calculated for each of the 4 degrees of rarity (common, moderately rare, highly rare, extremely rare) to all Sierran species to estimate sampling adequacy for various species groups.

Finally, we compared the results of the field test to the results obtained through simulations (Manley et al. 2004). We compared detection probability values from the field data for the 183 Lake Tahoe species to the estimates of Manley et al. (2004) for these same species. We also compared field-based estimates of sampling adequacy to the predictions of Manley et al. (2004) for the Lake Tahoe species and all Sierra Nevada species.

RESULTS

Species Detections

In our field test, we detected 185 species, including 163 (89%) of the 183 Lake Tahoe species. The proportion of species detected varied somewhat by taxonomic group, with all but the larger mammals exceeding 80%: 95 breeding landbird species (93% of those expected), 24 breeding waterbird species (89%), 8 reptile and amphibian species (100%), 20 small mammal species (87%), 8 larger mammal species (62%), and 9 bat species (90%). The 20 species that we expected to detect but missed included 7 landbirds, 3 waterbirds, 4 small mammals, 1 bat, and 5 larger mammals. We did not detect 15 of these species due to low numbers of individuals in the basin, and the remaining 5 species were either extreme habitat specialists, difficult to detect, or difficult to discern from more common congeners.

Of the 22 species detected that were not expected, 15 were species that do not breed in the Lake Tahoe basin, and we excluded them from further analysis in this paper. Notably, however, most of these species were rare or novel sightings for the Lake Tahoe basin, and their occurrence in the basin has important ecological significance beyond the scope of this paper. For example, hoary bat (*Lasiurus cinereus*) and long-legged myotis (*Myotis volans*) were not expected to occur in the basin, but they appeared to be fairly common given that we detected them at multiple sites. Similarly, Great Basin pocket mouse (*Perognathus parvus*) was detected by Sherman trapping in the southern portion of the basin in both years, but the basin was previously outside its known geographic range (Hall 1995). Point counts detected a number of bird species rarely recorded in the basin, such as blue-gray gnatcatcher (*Polioptila caerulea*), common poorwill (*Phalaenoptilus nuttallii*), and northern harrier (*Circus cyaneus*), and these detections helped document and confirm their ongoing occurrence in the basin.

The 7 remaining species detected but not expected were known to occur in the basin, but were thought to be best detected with survey methods we did not use (i.e., nocturnal broadcast surveys, Tomahawk live trapping, terrestrial visual encounter surveys; Manley et al. 2004). We reliably detected birds and mammals with the secondary methods we used, so we included them in our analysis. We did not include lizards in any further analyses because we only detected them with aquatic visual encounter surveys, which are not an appropriate method for them.

We detected nearly 60% (n = 108) of the 183 Lake Tahoe species with secondary survey methods. For 40 species, the secondary survey methods yielded equal or higher p_ds than we expected with the primary survey methods. Most of the crossover was between terrestrial and aquatic point counts, where differences in the survey methods were a combination of count duration, site type, and site location. We detected 28 landbird species more readily in aquatic habitats, and we detected 8 waterbird species more readily in terrestrial habitats. Aquatic sites had 30% more total survey time (40 min) than did terrestrial sites (30 min), so higher p_ds were not surprising.

We detected 10 of the 36 species with only the secondary survey method; 6 landbird species were detected only with aquatic counts, and 4 waterbird species were detected only with terrestrial counts. We also detected 4 mammal species as well or better with secondary detection methods. Ermine (*Mustela erminea*), long-tailed weasel (*Mustela frenata*), and pika (*Ochotona princeps*) all had p_ds that were higher with Sherman trapping (0.05, 0.61, and 0.95, respectively) than observed with their primary survey methods; we made no detections of ermine or long-tailed weasel with trackplates and cameras, and pika had a p_d of 0.53 with terrestrial point counts, just over half that obtained with Sherman trapping. California ground squirrel (*Spermophilus beecheyi)* had a p_d of 0.93 with trackplates and cameras, nearly twice that obtained with Sherman trapping (p_d = 0.58).

We chose not to derive p_ds from multiple survey methods because we wanted our estimates to reflect detection probabilities for individual survey methods. Thus, for all species we used the p_d from the primary survey method unless the species was only detected with a secondary method, in which case we used the p_d from the secondary method. We realize that this approach resulted in lower estimated detection probabilities than necessary for some species.

Estimates of Detectability and Sampling Adequacy

Most species were either readily detected or difficult to detect over the multiple visits specified in each survey method, with an equivalent number of species at each end of the detectability spectrum: 38% ($CI_{90\%}$ = 36% to 48%) of species had a high probability of detection (≥0.70), 40% ($CI_{90\%}$ = 38% to 43%) had a low probability of detection (<0.30), and only 21% ($CI_{90\%}$ = 14% to 21%) had a moderate probability of detection (≥0.30 to 0.69; Table 1). Birds and mammals had approximately twice as many species with high and low probability of detection than species with moderate probability of detection (Fig. 2). Landbirds generally had higher probability of detection than waterbirds, and small mammals generally had higher probability of detection than larger mammals. Bats, amphibians, and reptiles had relatively high probability of detection compared to other species groups, with 89% of the species in these 3 groups having moderate or high probability of detection (Fig. 2).

Table 1. Observed and expected probability of detection (p_d) values for species known to breed in the Lake Tahoe basin and that were detected by 1 of the following survey methods: terrestrial and aquatic point counts, Sherman live trapping, baited trackplate and camera arrays, aquatic visual encounter surveys, and bat mistnetting. Expected values are predictions from Manley et al. (2004) and observed values are empirical estimates from field sampling conducted in the Lake Tahoe basin, California and Nevada, USA, 2002–2004.

	Observed p_d				
Expected p_d	0.0	0.1	0.5	0.8	Total
0.1	6	7	1	2	16
0.5	3	13	15	11	42
0.8	12	33	23	57	125
Total	21	53	39	70	183

We estimated that 66% of the Lake Tahoe species would be adequately sampled if all FIA points on federal lands in the Sierra Nevada were sampled in each of 2 survey periods (Table 2). We also found that, based on the number of FIA points on an average-sized (600,000 ha) National Forest or National Park ($n \approx 250$ points), 54% of the 183 Lake Tahoe species required ≤250 sample points to be adequately sampled. Thus, the MSIM survey protocols could detect population change for these species at the scale of an individual administrative unit.

The proportion of species adequately sampled varied moderately among taxonomic groups (Table 2). We estimated that all 8 of the reptile and amphibian Lake Tahoe species (n = 3 and 5 species, respectively) were adequately detected throughout the Sierra Nevada. Birds and mammals comprised most (96%) of the Lake Tahoe species, and the proportion of species adequately detected varied little among these 2 primary species groups (64% and 65%, respectively). A

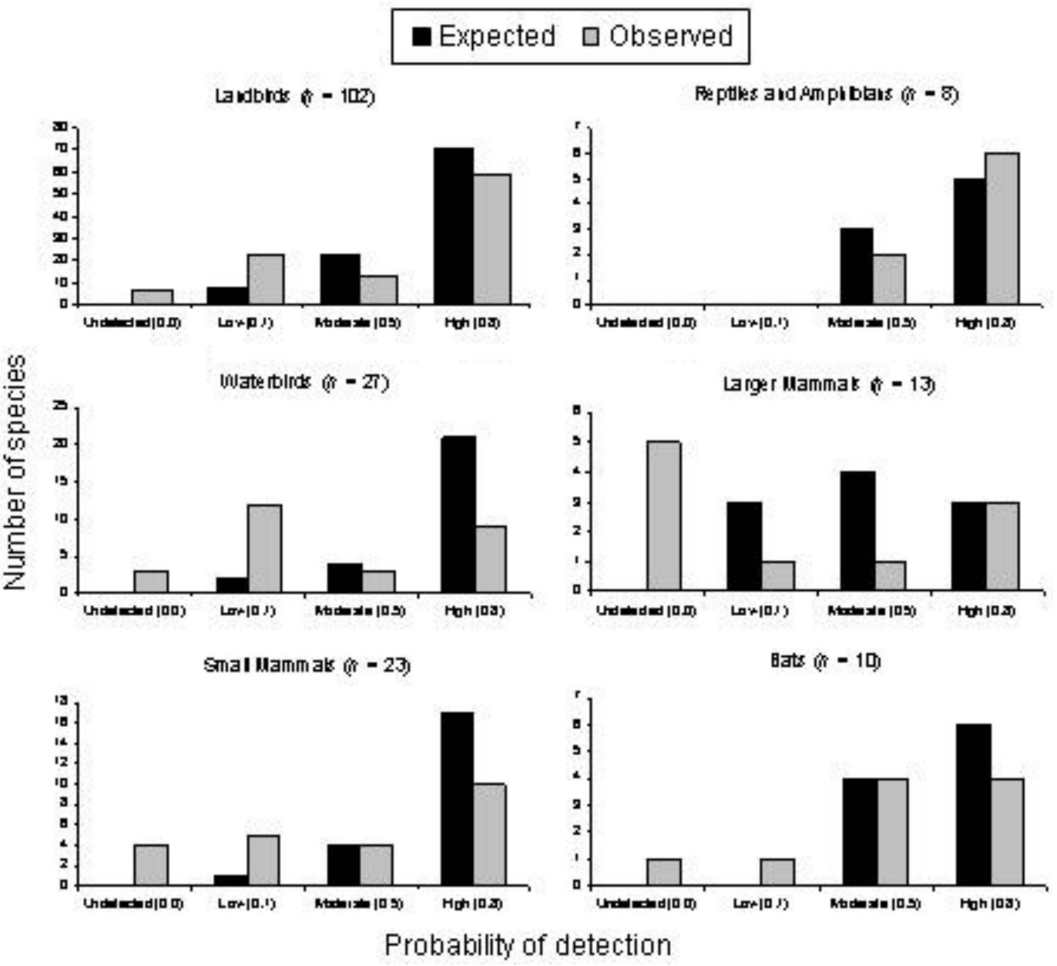

Fig. 2. Observed values of probability of detection based on 6 survey methods conducted in the Lake Tahoe basin, California and Nevada, USA, 2002–2004, and expected values for the same species based on predictions from Manley et al. (2004).

Table 2. The number and proportion of species observed and expected to be adequately sampled for 2 sets of species: species occurring in the Lake Tahoe basin that are readily detected with 6 survey methods, and all species in the Sierra Nevada. Expected values are predictions from Manley et al. (2004). Observed values were derived from detection probabilities generated from field data collected in the Lake Tahoe basin, California and Nevada, USA, 2002–2004. Adequate sampling was based on the ability to detect 20% relative change in proportion of sample points occupied between 2 points in time with 80% confidence and power.

			Proportion adequately sampled			
	Number of species		Lake Tahoe species		Sierra Nevada species	
Species group	Lake Tahoe	Sierra Nevada	Observed	Expected	Observed	Expected
All vertebrates	183	465	0.66	0.96	0.45	0.76
Birds	129	246	0.64	0.97	0.49	0.83
Landbirds	102	195	0.73	0.99	0.55	0.85
Waterbirds	27	51	0.30	0.89	0.31	0.78
Mammals	46	132	0.65	0.93	0.50	0.76
Small mammals	23	77	0.61	0.96	0.44	0.73
Larger mammals	13	38	0.62	0.83	0.52	0.76
Bats	10	17	0.80	1.00	0.76	0.88
Reptiles	3	55	1.00	1.00	0.39	0.65
Amphibians	5	32	1.00	1.00	0.29	0.44
Geographic range[a]						
Small	95	266	0.32	0.89	0.26	0.60
Large	88	199	0.75	0.99	0.75	0.99
Habitat specificity						
Narrow	97	335	0.51	0.93	0.33	0.68
Broad	86	130	0.83	1.00	0.81	0.99
Population size[b]						
Small	32	120	0.31	0.82	0.25	0.55
Large	150	283	0.73	0.99	0.61	0.90
Species of concern	6	38	0.50	0.83	0.32	0.47

[a] Geographic range represented by the number of Forest Inventory and Analysis systematic grid points (see Roesch and Reams 1999) within their range in the Sierra Nevada; small ranges had ≤1,000 FIA points, and large ranges had >1,000 FIA points.
[b] Population size was an estimate of the number of individuals in the Sierra Nevada obtained from U.S. Forest Service (2001); small populations had 1 to 1,000 individuals, and large populations had >1,000 individuals. Population size values were available for all Lake Tahoe species, but they were not available for 56 of the remaining 282 species in the Sierra Nevada, most of which had little overlap of their geographic range with the Sierra Nevada.

greater proportion of landbirds (73%) was estimated to be adequately sampled than waterbirds (30%). Among mammal species, the 3 groups we targeted were similarly represented in the 65% of mammals adequately sampled: 61% of small mammals, 62% of larger mammals, and 80% of bats.

The reduced sampling adequacy for waterbirds appeared to be driven by estimates of p_p, given that the average p_d was only slightly higher for landbirds ($\bar{x}$ = 0.524, SD = 0.403) compared to waterbirds ($\bar{x}$ = 0.304, SD = 0.366), but the average p_p was 3 times higher for landbirds ($\bar{x}$ = 0.427, SD = 0.212) compared to waterbirds ($\bar{x}$ = 0.130, SD = 0.107). Our estimates of p_p were based on the proportion of the ecoregion's landscape occupied by aquatic habitats. If aquatic habitats were targeted for monitoring as they were in our field test, then the proportion of sample points that occured in suitable habitat would be near 100% (i.e., p_p = 1.0) for many species, greatly enhancing the sampling adequacy for water-associated species.

Each of the rarity factors affected sampling adequacy to a similar degree (Table 2). We adequately sampled 75% of the 88 species with large geographic ranges, compared to 32% of 95 species with small ranges. Similarly, we adequately sampled 73% of the 150 species with large population sizes compared to 31% of the 32 species with small population sizes. Habitat specificity had slightly less of an impact on sampling adequacy; we adequately sampled 83% of the 86 species with broad habitat specificity compared to 51% of the 97 species with narrow habitat specificity. Sampling adequacy for species of concern was commensurate with that for other vulnerable species groups, with 50% estimated to be adequately sampled.

The proportion of species adequately sampled generally declined with increasing rarity (Table 3). Common species (i.e., species with no constrained factors: large range, broad habitat specificity, large population size) had the highest proportion of species adequately sampled (84% of 76 species). The proportion of moderately rare species (1 constrained rarity factor) adequately sampled was 72%, ranging from 60 to 80% among the 3 associated categories; species with small ranges were least affected with 80% of the 5 species ade-

Table 3. The proportion of species in the Lake Tahoe basin adequately sampled by 6 survey methods that occur in each of 8 rarity categories[a] (Rabinowitz 1981, Rabinowitz et al. 1986). Adequate sampling was based on the ability to detect ≥20% relative change in occupancy between 2 points in time at Forest Inventory and Analysis-based sample points on federal lands in the Sierra Nevada with 80% confidence and power. Estimates were based on data collected in the Lake Tahoe basin, California and Nevada, USA, 2002–2004.

Geographic range	Large		Small	
Habitat specificity	Broad	Narrow	Broad	Narrow
Large population size	0.84	0.73	0.80	0.28
Small population size	0.60	0.30	n/a	0.22

[a] No Lake Tahoe species had the combination of small geographic range, broad habitat specificity, and small population size.

quately sampled, followed by 73% of the 51 species with narrow habitat specificity, and 60% of the 5 species with small population size. The proportion of highly rare species (2 constrained rarity factors) adequately sampled was 30%, varying only slightly between the 2 associated categories populated by Lake Tahoe species: 30% and 28% of species with large ranges (n = 10 species) and large population sizes (n = 18 species), respectively. No species in Lake Tahoe had the combination of small range, broad habitat specificity, and small population size. We adequately sampled just 22% of the 18 extremely rare species (all 3 factors constrained).

We used the observed sampling adequacy for Lake Tahoe species relative to the 4 degrees of rarity to estimate sampling adequacy for all species in the Sierra Nevada in the following manner: 85% of common species, 70% of moderately rare species, 30% of highly rare species, and 20% of extremely rare species. Based on these relationships, we estimated that 47% (n = 217) of the 465 vertebrate species in the Sierra Nevada would be adequately sampled, with narrow variation among taxonomic groups: 49% of the 246 bird species, 50% of the 132 mammal species, 29% of the 32 amphibian species, and 39% of the 55 reptile species (Table 2).

Sierra Nevada-wide estimates of sampling adequacy in regard to rarity factors were quite similar to those observed for Lake Tahoe species. Based on individual rarity factors, sampling adequacy ranged from 61 to 81% for species with unconstrained conditions (i.e., large range, broad habitat specificity, large population size), and it ranged from 25 to 33% for species with the juxtaposed constrained conditions. Species of concern were again commensurate with rare species, with 32% being adequately sampled.

Observed and Expected Sampling Adequacy

We compared our field-based estimates of detectability and sampling adequacy to the predictions of Manley et al. (2004). Observed p_d values for Lake Tahoe species were lower than expected for most species and species groups (Table 1). Based on Manley et al. (2004), 68% of species were expected to have a p_d of 0.8, as opposed to the observed 38% (CI $_{90\%}$ = 36 to 48%). Given that the percentage of species observed and expected to a have a p_d of 0.5 were quite similar (21% and 23%, respectively), it is clear that many more species were observed to have a p_d of 0.1 than expected (40% [CI $_{90\%}$ = 38 to 43%] and 8%, respectively). Observed p_ds conformed closely to estimates for landbirds, amphibians, reptiles, and bats. The deviation of observed from expected p_ds across all species was driven by shortfalls in detectability in waterbirds, larger mammals, and small mammals. In terms of species-specific predictions by Manley et al. (2004), observed p_ds were equal to that expected for 46% of species, higher for 8% of species, and lower for 46% of species (including the 21 species not detected).

Manley et al. (2004) predicted that 96% of all Lake Tahoe species would be adequately sampled throughout their range in the Sierra Nevada. Our field-based observations were 30% lower, with 66% of the species estimated to be adequately sampled. The observed proportion of species adequately sampled varied among the taxonomic groups and rarity factors we examined (Table 2). Based on the protocols we tested and species breeding in the Lake Tahoe basin, the observed proportion of species adequately sampled was 0.33 lower for birds, 0.28 lower for mammals, and no different for amphibians and reptiles compared to the expected proportion. Greater discrepancies existed for rare species. Based on individual rarity factors, our observed sampling adequacy ranged from 73 to 83% for species with unconstrained conditions compared to the expected sampling adequacy of 99 to 100%. For species with constrained conditions, observed sampling adequacy ranged from 32 to 51% compared to expected values of 82 to 93%. The observed proportion of species of concern, 50%, was 0.33 lower than expected.

A similar degree of difference between observed and expected sampling adequacy was evident for all Sierra Nevada species. Manley et al. (2004) predicted that 76% of all species in the Sierra Nevada would be adequately sampled;

field-based observations were 0.30 lower, with 45% of the species estimated to be adequately sampled. Likewise, observed values were consistently 0.25 to 0.30 lower than predicted by Manley et al. (2004) across all taxonomic groups and rarity factors, with few exceptions. Expected values were closest to observed values for bats (88% vs. 76%, respectively), amphibians (44% vs. 29%, respectively), species with broad habitat specificity (99% vs. 81%, respectively), and species of concern (47% vs. 32%, respectively).

DISCUSSION

Our field test was successful in detecting a greater number of species than expected, and generating estimates of probability of detection that were used to evaluate the ability of the MSIM protocol to adequately sample vertebrates throughout the Sierra Nevada to meet minimum monitoring objectives. Our evaluation yielded a plethora of valuable information about the potential of multiple species protocols to meet land management monitoring objectives. Our results indicated that the dataset generated by the MSIM protocol provides reliable population status and change information for scores of species, but it also reflects the status and change of community composition and structure, along with many other facets of biological diversity. Data yields appeared rich despite the consistently conservative approach we took in evaluating detections: (1) we excluded from all calculations the 22 species that we unexpectedly detected, which if included would have inflated the number of species detected in all species groups; (2) we only used detections from 1 survey method per species to calculate p_d, even though 60% of the species were detected with multiple survey methods; (3) we used a 2-tailed test to make the results more relevant to a variety of potential MSIM applications, and for each species we used the direction of change that required the greater sample size; and (4) we used the lower bound of a 90% confidence interval without the benefit of covariates (MacKenzie et al. 2002) for estimates of probability of detection to derive the minimum number of sample points needed to detect the specified change. These precautions, combined with the limited population sizes in the basin, served to present a worst-case scenario that reflected the minimum one might expect to accomplish by implementing MSIM at the ecoregional scale.

Despite modest sample sizes, the multiple-species survey methods we tested detected 89% of the 183 target species in the study area, including 50% of all species of concern. The small geographic area but high diversity of environments in the Lake Tahoe basin results in many species with small populations, and our modest sample sizes yielded only a few observations of many species. The 20 species not detected generally had small populations (low p_p), but when present should be readily detected (high p_d) by the survey methods we used. Conversely, the MSIM survey methods were able to detect a large number of novel species. We detected 15 species so rarely documented in the Lake Tahoe basin that they were not considered to have established populations. Such novelties may serve as important early harbingers of larger scale changes in environmental conditions, such as those precipitated by land management, natural disturbances, or climate trends.

The nearly 200 species we estimated to be adequately sampled in the Sierra Nevada represented a variety of taxonomic groups and rarity characteristics. A lower percentage of the amphibian and reptile species were estimated to be adequately sampled (29 and 39%, respectively) compared to bird and mammal species (49 and 50%, respectively). However, amphibians and reptiles are likely to have higher detectability and proportion of species adequately sampled with the addition of terrestrial visual encounter surveys, which are the primary survey method for amphibians and reptiles. Terrestrial vertebrate encounter surveys have the added benefit of yielding additional detections of many species infrequently detected with the 6 methods we employed, specifically large mammals and raptors.

As expected, species with a higher degree of rarity suffered a lower sampling adequacy. The rarity factors we evaluated pose different challenges to monitoring. Species with small population sizes may have small or large home ranges. For most species, the primary survey methods cover areas equivalent to or greater than their home ranges, thus these survey methods are detecting occupancy (consistent presence within the survey area and period), as opposed to use (periodic presence within the survey area and period; MacKenzie 2005). Therefore, a greater number of sample sites (e.g., across multiple National Forests or an entire ecoregion) is likely to improve detection rates for most species. The combination of detections from multiple survey methods (i.e., Sherman trapping and trackplate and camera surveys) is also likely to greatly improve species detectability without alterations to either survey method.

In general, larger mammals and waterbirds had small population sizes and large home ranges or use areas that exceeded the area sampled per site. Thus, sample sites are used by them but are not continuously occupied (MacKenzie 2005). For example, waterbirds use many different water bodies for foraging and resting, so they are not always present during surveys at a given sample site. Large mammals have the largest average home-range sizes of any taxonomic group examined, so despite the large area effectively sampled per trackplate and camera array (≈150 ha), ≥50% of all large mammal species in the Lake Tahoe basin and the Sierra Nevada have home ranges over 3 times the size (≥500 ha) of the sample area. Notwithstanding the inconsistent presence of an individual at sample sites within its home range, additional sample effort per site may still increase detection probabilities, but it should be applied with caution. Sample effort per site was high for the trackplate and camera survey method (10 detection devices and 5 visits), so beyond minor adjustments to the survey method, the best option for improving the probability of detection for larger mammals would be the addition of complementary detection methods, such as terrestrial visual encounter surveys. We only conducted waterbird surveys twice per season, so a third visit may be advisable (MacKenzie et al. 2002) unless aquatic and terrestrial point count data are combined. However, increasing the number of sample sites also improves the probability of detection by increasing the number of sites with detections for species with no or few detections, resulting in higher and more precise estimates of probability of detection and proportion of sites occupied.

The primary survey method selected by Manley et al. (2004) for each species appeared to be the most effective in detecting the target species relative to the other survey methods tested. The substantial overlap in species detections among multiple methods, particularly between terrestrial and aquatic point counts, suggests that estimates of probability of detection and proportion of points occupied would be improved by combining detections from multiple methods. It is likely that the p_ds of many species detected with multiple methods in the field test would increase and substantially improve the number and proportion of species adequately sampled.

Estimated sampling adequacy for all species was lower than predicted by Manley et al. (2004), driven primarily by the greater proportion of species with p_ds of <0.3 than expected. Nearly one-third of the species assigned a low p_d by Manley et al. (2004) were estimated to be inadequately sampled. If the field test had been conducted across a larger or more homogenous geographic area, it is reasonable to assume that a smaller proportion of species would have gone undetected, and the estimated proportions of species detected within each range interval would be higher. Thus, implementation of the MSIM protocol at the ecoregional scale is likely to result in sampling adequacy between 50% and 75% of all species (Manley et al. 2004).

Temporal change in biological diversity can be evaluated in myriad ways, including change in richness, diversity, and abundance of all species or species groups based on taxonomy, ecological characteristics, or rarity (Magurran 2004). Although various steps can be taken to improve detectability per survey method (e.g., more visits, more sites), these are not likely to balance representation of common to rare species among those adequately sampled. Probability of detection estimates tend to be low, regardless of true detectability, if species are present at a few sample sites only (MacKenzie et al. 2002). Richness and diversity measures are not affected by per-species sampling adequacy to detect a trend, but they are affected by variation in detectability among species and sites, which can be accounted for in their calculation (Yoccoz et al. 2001). However, population changes among species within particular groups will be affected by sampling adequacy, and differential representation of species based on rarity will need to be considered in reporting change. For example, one way to reflect change in biological diversity would be to report the number or proportion of species in each of several ecological groupings that exhibited an increase, decrease, and no change in proportion of sites occupied between 2 points in time. There is variation in rarity and precision of estimates among species within groups. Rare species may be greater in number or proportion in some groups than others, biasing the interpretation of how species with various ecological characteristics are faring. Simple approaches, such as weighting within-group results by the number of species in the sample area with the same rarity characteristics, may be used to improve the representation of all species in evaluations of biological diversity monitoring data. Similarly, the precision of estimates for species detected within a group will also vary. Hierarchical models for composite change associated with species groups may be used to derive composite mean trend estimates using Markov Chain Monte Carlo methods (Sauer and Link 2002, Sauer et al. 2003).

MANAGEMENT IMPLICATIONS

The data yields indicated by our field test results are impressive relative to the current number and diversity of species currently being targeted by monitoring programs on most NFS lands (GAO 1997). Further, although the proportion of species adequately sampled varied from 30 to 80% among species groups, representation per group was sufficient to make inferences about changes in community structure and biological diversity.

Rare species are of particular concern in land management. Rare species are especially vulnerable to population extirpation (Manne and Pimm 2001, Henle et al. 2004), and their status is a measure of the effectiveness of management solutions for sustaining the large numbers of extant species with small or declining populations. Rare species present a challenge to any monitoring effort, particularly multiple-species approaches that cannot be tailored to the needs of individual species. Nonetheless, adequately sampling 25% ($n \approx 70$ in the Sierra Nevada) of the rarest vertebrate species from a diversity of taxonomic groups would be a significant increase in data yields compared to most monitoring programs on NFS lands. Techniques for borrowing information (MacKenzie 2005) from co-occurring species and environmental conditions across sample sites can be used to improve the precision of estimates for rare species, as well as species that are difficult to detect (MacKenzie et al. 2002).

A fundamental concern about large-scale monitoring programs is their applicability to management needs at smaller administrative scales such as an individual National Forest or National Park. Our results represent a conservative estimate of the substantial yields that could be expected from implementation of the MSIM protocol at an individual administrative unit because the Lake Tahoe basin was 10% the size, had 90% fewer sample sites, and had a subset of survey methods relative to full implementation of MSIM on an average-sized National Forest or National Park. Specifically, we believe the following expectations are reasonable for full implementation of MSIM on a National Forest or National Park: (1) Multiple plant and animal survey methods can be conducted in the same sample season at the same sites without prohibitive conflict among the methods. (2) An individual administrative unit could detect most species targeted by each of the survey methods, resulting in a reliable inventory of most targeted species and distribution data for many of these species. (3) An administrative unit could monitor population change, as we modeled, for approximately 50% of the species targeted by the survey methods. Monitoring data for these and other species at the ecoregional scale provide a context for evaluating the scale of population change (local or regional) and potential causal factors, thereby identifying the appropriate type and scale of response. (4) Local population change data can inform the potential effects of changes on biological integrity as reflected in changes in community composition and differential population change among species with differing habitat associations and environmental sensitivities.

We collected habitat data at all sites, and, although they were not our focus, they clearly meet important regulatory and management information needs. The difficulty in detecting many species adequately enough for modest monitoring programs to identify important population changes continues to drive the need to turn to habitat monitoring as a companion, if not a surrogate, monitoring effort. Presence–absence data can be used to describe and model habitat relationships for many of the species detected (e.g., Morrison et al. 1998, Carroll et al. 1999, Scott et al. 2002, Stauffer 2002, Van Horne 2002), identify or validate indicator species based on observed relationships among species and between species and the environment (Hill 1979, Lebreton et al. 1991, Dufréne and Legendre 1997), and evaluate the efficacy and effectiveness of conservation strategies at a variety of scales (Wilcove 1993, Noss and Cooperrider 1994).

In conclusion, our test indicates that a multi-taxonomic presence–absence monitoring approach such as the MSIM protocol is a reliable and valuable, and, we argue, essential, foundation for large-scale monitoring programs that target plant and animal species, communities, and biological diversity. Further, such multi-species approaches can provide multi-taxonomic population and biodiversity change information at multiple scales, including an individual National Forest or National Park, which is the primary scale of accountability.

ACKNOWLEDGMENTS

This manuscript would not be possible without the programming support and spiritual guidance of our statisticians J. A. Baldwin and S. R. Mori, the contributions of W. J. Zielinski to the preceding paper, the data queries generated by J. T. Gallagher, and the over 30 indefatigable field crew members who collected the data. Finally, thanks to C. D. Vojta and the anonymous reviewers for improving the manuscript.

LITERATURE CITED

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. General Technical Report GTR-PSW-124. U.S. Forest Service, Pacific Southwest Research Station, Berkeley, California, USA.

BISHOP, Y. M., S. E. FIENBERG, AND P. W. HOLLAND. 1975. Discrete multivariate analysis: theory and practice. MIT Press, Cambridge, Massachusetts, USA.

CARROLL, C., W. J. ZIELINSKI, AND R. NOSS. 1999. Using presence–absence data to build and test spatial models for the fisher in the Klamath Region, U.S.A. Conservation Biology 13:1344–1359.

CDFG (CALIFORNIA DEPARTMENT OF FISH AND GAME). 2000. California wildlife habitat relationships program database, version 7.0. California Department of Fish and Game, Sacramento, USA.

CRUMP, M. L., AND N. J. SCOTT JR. 1994. Visual encounter surveys. Pages 84–92 *in* W. R. Heyer, M. A. Donnelly, R. W. McDiarmid, L. C. Hayek, and M. S. Foster, editors. Measuring and monitoring biological diversity: standard methods for amphibians. Smithsonian Institution Press, Washington, D.C., USA.

DAVIES, K. F., C. R. MARGULES, AND J. F. LAWRENCE. 2004. A synergistic effect puts rare, specialized species at greater risk of extinction. Ecology 85:265-271.

DUFRENE, M., AND P. LEGENDRE. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67:345–366.

FELLERS, G. M., AND K. L. FREEL. 1995. A standardized protocol for surveying aquatic amphibians. Technical Report NPS/WRUC/NRTR-95-01. U.S. Department of the Interior, National Park Service, Western Region, San Francisco, California, USA.

FLATHER, C. H., S. J. BRADY, AND M. S. KNOWLES. 1999. Wildlife resource trends in the United States: a technical document supporting the 2000 USDA Forest Service RPA Assessment. General Technical Report GTR-RMRS-33. U.S. Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, USA.

FLEISS, J. L. 1981. Statistical methods for rates and proportions. Second edition. John Wiley and Sons, New York, USA.

GAO (GOVERNMENT ACCOUNTING OFFICE). 1997. Forest Service decision-making. Government Assessment Office GAO/RCED 97-71. U.S. Government Printing Office, Washington, D.C., USA.

GILLESPIE, A. J. R. 1999. Rationale for a national annual forest inventory program. Journal of Forestry 97(12):16–20.

HALL, E. R. 1995. Mammals of Nevada. University of Nevada Press, Reno, USA.

HALL, R. J., AND C. A. LANGTIMM. 2001. The U.S. national Amphibian Research and Monitoring Initiative and the role of protected areas. The George Wright Forum 18:17–25.

HENLE, K, K. F. DAVIES, M. KLEYER, C. MARGULES, AND J. SETTELE. 2004. Predictors of species sensitivity to fragmentation. Biodiversity and Conservation 13:207–251.

HEYER, W. R., M. A. DONNELLY, R. W. MCDIARMID, L. C. HAYEK, AND M. S. FOSTER, editors. 1994. Measuring and monitoring biological diversity: standard methods for amphibians. Smithsonian Institution Press, Washington, D.C., USA.

HILL, M. O. 1979. TWINSPAN: a FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Department of Ecology and Evolutionary Biology, Cornell University, New York, USA.

HOEL, G. H., S. C. PORT, AND C. J. STONE. 1971*a*. Introduction to statistical theory. Houghton Mifflin, Boston, Massachusetts, USA.

———, ———, AND ———. 1971*b*. Introduction to probability theory. Houghton Mifflin Company, Boston, Massachusetts, USA.

JONES, J., W. J. MCSHEA, M. J. CONROY, AND T. H. KUNZ. 1996. Capturing mammals. Pages 115–155 *in* D. E. Wilson, F. R. Cole, J. D. Nichols, R. Rudran, and M. S Foster, editors. Measuring and monitoring biological diversity: standard methods for mammals. Smithsonian Institution Press, Washington, D.C., USA.

KINGSBURY, B., AND J. GIBSON. 2002. Habitat management guidelines for amphibians and reptiles of the Midwest. Partners in amphibian and reptile conservation. Center for Reptile and Amphibian Conservation and Management, Indiana–Purdue University, Fort Wayne, Indiana, USA.

KUNZ, T. H., G. C. RICHARDS, C. R. TIDEMANN, E. D. PIERSON, AND P. A. RACEY. 1996. Observational techniques for bats. Pages 105–114 *in* D. E. Wilson, F. R. Cole, J. D. Nichols, R. Rudran, and M. S. Foster, editors. Measuring and monitoring biological diversity standard methods for mammals. Smithsonian Institution Press, Washington, D.C., USA.

LANDRES, P., D. COLE, AND A. WATSON. 1994. A monitoring strategy for the national wilderness preservation system. Pages 192–197 *in* J. C. Hendee, and V. G. Martin, editors. International wilderness allocation, management, and research. International Wilderness Leadership Foundation, Fort Collins, Colorado, USA.

LAWTON, J. H. 1994. Population dynamic principles. Philosophical Transactions of the Royal Society of London Series B. Biological Sciences 344:61-68.

LEBRETON, J. D., R. SABATIER, G. BANCO, AND A. M. BACOU. 1991. Principal component and correspondence analyses with respect to instrumental variables: an overview of their role in studies of structure–activity and species–environment relationships. Pages 85–114 *in* J. Devillers, and W. Karcher, editors. Applied multivariate analysis in SAR and environmental studies (Eurocourses, Volume 2). Kluwer Academic, Dordrecht, Netherlands.

———, K. P. BURNHAM, J. CLOBERT, D. R. ANDERSON. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62:67–118.

MACKENZIE, D. I. 2005. What are the issues with presence–absence data for wildlife managers? Journal of Wildlife Management 69(3):849–860.

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

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

Magurran, A. E. 2004. Measuring biological diversity. Blackwell Publishers, Malden, Massachusetts, USA.

Manley, P. N., J. A. Fites-Kaufman, M. G. Barbour, M. D. Schlesinger, and D. M. Rizzo. 2000. Biological integrity. Pages 401–598 *in* D. D. Murphy and C. M. Knopp, editors. Lake Tahoe watershed assessment, Volume I. General Technical Report GTR-PSW-175. U.S. Forest Service, Pacific Southwest Research Station, Albany, California, USA.

———, and B. Van Horne. 2005. The Multiple Species Inventory and Monitoring protocol: a population, community, and biodiversity monitoring solution for National Forest System lands. Proceedings of the International Monitoring Science and Technology Symposium. General Technical Report GTR-RMRS-in press. U.S. Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado, USA.

———, W. J. Zielinski, M. D. Schlesinger, and S. R. Mori. 2004. Evaluation of a multiple-species approach to monitoring species and ecosystem conditions at the ecoregional scale. Ecological Applications 14:296–310.

Manne, L. L., and S. L. Pimm. 2001. Beyond the eight forms of rarity: which species are threatened and which will be next? Animal Conservation 4:221–229.

McDonald, M. E., S. Paulsen, R. Blair, J. Dlugosz, S. Hale, S. Hedtke, D. Heggem, L. Jackson, K. B. Jones, B. Levinson, A. Olsen, J. Stoddard, K. Summers, and G. Veith. 2002. Research strategy, environmental monitoring and assessment program. U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1998. Wildlife habitat relationships: concepts and applications, second edition. University of Wisconsin Press, Madison, USA.

Noss, R. F., and A. Y. Cooperrider. 1994. Saving nature's legacy: protecting and restoring biodiversity. Island Press, Covelo, California, USA.

Pawitan, Y. 2001. In all likelihood: statistical modeling and inference using likelihood. Oxford University Press, New York, USA.

Preston, F. 1948. The commonness and rarity of species. Ecology 29:254–283.

———. 1962. The canonical distribution and commonness and rarity: part I. Ecology 43:185–215.

Rabinowitz, D. 1981. Seven forms of rarity. Pages 205–217 *in* H. Synge, editor. The biological aspects of rare plant conservation. John Wiley and Sons, New York, USA.

———, S. Cairns, and T. Dillon. 1986. Seven forms of rarity and their frequency in the flora of the British Isles. Pages 182-204 *in* M. E. Soulé, editor. Conservation biology: the science of scarcity and diversity. Sinauer Associates, Sunderland, Massachussetts, USA.

Ralph, C. J., G. R. Geupel, P. Pyle, T. E. Martin, and D. F. DeSante. 1993. Handbook of field methods for monitoring landbirds. General Technical Report GTR-PSW-144. U.S. Forest Service, Pacific Southwest Research Station, Albany, California, USA.

———, J. R. Sauer, and S. Droege, editors. 1995. Monitoring bird populations by point counts. General Technical Report GTR-PSW-149. U.S. Forest Service, Pacific Southwest Research Station, Albany, California, USA.

Reed, J. M. 1992. A system for ranking conservation priorities for Neotropical migrant birds based on relative susceptibility to extinction. Pages 524–536 *in* J. M. Hagan, and D. W. Johnston, editors. Conservation of Neotropical migrant landbirds. Smithsonian Institution Press, Washington, D.C., USA.

Roesch, F. A., and G. A. Reams. 1999. Analytical alternatives for an annual inventory system. Journal of Forestry 97:33–37.

Southern Appalachian Man and the Biosphere. 1996. The Southern Appalachian Assessment Summary Report. Report 1 of 5. U.S. Forest Service, Southern Region, Atlanta, Georgia, USA.

Sauer, J. R., and W. A. Link. 2002. Hierarchical modeling of population stability and species group attributes using Markov Chain Monte Carlo methods. Ecology 86:1743-1751.

———, ———, and J. D. Nichols. 2003. Estimation of change in populations and communities from monitoring survey data. Pages 227–253 *in* D. E. Busch and J. C. Trexler, editors. Monitoring ecosystems: interdisciplinary approaches for evaluating ecoregional initiatives. Island Press, Washington, D.C., USA.

Schlesinger, M. D., and J. S. Romsos. 2000. Vertebrate species of the Lake Tahoe basin. Pages G-1 to G-15 *in* D. D. Murphy, and C. M. Knopp, editors. Lake Tahoe watershed assessment, Volume II. General Technical Report PSW-GTR-176. U.S. Forest Service, Pacific Southwest Research Station, Albany, California, USA.

Scott, J. M., P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. 2002. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

Simberloff, D. 1986. The proximate causes of extinction. Pages 259-27 *in* D. M. Raup and D. Jablonski, editors. Patterns and processes in the history of life. Springer-Verlag, New York, USA.

Slade, N. A., M. A. Eifler, N. M. Gurenhagen, and A. L. Davelos. 1993. Differential effectiveness of standard and long Sherman live traps in capturing small mammals. Journal of Mammology 74:156–161.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research, Third edition. W. H. Freeman, New York, USA.

Stauffer, D. 2002. Linking populations and habitats: where have we been? Where are we going? Pages 53–62 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

Stephenson, J. R., and G. M. Calcarone. 1999. Southern California mountains and foothills assessment: habitat and species conservation issues. General Technical Report GTR-PSW-172. U.S. Forest Service, Pacific Southwest Research Station, Albany, California, USA.

TERBORGH, J. 1974. Preservation of natural diversity: the problem of extinction prone species. Bioscience 24:715–722.

THOMPSON, W. L., C. GOWAN, AND G. C. WHITE. 1998. Monitoring vertebrate populations. Academic Press, New York, USA.

UNITED NATIONS. 2002. Report on the World Summit on Sustainable Development: 2002 August 26 to September 4; Johannesburg, South Africa. United Nations Conference A/CONF.199/20. United Nations, New York, USA.

U.S. FOREST SERVICE. 2001. Assessment of species vulnerability and prioritization. Sierra Nevada forest plan amendment final environmental impact statement, Volume Four Appendix R. U.S. Forest Service, Pacific Southwest Region, Vallejo, California, USA.

VAN HORNE, B. 2002. Approaches to habitat modeling: the tensions between pattern and process and between specificity and generality. Pages 63–72 *in* J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

WEMMER, C., T., H. KUNZ, G. LUNDIE-JENKINS, AND W. J. MCSHEA. 1996. Pages 156–176 *in* D. E. Wilson, R. F. Cole, J. D. Nichols, R. Rudran, and M. S. Foster, editors. Measuring and monitoring biological diversity: standard methods for mammals. Smithsonian Institution Press, Washington, D.C., USA.

WILCOVE, D. 1993. Getting ahead of the extinction curve. Ecological Applications 3:218–220.

WILSON, D. E., R. F. COLE, J. D. NICHOLS, R. RUDRAN, AND M. S. FOSTER, editors. 1996. Measuring and monitoring biological diversity: standard methods for mammals. Smithsonian Institution Press, Washington, D.C., USA.

WISDOM, M. J., R. S. HOLTHAUSEN, B. C. WALES, C. D. HARGIS, V. A. SAAB, D. C. LEE, W. J. HANN, T. D. RICH, M. M. ROWLAND, W. J. MURPHY, AND M. R. EAMES. 2000. Source habitats for terrestrial vertebrates of focus in the interior Columbia Basin: broad-scale trends and management implications: overview. General Technical Report GTR-PNW-485, Volume One U.S. Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA.

YOCCOZ, N. G., J. D. NICHOLS, AND T. BOULINIER. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16(8):446-453.

ZAR, J. H. 1998. Biostatistical analysis, Fourth edition. Prentice-Hall, Englewood Cliffs, New Jersey, USA.

ZIELINSKI, W. J., AND T. E. KUCERA, EDITORS. 1995. American marten, fisher, lynx, and wolverine: survey methods for their detection. General Technical Report GTR-PSW-157. U.S. Forest Service, Pacific Southwest Research Station, Albany, California, USA.

Associate Editor: Vojta.