Estimating Detection Probabilities for Fishers Using Non-Invasive Methods and Implications for Survey Protocols

29 January, 2009

Draft Final Report

Prepared for: U.S. Fish and Wildlife Service, Yreka Field Office and the Fisher Science Team

Keith M. Slauson[1], Jim Baldwin[2], William J. Zielinski[1], and Michael Schwartz[3].

U.S.D.A. Forest Service, [1] Pacific Southwest Research Station, Redwood Sciences Laboratory, Arcata, California. [2] Pacific Southwest Research Station, Environmental Statistics Unit, Albany, California. [3] Rocky Mountain Research Station, Carnivore Genetics Laboratory, Missoula Montana.

ABSTRACT

Conservation concern for the fisher in the western United States has become heightened in the last decade, culminating in the listing in 2004 as 'warranted but precluded' under the U. S. Endangered Species Act. Occupancy surveys using non-invasive detection devices have, and continue to, serve a major role in the research and management of fishers in the western U.S. However, the standard device-specific survey protocols have not been subjected to quantitative evaluation using the recently developed methods for occupancy estimation. These methods involve the estimation of two parameters (P and ψ) from the collection of detection histories, the sequences of 0's and 1's representing detection or non-detection of a fisher, over the duration of a survey at each sample unit. P refers directly to the ability of the survey protocol to detect a fisher given its presence in the sample unit, which directly refers to the characteristics (e.g., number of stations, number of visits, survey duration) of the survey protocol. ψ refers to the probability that one or more fishers occupies the area surveyed, which is directly related to the ecological characteristics of the locations surveyed (e.g., habitat structure, carnivore community). The focus of this paper is to address issues affecting P, not probability of occupancy, thus we did not attempt to explore habitat or community covariate effects on ψ.

We refer to this enterprise as occupancy estimation and address the application of these methods for two specific survey objectives: (1) assessing the occurrence of fishers at a single location and (2) a non-spatial method for assessing distribution. We used data from standardized occupancy survey protocols using track plates, remote cameras, and hair snares to detect fishers in the western United States to: (1) estimate their respective detection probabilities (2) investigate factors causing detection heterogeneity, and (3) evaluate the effects of spatial or temporal changes to survey protocols. We found three significant sources of detection heterogeneity: (1) within sample unit visit dependency, (2) visit-specific detection probabilities, and (3) survey season. Incorporating covariates to account for these sources of detection heterogeneity significantly increased the accuracy of estimates and inferences from previous analyses. Visit dependency was accounted for by adding a Markov chain persistence factor. The probability of detection varied by visit, generally increasing throughout the survey duration. Fishers had a lower probability of detection during the summer season (July-September) compared to the rest of the year. We developed new protocol recommendations largely on the basis of the analysis of the track plate data but we assume it would apply also to the use of cameras and hair snares. Our primary conclusion is that to achieve the same probability of detection, surveys in the summer (July-September) require longer durations than non-summer seasons (October-June). Based on our analysis and review of the literature, we propose that survey effort (# stations x survey duration) be used to define quantitative thresholds for non-invasive detection device protocols. Distributional surveys or monitoring efforts should use protocols to achieve a ≥0.80 detection probability and surveys to assess the occurrence of fishers at a single-location (e.g., pre-project surveys) use protocols to achieve a ≥0.95 detection probability. Non-summer season protocols need 50 and 60 functional survey days of survey effort to achieve probabilities of detection of at least 0.80 and 0.95, respectively. Summer season protocols need 120 and 200 functional survey days of survey effort to achieve detection probabilities of at least 0.8 and 0.95, respectively. These recommendations assume the use of 6 track plate stations or 2 camera stations per sample unit, with minimal spacing of 500m and 1.6 km between stations, respectively. We recommend that the minimally valid survey effort for evaluating previous survey results be set to achieve a detection probability of at least 0.95, resulting in a minimum of 200 and 60 survey days needed per sample unit during the summer and non-summer seasons, respectively.

INTRODUCTION

Background

One of the fundamental techniques in wildlife research and management is the species-specific **survey** (bold terms defined in glossary section) to determine whether a species of interest occupies a specific area. Inferences about a species' distribution can then be made to address research (e.g., multi-scale habitat relationships: Carroll et al. 1999, Slauson et al. 2007) or management objectives (e.g., population monitoring: Zielinski and Stauffer 1996). Ideally the survey method used to determine **occupancy** (hereafter occupancy survey) for a species should have a probability of detecting the species given occupancy equal to 1, but this is rarely the case. When the probability of detecting a species is <1, a non-detection can result for one of two reasons; either the species was not present in the **sample unit** or the species was present but was not detected during the survey. Thus the inferential problem when dealing with imperfect detection is to properly partition the observed non-detections into true absences (nonpresence, fixed zeros) and false absences (sampling zeros, present but missed; Royle et al. 2008). The consequences of imperfect detection of a species involve biases in the observed species-associated parameter estimates and any modeling of the data represent the surveyor's abilities to detect the species on the landscape, not where the species actually occurs (MacKenzie 2005). Furthermore, the reliability of data derived from imperfect detection methods and the management decisions and ecological inferences drawn from it are questionable.

For species that are difficult to observe, wide-ranging, and of significant conservation concern, occupancy surveys are one of the most cost-effective methods to determine **distribution**, especially over large spatial scales. Forest carnivores, including the marten (*Martes americana*), fisher (*M. pennanti*), lynx (*Lynx canadensis*), and wolverine (*Gulo gulo*), represent such a group of species, whose secretive nature, rarity, and conservation concern are collectively high. The fisher (*Martes pennanti*) is a medium-sized carnivorous mammal in the family Mustelidae. Fishers in western North America are associated with closed-canopy, typically mid to late-successional, coniferous forests (Powell and Zielinski 1994). Conservation concern over the status of fisher populations in the western United States became heightened in the last decade as new information revealed that fishers had become extirpated from Washington, nearly extirpated from Oregon (Aubry and Houston 1992, Aubry and Lewis 1993), and existed in two geographically disjunct populations in California (Zielinski et al. 2005). In 2004 the USDI Fish and Wildlife Service found that the fisher was warranted for protection under the Endangered Species act in the Pacific states but that its listing was precluded by higher priority species (USDI Fish and Wildlife Service 2004).

Existing standard survey methods for determining occupancy of fishers in the Pacific states were developed using a combination of quantitative information (e.g., average number of days until first detection) and professional judgment. A published survey **protocol** (Zielinski and Kucera 1995) provides important guidelines for **detecting** fishers and other forest carnivores using non-invasive detection devices. This protocol identifies $4mi^2$ areas as the sample unit and features two detection devices, track plates (Ray and Zielinski 2008) and remote cameras (Kays and Slauson 2008). There are 6 track plate stations per sample unit with a minimum of 12 days of survey effort or 2 camera stations spaced a minimum of 1.6 km and surveyed for a minimum of 28 days.

Based on available data on the current distribution of fishers in the Pacific states, the results of surveys using this protocol appear to be relatively robust; however, we lack a quantitative assessment of the efficacy of standard survey methods for determining fisher occupancy using a standard sample unit. A large-scale systematic track plate protocol used to determine the distribution of fishers in California (Zielinski et al. 2005) has been evaluated 3 times with subsets of the field data we will be using here. However, these prior efforts used early methods (Zielinski and Mori 2001, Campbell 2004) or simple models (Royle et al. 2008) to estimate **detection probability**. We will employ the most recent analytical tools and incorporate complex models to estimate detection probability and occupancy on the systematic track plate dataset and available remote camera datasets. In addition, survey methods are currently being developed and employed that involve the snaring of hair and collection of scats to obtain DNA samples. These techniques may eventually replace remote-camera or track plate surveys for determining fisher occupancy, because they also provide DNA for use in various research applications (e.g., individual identification, landscape genetics). Consequently, the efficacy of hair-snare and scat collection surveys also needs quantitative evaluation.

Occupancy Estimation

The goal of an occupancy survey is to identify, as accurately as possible given time and resource constraints, if one or more **sites** are occupied by fishers. The result of an occupancy survey is the occupancy state for fishers, either occupied or not-occupied. The most basic interpretation of fisher occupancy is that at least one fisher was present during the survey period. Because fishers are territorial and they maintain year-round home ranges, once presence has been established it can usually be extended to infer the presence of one or more territorial individuals. Some caution must be used here, due to the potential for dispersing individuals to temporarily occupy sites not occupied by territorial individuals; however the greatest concern for this is during the fall and early winter. When one or more fishers are present, additional information about occupancy at the site is possible and includes the number of individuals present, their genders, ages, and potentially reproductive status.

The characteristics, or occupancy states, of the fishers (e.g., age, sex, reproductive status) detected during occupancy surveys are limited by the detection methods currently available: track plate (species presence, gender), remote cameras (species presence), hair snare (species presence, gender, individual), scat detector dogs (species presence, gender, individual, potentially reproductive status). Furthermore, the results produced by an occupancy survey are dependent on the four design elements of a survey protocol: (1) the detection device used determines the occupancy states that can be detected, (2) the survey protocol used with the devices determines the certainty by which the occupancy state can be determined, (3) the spatial location of the surveys determine the habitat context for the occupancy states observed, and (4) the temporal period when the survey occurs determines the temporal context in which the occupancy state was determined. Each of these design elements should be carefully considered when designing or interpreting the results of occupancy surveys. Discussion of survey design is beyond the scope of this report, however a recent treatment of this topic for carnivores can be found in Long and Zielinski (2008).

Distinguishing P and ψ: The survey protocol used to determine occupancy inherently involves a detection probability parameter, $\boldsymbol{P}$, which is the probability of detecting a fisher when it is present *at a sample unit* using the survey protocol. The relationship between P and the protocol depends on the nature of the protocol (e.g., number of **stations** and their spacing in the sample unit, **survey duration**). As an example, suppose that a survey is conducted using a sample unit composed of 6-track plate stations placed in a 4 mi^2 area, where the stations are visited every 2 days (to collect tracks and add bait) for 16 consecutive days, resulting in a total of 8 visits. In this example, each **visit** (v) represents a replicate **sampling occasion**. Thus, the probability of detecting a fisher using this survey protocol, is $P = 1-(1-p)^V$, where ***p*** is the *per visit* probability of detecting a fisher given a fisher is present. Clearly, P increases as both p and v increase, allowing for the evaluation of alternative existing and potential survey protocols. p is influenced by attributes of the sampling unit, the number of replicate stations present, their spatial distribution, the amount of time (e.g., number of days) elapsed per visit, and the attractants present (e.g., bait and olfactory lure types).

The other parameter of interest is ψ, the **probability of site occupancy**, also interpreted as the **proportion of sites occupied**. This is often the parameter of interest for occupancy surveys as it can be used to investigate habitat covariates that influence occupancy or be used to determine the number of sample units occupied for comparison over time for **monitoring** purposes. However, when $P <1$, the true occupancy state of non-detection sites is confounded with P and thus the true occupancy state is the result of the product $\psi * P$ for each site. To obtain information to estimate P, replicate observations for each sample unit are compiled into a **detection history** (MacKenzie et al. 2002). Importantly, this approach uses data only at sites where the target species has been detected at least once. Using the 8-visit example above, a potential detection history could be 00010111 where each 0 and 1 represents a visit-specific result. Thus, a sample unit surveyed $v = 1,2,\ldots V$ visits, yields the detection history for sample unit i: $y_i = (y_{i1}, y_{i2},\ldots y_{iV})$. To estimate p and ψ the site-specific detection histories can be summarized by the following: $y_{i.} = \sum_{v=1}^{V} y_{iv}$ where the probability distributions of the total number of detections, y_i , can be specified as a function of the parameters p and ψ (Royle et al. 2008). Then, estimates of the parameters p and ψ can be obtained by numerically maximizing the likelihood function.

While P and ψ are closely linked in the computation of occupancy estimation, it is important to understand that they represent distinct phenomena. P refers directly to the ability of the survey protocol to detect a fisher given its presence in the sample unit, which directly refers to the characteristics (e.g., number of stations, number of visits, survey duration) of the survey protocol. ψ refers to the probability that one or more fishers occupies the area surveyed, which is directly related to the ecological characteristics of the locations surveyed (e.g., habitat structure, carnivore community). The focus of this paper is to address issues affecting P, not probability of occupancy, thus we did not attempt to explore habitat or community covariate effects on ψ.

<u>Detection Mechanics and Occupancy Estimation Assumptions</u>: Detection of fishers using detection **devices** involves a sequence of probabilistic events: (1) an individual that occupies an area must find at least one device, (2) then the individual must approach and enter/trigger the device, and (3) the device records the presence of the fisher (via quality tracks, photographs, or a hair sample). Each event is directly linked to fisher ecology, behavior, and survey design in the following ways: Event 1 relates directly to survey design and fisher ecology and behavior (e.g., fisher movements and sensory abilities to detect and find stations), Event 2 relates directly to fisher behavior (e.g., their willingness to enter enclosed spaces) and survey device design, and Event 3 relates to detection device efficacy. By using detection data only from sample units where fishers have been

detected, as in the occupancy estimation approach (Mackenzie et al. 2005), the probabilities of event 1-3 are all assumed to be 1 for all individuals and cohorts (e.g., males and females, juveniles and adults). It is important to keep in mind the uncertainties in the magnitude of these three probabilistic events. If any are significantly <1, estimates of probabilities of detection may be inflated. However, concern for this bias has remained low because average **latency to first detection** (**LFD**; the number of days until the first detection of a fisher at a sample unit) is typically <50% of the total survey duration (Zielinski and Mori 2001). This suggests that, on average, animals are able to find, and are willing to enter or approach, detection devices early to midway through the survey duration.

Occupancy estimation includes four assumptions (Mackenzie et al. 2005): (1) Occupancy status at each site over the survey season is constant; that is sites are 'closed' to changes in occupancy, (2) Probability of occupancy is constant across all sites, or differences in occupancy probability are modeled using covariates, (3) Probability of detection is constant across all sites and surveys or is a function of site-survey covariates; there is no unmodeled heterogeneity in detection probabilities, and (4) Detection of species and detection histories at each location are independent. If these assumptions are not met, estimators may be biased and inferences about factors that influence either occupancy or detection may be incorrect. Assumption 2 pertains only to estimating ψ, not P, and thus will not be discussed further.

For the majority of the year, and for the majority of sites, the assumption of closure is valid. The exceptions are for males that leave their home ranges during the mating season in search of females and for dispersing juveniles that occupy areas previously unoccupied in the same year. Both of these phenomena are time-specific, with males searching for females in the spring and juveniles dispersing in the late summer and fall. In addition, mortality can also result in changes in occupancy if it occurs significantly during the survey period. The potential for these cases to cause major violations is likely small, especially if the total number of sites sampled is large (e.g., >50) or surveys occur outside these periods.

The third assumption, that probability of detection is either constant or that variation is a product of site and/or survey covariates, is the focus of this paper. Unmodeled heterogeneity in detection probabilities, hereafter **detection heterogeneity**, will often result in negatively biased occupancy estimates. Negatively biased occupancy estimates will underestimate true occupancy rates. Detection probability may also vary among surveys as a result of factors such as season, environmental (e.g., weather conditions-temperature, precipitation), and social behavior patterns.

Detection probabilities may also vary among sites because of habitat features, and possibly due to the size of the local population at each site. This latter factor is not likely an issue due to the size of standard sample units relative to the size of fisher home ranges and intrasexual territoriality. We will explicitly evaluate whether detection probability is constant and test multiple covariates hypothesized to cause detection heterogeneity.

Although the occupancy status of home ranges is not likely to change over the survey season, fishers may not be equally available for detection in a sample unit over the survey duration. Fishers have large home ranges and may not be within detection range of a sample unit. If substantial regional variation exists in home range sizes, detection probabilities could be lower in areas with larger than average home ranges because the density of sample units relative to the density of fishers is low. The relationship between home range size and the portion of home range occupied by a sample unit also varies by sex. Because male home ranges are larger, sample units represent a higher proportion of female versus male home ranges. The effects of these two factors may be revealed by investigating sex-specific detection probabilities and variation in detection probabilities with factors that are associated with home range size (e.g., habitat suitability, geographic location, population cores versus peripheries). We will investigate several factors hypothesized to be associated with variation in home range size, but investigating sex-specific detection probability is beyond the scope of our analysis.

The final assumption, that species and detection histories at each location are independent, is an assumption to be concerned about. If detection is not independent among sites, the precision of the occupancy estimates is usually overstated. Non-independence can arise when an individual fisher can visit >1 site per occasion. As a general rule, if the spacing between sample units is less that the diameter of a typical home range for a male, the assumption is potentially violated. In these instances, the 'effective sample size' (the number of independent sites or detection histories) is actually smaller than the number of sites surveyed, and the estimated standard errors obtained from the above model are too small (Mackenzie and Bailey 2004). While this problem should be addressed in the design phase, if this problem is suspected, standard errors can be adjusted with an estimated **variance inflation factor** (Mackenzie et al. 2005).

Lack of independence in detection histories, when a visit to a station either increases or decreases the chance of a revisit, may also create biases in the parameter estimates. We suspected that the assumption of independence of detection histories (among visits) might be violated for several reasons. We hypothesized that once individual fishers visited one station in a sample unit, they would be more likely to revisit that or other stations in the same sample unit due to a positive experience of receiving a bait reward. Second, in most sample units where fishers are detected, the same individual will be detected on multiple occasions. Thus the response from individual fishers detected at each sample unit will likely dominate the detection data collected from each sample unit. We will address the assumptions of spatial and temporal independence for each dataset by considering their design relative to fisher spacing and explicitly test for independence within detection histories by including survey-specific covariates to account for dependency.

Survey Objectives & Detection Confidence

The desired accuracy level of detection, or the confidence in assigning detection certainty, depends on the objectives of the survey or research effort. Long and Zielinski (2008) identified the two survey objectives most often involved when occupancy surveys are conducted for carnivores: (1) assessing the occurrence of at least one individual at a single location (hereafter single-location surveys) and (2) a non-spatial method for assessing distribution (hereafter distributional surveys). Surveys focused on one or a few sites usually demand a high level of confidence that a fisher is detected if present. This is because the typical motivation for conducting these single-location surveys is to determine if one or more fishers occupy an area proposed for a land management activity that may affect the habitat. Conversely, distributional surveys seek to estimate the proportion of locations occupied by fishers over large geographic areas. For these 'large-scale' surveys, 2 goals must be optimized: achieving the highest confidence of detection at each site and surveying as many sites as possible. This is because the estimation of ψ typically becomes more precise by sampling more locations rather than increasing the per location probability of detection (MacKenzie and Royle 2005).

To incorporate both of these survey objectives, we have included multiple levels of precision of P (e.g., 0.80 to 0.95) in order to provide guidance for survey protocols for the detection of fishers. We assume that surveyors with the goal of maximizing detection probability at one or a few sites will desire guidance as to the survey effort necessary to achieve a $P \geq 0.95$. Conversely, we assume that surveyors with the goal of determining the distribution of fishers over a large area will desire guidance as to the survey effort necessary to achieve a $P \geq 0.80$, which represents the optimization of effort to have high detection probabilities at single sites and spatial replication of sample units.

The objectives for this report include:

1. Review and synthesize the relevant literature on detection probability for fishers and closely related species (e.g., *Martes americana*).

2. Develop models that incorporate covariates that account for detection heterogeneity to estimate detection probabilities for fishers and to compare these models using statistical techniques. Covariates include a wide variety of temporal and spatial factors:

   Month: Detection probability is not equal among all months. Multiple factors that change during the year, including environmental (e.g., weather extremes), social (e.g., reduced use of home ranges by females during kit rearing), and population changes (e.g., addition of the young of the year fishers into the population) and have the potential to cause variation in the detection probability of fishers.

   Population: Detection probability differs in the Klamath and southern Sierra fisher populations in California.

   Distance to Coast: Detection probability is a decreasing function of distance from coast due to the hypothesized higher densities of fishers near the coast versus the interior. This effect is hypothesized to be relevant only for the Klamath fisher population.

Elevation: Detection probability declines at low and high elevations, due to lower densities of fishers at the elevational margins of their distribution.

Population Core versus Periphery: Detection probability will be lower along the periphery of populations, due to the hypothesized lower density of fishers, compared to population centers.

Habitat Suitability: Detection probability will be higher at sites with high habitat suitability, assuming fisher density is an increasing function of habitat suitability.

Visit Dependency: Detection histories for each sample unit are not independent. We hypothesize that once fishers find stations they will be detected more frequently in subsequent visits, creating clumped patterns of 0's and 1's in the detection histories. The consequence of this will be increased detection heterogeneity and biased estimates of p if left unaccounted for.

Device Inoperability: Increasing levels of device inoperability will decrease p. Inoperability is defined as human, animal or environmental factors that render detection devices unable to record the presence of a fisher. Typically, the major contributor to inoperabilities is American black bear (*Ursus americanus*) damage to detection devices.

3. Determine how changes in spatial and/or temporal survey effort affect the probability of detection for each method-specific protocol at confidence levels suitable to meet single-location ($P > 0.95$) and distributional ($P > 0.80$) survey objectives.

4. Determine how the findings from objectives 1 and 2 affect estimation of ψ.

5. Use the results of these analyses to identify minimum survey effort needed to detect fishers over a range of probabilities. We will evaluate current protocols and make recommendations on their future use for conducting occupancy surveys for fishers.

METHODS

Literature Review

We began this work by thoroughly reviewing literature that either experimentally estimated detection probability or that addressed issues influencing the detection of fishers, or the closely related American marten (*Martes americana*). The review focused specifically on summarizing, comparing, and contrasting experimental versus post-hoc approaches to addressing detection probability. Secondarily, we synthesized all significant findings relevant to issues affecting the detection of fishers and related taxa.

This review and synthesis provides the foundation for our work. It also allowed us to understand the foundations that had been established in regard to our 2 primary objectives: estimating detection probability and the factors that affect it and understanding how variation in effort affects detection probability.

Estimating Detection Probability and Evaluating the Factors that Affect It

We first describe the datasets included in the analysis for each detection device. Then we describe the summary statistics that will be used to describe each dataset. We then identify the covariates that may cause detection heterogeneity and how they will be evaluated. Those variables that meet the described levels of statistical significance will be included in the model development stage to develop candidate models to estimate p and ψ.

Track Plate: Track plate data came from the large-scale systematic survey effort within the historical range of the fisher in California and southwestern Oregon (hereafter systematic track-plate survey; Zielinski et al. 2005). In the area surveyed, all elevations and forested habitat types were sampled. These surveys were conducted on a systematic grid with 7 - 10 km between adjacent sample units, a distance that likely satisfies assumption 4, independence of species detection and detection histories between sample units. At each point on the grid a sample unit was established, comprised of 6 track plate stations in a pentagonal array with one station in the center and 5 around the perimeter with 500-600 m between adjacent stations (Figure 1). Once established, each sample unit was run for 16 consecutive days. Each station was baited with a single drumstick-sized piece of chicken and a commercial trapping lure (Gusto, Minnesota Trapline Products, Pennock, Minnesota, USA) was applied to a nearby tree when each station was established and re-applied after 8 days if a marten or fisher had not yet been detected. Each station was visited every 2 days to replace bait and remove tracks.

The systematic track-plate survey include 530 sample units and was largely conducted by the Pacific Southwest Research Station from 1996-2002, with the most recent efforts conducted by the U.S. Fish & Wildlife Service from 2005-2006 (S. Yeager, unpublished data). Eighty-four (16%) sample units had fisher detections, 56 in northwestern California region and 28 in the southern Sierra Nevada region. Fishers were detected at a total of 220 stations, averaging 2.62 stations per sample unit.

Remote Camera: Remote camera data consisted of 2 surveys located in the interior Klamath region of northwestern California. The first survey effort (hereafter Collins-Baldy) was conducted in an approximately 124 km^2 ($77 mi^2$) study area composed of a checkerboard ownership of Forest Service and private timberlands (Farber and Franklin 2005). The second survey effort (hereafter, Mt. Ashland) was conducted in a 210 km^2 study area which also comprised a checkerboard ownership of USFS and private timberlands (Farber and Criss 2006). Each study area was completely saturated with 4 mi^2 sample units, potentially violating assumption 4, independence between species detection and detection histories between sample units. This violation is less of a concern because the objective of these surveys were to determine the distribution of fishers throughout each study area (Farber and Franklin 2005, Farber and Criss 2006), not to survey with the intent to employ occupancy estimation methods. Each sample unit was surveyed using the Zielinski and Kucera (1995) remote camera protocol, where two remote camera stations are placed approximately 1.6 km apart (Figure 1), surveyed for 28 consecutive days, and checked once per week. The Collins-Baldy survey modified this protocol by removing a sample unit once a fisher was detected resulting in truncated detection histories. Each station consisted of a single passive-infrared film camera (Stealth Cam Model MC2-GRT or Moultrie Game Spy II camera) mounted on a single tree facing a second tree with the bait attached to it. Each station was baited with three drumstick-sized pieces of chicken and two perforated cans of cat food placed in a chicken-wire cage nailed to either the base of a tree or 2.5 m up a tree. No olfactory lures or visual attractants were used. In total there were 42 sample units, 21 in the Collins-Baldy study area and 21 in the Mount Ashland study area. Fishers were detected at a total of 20 (49%) sample units, 15 (71%) in the Collins-Baldy and 5 (24%) in the Mount Ashland study areas.

We also evaluated remote camera surveys conducted in the southern Cascades of Oregon for inclusion in the analysis. These surveys generally followed the same Zielinski and Kucera (1995) protocol, including two remote cameras per 4 mi^2. However, bait stations initially without cameras were established in a number of locations and a camera added only if potential fisher tracks were found at the site. Sample units with fisher detections included a mixture of four different survey protocols, non-systematic design, variable visit intervals, variable total survey durations, and placement largely coincident with areas where 'fisher' snow tracks were sighted. Due to the amount of variability in the methods relative to the number of sample units with fisher detections (11 total sample units), these data were not included in this analysis. This underscores the importance of consistently following a well-designed protocol in order to produce results that can be analyzed and contribute to understanding beyond individual survey areas.

Hair Snare: Hair snare data consisted of one survey in the Locsha Corridor region of northern Idaho. This survey was conducted in a 300 km^2 study area composed of U. S. Forest Service lands (M. Schwartz unpubl. data). Individual stations were established along a highway corridor at 3.2 km intervals, and at least 150 m from the road (Figure 1). This spacing also potentially violates assumption 4, independence between stations. While this is a violation of an assumption of occupancy estimation, the original objective of this survey was to determine the distribution of fishers throughout the study area, not to survey with the intent to employ occupancy estimation methods. Once established, each station was run for a total of 22 consecutive days and checked every other day for 11 visits. Each station consisted of a 81.3 cm (32") triangular coroplast enclosure open at both ends (Schwartz et al. 2006). Three 30-caliber brass gun brushes, one at the center of each side, were placed at each end of the enclosure. Each station was baited with a small (3"x1"x1") piece of beaver (*Castor canadensis*) meat wired into the center of each hair snare. Scent lure consisted of a combination of rotten fish and beaver castor infused into a sponge and suspended from a tree branch.

Genomic DNA was extracted from hair samples using *Qiagen DNeasy Tissue Kit* according to manufacturer protocols (Qiagen Inc., Valencia, CA). Species identification was determined using restriction digest of a 442 bp segment of the *cytochrome b* region diagnostic for fisher (Riddle et al. 2003). All samples producing fisher identification using restriction digest were subsequently sequenced. Both strands were sequenced and analyzed on a Li-Cor 4300 DNA imager using standard protocols. Sequences were compared to reference samples collected by the lab or to sequences located in the National Institutes of Health (NIH) Genbank.

There were a total of 23 hair snare stations in the Lochsa study area and fishers were detected at 10 (43%) of these sample units. A mean of 1.74 (SD = 1.7; range 0-6) hair samples were collected per sample unit, of those 19 (47.5%) were confirmed to be fisher using genetic analysis.

Statistical Analysis

Summary Statistics: Diagnostic summary statistics (mean, SE, range, 95% confidence intervals) were generated for the following variables: latency to first detection (LFD), the number of sample unit visits with a fisher detection and the total number of station visits with fisher detections. For the track plate data set, the entire suite of summary statistics were also calculated by month and season. The average number of inoperable intervals (track plate and hair snare) or days (remote camera) when devices were rendered unable to detect a fisher (e.g., due to bear disturbance or equipment failure) and the percentage of total survey days per protocol they represented were summarized for each device type.

Univariate Tests: Investigating Detection Heterogeneity: The systematic track plate surveys were conducted over a substantial geographic area (>30,000 km^2), and were conducted over different months and years. We first investigated the effect of month on: (1) LFD, (2) total visits to sample units, and (3) total number of station visits using summary statistics and t-tests. We used these results to help determine whether sufficient difference occurred in the data to identify season-specific estimates of p. Differences between detection probabilities were estimated separately for each fisher population (Klamath and southern Sierra) and compared using t-tests. For the season and population comparisons, we used a simple model (constant p: assumes p does not vary among visits) to estimate the single visit detection probabilities for the 2 groups of interest, season and population, and compared single visit detection probability estimates using t-tests. We used correlation analysis to evaluate the effect of distance to coast and elevation on the number of visits with fisher detections at sample units.

Understanding the effect of core versus periphery required first that we define these locations. We did so in two ways, by conducting a neighborhood analysis to define local cores and peripheries and regionally by designating periphery as those sample units along the edge of the two contiguous populations in California. The neighborhood analysis was conducted to determine the total number of sample units within 8.5 and 17.0 km radii circles (i.e., 'neighborhoods') that detected fishers. These distances captured the most proximate neighborhood (within 4 sample units) and a regional neighborhood (within 12 sample units), respectively. For each sample unit where fishers were detected, we examined the correlation of the total number of fisher visits to the sample unit (1-8) with the total number of sample units in the neighborhood that detected fishers. The regional neighborhood analysis used LFD and the number of sample unit visits that detected fishers to compare sample units on the geographical edges of known fisher distribution and those in the cores using t-tests. These two approaches allowed us to investigate the relationship of adjacent sample units with detections and the detection characteristics of sample units on the core versus periphery of the distribution of fishers.

We investigated the influence of habitat suitability by correlating the number of fisher visits at a sample unit and the mean habitat suitability value in a 1.0 km radius around sample units where fishers were detected. To derive habitat suitability values, we used the predictive models in Davis et al. (2008) for each region (Klamath and southern Sierra) occupied by fishers in California.

To limit the number of potential candidate multivariate models for the detection probability analysis (described below), we used the results of the univariate comparisons to exclude variables that did not affect estimation of *p* or diagnostics related to *p* (e.g., LFD, total sample unit visits). All variables showing moderate to significant effects (e.g., $\alpha \leq 0.10$, $r > 0.6$) were included for developing competing models to fit to the detection history data.

<u>Visit Dependency:</u> One of the assumptions of occupancy modeling is that detection histories are independent. Lack of independence, due to positive dependency in visit results, causes biased estimates and usually results in underestimates of variability because the true number of observations (8 visits) is less than what is observed. If dependence exists, it means that the status (0 or 1) for any visit depends on the status of the previous visit. To evaluate whether dependence was an issue in the detection histories, we constructed a 2x2 frequency table to compare sequential visit outcomes (e.g., 00, 01, 10, 11) for each visit transition (from visit 1 to visit 2 and visit 2 to visit 3 and visit 3 to visit 4, etc.) observed in the field data with the expected counts. For comparison, we used a single candidate model (2-group, visit-specific *p*), and this same model with a Markov-chain persistence factor added to account for any dependence in the detection histories. A Markov-chain persistence factor is one mathematical model to account for dependency. The observed and expected frequencies of transitional values were compared between the model with and without the Markov-chain persistence factor.

<u>Model Development and Selection</u>: The first step in estimating detection probability and occupancy is to fit a model to the data. For each detection device data set we developed a candidate set of models using both standard models included in program PRESENCE (version 2.0, MacKenzie and Hinze 2006) and custom models representing hypotheses on how covariates may influence detection probability. Program PRESENCE uses detection histories to estimate patch occupancy rates and related parameters (e.g., *P*) for single to multiple year surveys and can be downloaded for free at http://www.mbr-pwrc.usgs.gov/software/presence.html. Standard models included covariates that assume *p* is constant across visits (constant *p*), that estimate *p* individually for each visit (visit-

specific p), and assume that heterogeneity exists in the data such that the data are better modeled if partitioned into two groups (2-groups). Custom models included only variables with moderate to significant effects in the univariate analysis (see above).

We evaluated the relative performance of all models to describe the detection history data by comparing their AIC_c values and AIC_c weights (w) (Burnam and Anderson 2002). We used the top model(s) to estimate the probability of detection (p) and occupancy (ψ).

Effects of Spatial and Temporal Changes in Survey Protocols on Detection Probability

We used the estimates of p from each of the top model(s) to calculate how many visits (track plates and hair snares) or survey days (cameras) needed to be added or subtracted to achieve a $P \geq 0.80, 0.85, 0.90, 0.95$ for an individual sample unit. If independence of status (detection/non-detection) between visits exists ($\phi = 0$ in the Markov chain model), p can be used in the following equation to estimate the single-station p (sp) for any visit (v):

$$sp_v = 1 - (1 - p_v^{(1)})^{1/S}$$

where S is the number of stations in the sample unit used to estimate p. This value for sp_v can then be used to determine P for any given combination of stations (s) and visits using the following equation:

$$P_{vs} = 1 - \prod_{i=1}^{v} ((1 - sp_i)^{(s)})$$

This model also assumes independence among stations and among visits. However, if there is positive dependence of status between visits ($\phi \geq 0$), a new equation is needed in order to account for this dependence. For example, the probability of at least one detection at a sample unit with two visits and one station (P_{21}) is given by the following equation:

$$\begin{aligned} P_{21} &= 1 - (1 - p_1)\prod_{i=1}^{v-1} ((1 - p_i)(1 - p_{i+1}) + \phi(\min(p_i, p_{i+1}) - p_i p_{i+1}))/(1 - p_i) \\ &= 1 - (1 - p_1)(1 - p_2) - \phi \min(p_1, p_2)(1 - \max(p_1, p_2)) \end{aligned}$$

In this equation, P_{21} decreases with increasing positive persistence ($\phi \geq 0$). Additional details on the derivations of these models can be found in Appendix 2.

Station Inoperability: For the track plate and camera data we also calculated the interactive effects on P of changing the number of total stations and total survey duration, while increasing the proportion of visits with station inoperabilities. To calculate these interactive effects it was necessary to derive new equations to calculate their effects on P.

The probability of not detecting any fishers irrespective of inoperability in a *d*-day period is $(1-p)^d$ (assuming day-to-day independence of inoperability). The probability of having the opportunity to detect at least one fisher but having those results negated because of inoperability on the first day is $(1-(1-p)^d)\omega$ with ω being the probability of inoperability on any given day. The probability of the first detection of a fisher occurring on day 2 and having inoperability on day 2 is $(1-p)(1-(1-p)^{d-1})(1-(1-\omega)^2)$. In general the probability at a station missing a fisher detection due to inoperability that occurs on day *k* is given by:

$$(1-p)^{k-1}(1-(1-p)^{d-k})(1-(1-\omega)^k).$$

The sum of the probabilities of no detections is

$$\begin{aligned} Q &= (1-p)^d + \sum_{k=1}^{d}(1-p)^{k-1}(1-(1-p)^{d-k})(1-(1-\omega)^k) \\ &= \frac{\omega - p(1-\omega)((1-p)(1-\omega))^d}{\omega - p(1-\omega)} \end{aligned}$$

For *t* *d*-day time periods the probability that a station at a sampling unit with presence has no detections is Q^t. The probability that all *s* stations at a sampling unit with presence has no detections is Q^{st} (assuming that stations are independent given that the sampling unit has presence); and the probability of at least one detection at the sampling unit with presence is given by $1-Q^{st}$. This equation, $1-Q^{st}$, was used to determine the effect of increasing amounts of inoperability on track plate and camera sample units with different combinations of the number of stations (track plate only) and survey duration (track plate and remote camera).

RESULTS

Literature Review

We found and reviewed 10 studies that addressed detection probability for fishers and American martens. Three were experimental and seven were post-hoc analyses of detection device data.

Experimental Studies: Ivan (2000) assessed the effectiveness of enclosed track plates for detecting American martens in Montana by monitoring individuals with radio collars during periods when track plates were present in their home ranges. A total of 15 martens were radio collared and toe branded, 1-2 in each of 10 different 10.44 km^2 sample units during the summer (June-August). Six track plates were deployed in two rows of three, separated by 0.3-0.8 km, and run for 12 consecutive days, following the Zielinski and Kucera (1995) protocol. Each station was baited with chicken, sardines, and an unknown olfactory lure. Track plates were established in each sample unit from 1-12 days after martens were captured. All radio collared martens were present and potentially detectable in their respective sample units during the survey period. Only 7 of 10 sample units detected martens, even though they were present in the vicinity of all 10. Only 1-2 of the 15 martens radio collared martens were detected on track plates. Eight of the 15 martens that were not detected were also monitored with remote telemetry systems; data from the remote systems indicated that 2 martens approached track plates but did not enter them and the other 6 were never detected but were in the vicinity of track plates. Probability of detecting martens, given presence, derived empirically (P = 0.7, 95% CI = 0.42 – 0.98) was lower than that derived from LFD (0.977). A minimum of 2-7 martens were present in each sample unit, based on trapping during the summer.

Fowler and Golightly (1994) estimated detection probability using track plates and line-triggered remote cameras placed in the home ranges of five radio collared and ear tagged martens in California. In the fall (October-November) martens were live trapped, fitted with radio collars and ear tagged. Radio collared martens were relocated daily via remote triangulation and these locations were used to estimate home range location and size. During the winter (December-January), >1 month after live capture, 20 detection stations, each station composed of a track plate and remote camera 20 m apart, were established at 1-km intervals across the area occupied by all 5 home ranges. Each detection device was baited with chicken and visited every 2 days for a total of 22 days. The number of stations present in each home range varied from 4-11. All 5 martens were detected at >1 station, yielding a 100% detection probability for individuals. The number of individual stations that each individual marten was detected at varied from 2-9. The proportion of stations available that detected each individual ranged from 0.4 to 1.0. 95% of the stations present in ≥1 marten home range detected at least one marked individual marten.

Smith et al. (2007) investigated the effect of the density of martens present in sample units on the estimation of detection probabilities in South Dakota. At different times during the winter through summer seasons, 8 locations were sampled using the 10.2 km^2 sample unit and 6-station track plate protocol recommended by Zielinski and Kucera (1995). Each sample unit was surveyed for 12 consecutive days and visited every 4 days; attractants consisted of chicken as bait and an olfactory lure placed on a nearby tree. To determine the density of martens at each of the 8 sample units surveyed, 20 live traps were baited similarly to track plates and opened for 10 days in each sample unit. Trapping was conducted 4 days to 7 months after track plate surveys concluded; 0-7 martens were captured at each sample unit. The top ranked model allowed p to vary as a function of density and produced estimates of P = 1.0 (SE = 0.001) for high density (>2 martens per sample unit) sample units and P = 0.70 (SE = 0.15) for low density (≤1 martens per sample unit) locations using the survey protocol.

<u>Post-Hoc Studies:</u> Zielinski and Mori (2001) used a subset (67%) of the systematic track -plate survey data we have used here (Table 1). A total of 357 sample units, 65 (18%)

that detected fishers, collected from 1996-2000 were used to estimate probability of detection. They used a simple model, assuming p is a constant across all visits, to estimate a 1-visit $p = 0.40$. Using this 1-visit estimate of p, a 6-station sample unit using 8-visits was estimated to have a $P = 0.98$. LFD was 6.85 days at sample units.

Campbell (2004) also used a subset (33%) of the same systematic track plate data we have analyzed here (Table 1). In all, 177 sample units from the Sierra Nevada were included, consisting of 95 of the standard 6-station enclosed track plates and 82 modified with the addition of 2 unenclosed track plates and 2 remote cameras. Each sample unit was run for 16 consecutive days to survey for fishers and other mesocarnivores during the summer and fall of 1996-1999 in the central and southern Sierra Nevada mountains of California. Fishers were detected at 10 (5.6%) sample units, and at a total of 14 stations. LFD was 2.8 days (SE = 0.7) at enclosed track plates and 5.3 days (SE = 0.88) at remote cameras; these differences were not statistically significant.

Campbell (2004) differed in the approach to analyzing the detection histories for each sample unit by disaggregating all stations within sample units and analyzing stations separately. This approach ignores the likely dependency between both the initial detection within a sample unit and the subsequent revisits to additional stations within a sample unit and thus results in optimistic outcomes. Her estimates of p represent those for a single visit to an *individual* station. These probabilities were: enclosed track plates ($p = 0.154$, SE = 0.062), open track plates ($p = 0.005$, SE = 0.002), and remote cameras ($p = 0.096$, SE = 0.078). Using these estimates, Campbell (2004) determined that the protocol with the least effort required to achieve a $P \geq 80\%$, given species presence, consisted of 2 enclosed track plate stations, run for 5 visits (10 days), yielding a $P = 0.81$. Five alternative sample unit designs that achieved $P > 0.80$ included: 2-closed track plates, 6 visits (12 days) = 0.87; 3-closed track plates, 4 visits = 0.87; 2 closed track plates, 8 visits = 0.90; 4 closed track plates, 3 visits = 0.93; 3 closed track plates, 1 remote camera, 3 visits = 0.84; and 6-closed track plates, 0-1 open track plates, 1-2 remote cameras, 7-8 visits = >0.99.

Royle et al. (2008) also used a subset (81%) of the systematic track plate dataset we have analyzed (Table 1). A total of 464 sample units were included, 64 (13.8%) with fisher detections, following the previous 6-station, 8-visit, 16-day survey duration protocol (Zielinski et al. 2005) mentioned previously. These surveys were conducted from southwestern Oregon south through northwestern California, southern Cascade and Sierra Nevada mountains of California during the summer and fall from 1996-2002. Royle et al. (2008) applied a simple occupancy model (assumes that p is constant across all visits and no additional covariates added), yielding the parameter estimates $p = 0.40$ and $\psi = 0.14$. They noted that there was little adjustment in (estimated) occupancy rate needed (observed $\psi = 0.138$) because the (estimated) power to detect fishers with the given survey method and design was high ($P = (1 - (1 - 0.40)^8 = 0.98$). Note that the estimates of P and p by Royle et al. (2008) and Zielinski and Mori (2001) are identical, despite using different subsets of the systematic track plate dataset.

Tallancy (2005) used track plates, remote cameras, and hair snares to determine the distribution of mesocarnivores in 8 National Park Units in the northeastern U.S. (Table 1). Sample units consisting of 5 detection devices (1 remote camera, 2 hair snares, 2 closed track plates) were surveyed at 47 locations for 2, 2-week periods during the summer and winter in 2004. Sample units were baited with chicken and cat food, olfactory predator lure (Cronk's Outdoor Supplies, Wiscasset, ME), and visited every 3 days to replace bait and remove tracks, hair, or film. Single-visit detection probabilities for fishers were higher in the summer (p = 0.357, SE = 0.05) than in the winter (p = 0.232, SE = 0.06). Accordingly, the protocol had a higher detection probability during the summer (P = 0.89, SE = 0.03) than in the winter (P = 0.73, SE = 0.08; Table 1).

Gompper et al. (2006) assessed the value of 4 non-invasive methods (track plates, remote cameras, snow tracking, and scat surveys) to detect fishers, martens, and other mesocarnivores in Adirondack State Park (Table 1). 5-km transects were selected at 54 sites in Adirondack State Park; on each transect 3 remote cameras were run for 30 days and 6 track plates were run for 12 days. Seasons when surveys occurred and whether camera and track plate stations were run simultaneously were not reported. Stations were baited with either chicken, deer, or beaver and each station received an olfactory lure (Gusto, Minnesota Trapline Products, Pennock, Minnesota). For fishers, remote cameras and track plates were approximately equivalent in detection efficiency; however they reported a higher LFD for track plates versus cameras, hypothesized to be due to the difference in the survey apparatus (closed box versus bait and cameras on trees). Snow tracking yielded the highest single visit probability of detection for fishers (p = ~0.45) compared to track plates or cameras (p = 0.13, 0.10, respectively). Remote camera sample units had a higher detection probability (P = ~0.90) than track plate sample units (P = ~ 0.80), however survey effort between the two sample units was not equivalent (Table 1). Track plates had both a higher probability of detection for martens and detected martens more often than cameras. Gompper et al. (2006) concluded that track plate, remote camera, and snow tracking can all be effective methods to detect fishers and martens.

Long et al. (2007) compared the abilities of scat detection dog surveys, single remote camera stations, and pairs of hair snares (scented carpet pads with nails) to detect fishers and other species (Table 1). Cameras and hair snares were left for 14 consecutive days and scat dog surveys covered a 2-km diamond shaped transect at each site. Remote cameras were baited with 2 pieces of chicken, a sack of fish food pellets, and an olfactory lure (Gusto). During the summers of 2003-2004, they compared these techniques at 168 sites throughout Vermont. Scat dogs had a higher estimated probability of detecting fishers during a single visit, p = 0.84 (SE = 0.07) versus single remote cameras (P = 0.28, SE = 0.07). The hair snares used in this study did not detect any fishers. Long et al. (2007) recommended that despite a higher per site cost, scat dog surveys were substantially more effective and efficient (cost/detection) than remote cameras and hair snares.

Slauson and Zielinski (2007) conducted strategic surveys for fishers and martens at 35 sites on the Mendocino National Forest from July-September of 2006 (Table 1). Six-station track plate sample units were established in 4mi^2 blocks following the protocol of Zielinski and Kucera (1995). Each station was run for 14 consecutive days, had lure applied on the set up day, and was visited every 3-4 days to replace bait and remove tracks. Fishers were detected at 20 (54%) sample units, yielding a mean LFD of 8.8 days (SE 0.8). The probability of detecting a fisher, given one was present, was 0.23 (SE = 0.07) for a single visit and 0.65 (SE = 0.08, 95% CI 0.50 – 0.80) for the entire 14-day, 4-visit protocol (Table 1). Using methods to discriminate sex, based on track measurement (Slauson et al. 2008), 10 male and 11 female fishers were identified at the 20 sample units. Although differences between the single-visit detection probability for male versus female fishers was not statistically significant (p = 0.26 versus 0.11, respectively), the probability of detecting a male versus female fisher using the 14-day protocol was nearly double (P = 0.71 versus 0.37, respectively). Slauson and Zielinski (2007) hypothesized that the overall low probability of detection they observed was due, in large part, to high temperatures experienced in the summer season during their survey effort.

Estimating Detection Probability and Evaluating the Factors that Affect it

Track Plate: Mean LFD was 6.9 days, occurring after 43% of the survey duration (Table 2). The 95% confidence interval for LFD was 5.9 to 7.8 days, representing 37 to 49% of the total survey duration. Track plate surveys occurred in most months of the year, however >80% of all sample units were surveyed between June and October (Figure 2). LFD, sample unit visits, and total station visits varied by month (Figure 3). LFD was lowest and sample unit and total station visits highest during May-June and October compared to July-September (Figure 3). The proportion of sample units with fisher detections, by month, revealed a similar pattern (Figure 4).

There was a significant effect of month on detection probability, sample units surveyed between October and March (n = 30 sample units with fisher detection) had a significantly higher single-visit detection probability (p = 0.5083, SE = .0322) than sample units surveyed between May and September (n = 51 sample units with fisher detections, p = 0.3234, SE = 0.023; t = -4.64, df = 81, P = <.0001). The months of July-September (hereafter the summer season) and October-June (hereafter the non-summer season) had the largest differences in detection probability (Figure 5). Visual inspection of detections during these two periods and their geographic distribution did not reveal any geographic variation that may have affected the results by season. We included *season*, defined as 'summer' (July-September) and 'non-summer' (October-June), in model development.

The number of inoperable station visits was summarized from the full collection of sample units (n =530) and their stations (n = 3,404), which included all sites where fishers were and were not detected. Because it cannot be determined precisely when the inoperability occurred (e.g., day 1 versus day 2), the entire *visit* was considered inoperable. The mean number of inoperable track plate station visits was 4.9 (SE = 0.26, 95% CI = 4.4 to 5.4, range = 0 to 38), equivalent to 10.2% of the total visits (n = 48) for the survey protocol. The mean number of inoperable station visits per sample unit varied significantly by season (F = 3.21, df = 3, P = 0.022, Table 3). As expected, due to bears becoming less active, the winter season (Dec.-Feb.) had significantly fewer inoperable station visits than other seasons (linear contrast, F = 9.03, df = 1, P = 0.0028). The mean number of inoperable station visits was significantly higher at stations where fishers were detected (1.1 visits, SE = 0.12) versus not detected (0.75 visits, SE = 0.024; t = -2.93, df = 191, p = 0.003). The correlation between higher rates of inoperability at stations where fishers were detected is expected due to the largely overlapping ranges and co-occurrence of fishers and black bears in California (Dark 1997, Campbell 2004).

There was no significant difference in the mean single-visit detection probability between the Klamath (p = 0.378) or southern Sierra (p = 0.422) fisher populations (t = 1.06, df = 81, p = 0.293). Distance from coast (in the Klamath population) was not correlated with the number of fisher visits to a sample unit (r = 0.29) or LFD (r = -0.06). The number of fisher visits to sample units in the southern Sierra and Klamath populations were not strongly correlated with elevation (r = -0.33, r = -0.12, respectively). For these reasons, neither population, distance to coast, or elevation were included in models developed to estimate probability of detection.

The number of fisher visits to sample units and the number of sample units within 8.5 km radii neighborhood were not correlated (r = 0.47). However, there was a significant correlation when sample units within 17 km radii neighborhood were considered (r = 0.79). Further examination of the sample units on the periphery versus the core of the known ranges of each population, revealed no significant differences between the means of either LFD (t = 10.45, df = 69, P = 0.65, Table 4) or the number of fisher visits to sample units (t = -1.03, df = 75, P = 0.30, Table 4). This suggests that on the periphery of fisher populations, where fishers potentially occur at lower densities, track station protocols perform equally well as in the cores of populations where fishers likely occur at higher densities. Thus, core versus periphery was not included in model development.

In neither the Klamath or Southern Sierra population was number of fisher visits or number of station visits significantly correlated with predicted habitat suitability (P>0.05). This suggests that predicted habitat suitability, at the landscape scale, does not influence the number of fisher visits to the sample unit and thus p. For this reason, habitat suitability was not included in models developed to estimate probability of detection. This result differs from that of Carroll et al. (1999) who found a significant relationship between the number of fisher visits and habitat suitability at the landscape scale in the Klamath region.

The resulting ratios of expected to observed visit transition counts ("00" and "11" transitions) for the 2-group, visit-specific p model are all somewhat lower than observed which suggests that there is still some dependence to account for (Table 5). Adding a simple Markov persistence factor produces ratios of observed and expected visit transition counts much closer to 1.00, reducing the impact of the dependency. For this reason, a Markov persistence factor was included in model development.

Using the results of the analysis described above, we included 8 custom models, 3 'pre-defined' models (1-group visit-specific p, 2-group constant p, 2-group visit-specific

p) included in Program PRESENCE, and a null model into the candidate model set (Table 6). The 8 custom models represented individual hypothesis about how combinations of season, visit dependence, survey-specific p, and surrogates for these (visit number and linear in logits) best describe the survey data. Season and Markov-chain persistence factors were significant in the univariate analysis, and thus are included as covariates in candidate custom models.

The model with the lowest ΔAIC_c value (Model 1, Table 6) contained the 3 variables with the highest importance weights (Table 7): 1 Markov chain persistence factor, visit-specific p, and 2 Seasons. Model 1 had nearly all the Akaike weight (w) (0.89) and was 9.9 times more likely than the next best model (Table 6). For the top model, ψ was estimated as 0.2071 (SE = 0.024) and the Markov chain persistence factor was estimated as 0.4528 (SE = 0.057). Estimates of the visit-specific detection probabilities within each of the 2 'seasons' differed significantly for most visits (Figure 5). The overall per sample unit estimate for P for the 16-day, 8-visit track plate protocol for the summer (July-September) was 0.607 (SE = 0.057, 95% CI 0.49 – 0.72) and non-summer (October-June) was 0.956 (SE = 0.003, 95% CI 0.95 – 0.96). To achieve a P of >0.80 and >0.95 per sample unit during the non-summer season a 10 and 16 day survey duration would be required per sample unit, respectively. To achieve a P of >0.80 and >0.95 per sample unit during the summer season a 30 and 58 day survey duration would be required per sample unit, respectively.

<u>Remote Camera:</u> Mean LFD per sample unit for the Collins-Baldy survey was 9.9 days and for the Ashland survey was 25.2 days, representing 35.3% and 90.0% of a 28-day survey duration, respectively (Table 2). The 95% confidence intervals for LFD at the Collins-Baldy and Ashland surveys were equal to 24-47% and 28-152% of a 28-day survey duration, respectively. The mean survey durations for the Collins-Baldy and Ashland surveys were 16.3 (SE = 2.0) and 34.0 (SE = 1.9) days per sample unit, respectively. The mean number of inoperable days per station for the Collins-Baldy and Ashland surveys were 3.5 (SE = 0.75, 95% CI 2.0 -5.0) and 1.6 (SE = 0.05, 95% CI 0.5 - 2.7), representing 22% and 0.05% of their mean survey durations, respectively.

The models with the lowest ΔAIC_c values for the Collins-Baldy and Ashland surveys both included a single group, constant p, and the Markov chain parameters (models 1, Table 9). The top model in each case accounted for all of the Akaike weight and were $\geq$29 ΔAIC units lower than the second best models (Table 9).

For the Collins-Baldy survey, model 1 estimated a P of 0.930 (SE = 0.007, 95% CI = 0.92 to 0.94) for the 28-day survey duration (Table 10). The estimate for the Markov-chain persistence factor was 0.518 (SE = 0.058), showing a similar level of dependency in the detection histories as the track plate dataset. Using the estimate for p, a survey duration of 17 and 32 days during the non-summer months would yield a P >80% and >95% per sample unit, respectively. However, the estimates for this study area are likely biased high due to the truncation of the survey duration after a fisher was detected, effectively excluding more survey days with non-detections from the full detection history than would be expected by chance.

For the Ashland survey, model 1 estimated a *P* of 0.671 (SE = 0.155, 95% CI = 0.36 to 0.98) for the 28-day survey duration (Table 9b). The estimate for the Markov-chain persistence factor was 0.098 (SE = 0.128), a low value, showing less dependence than the Collins-Baldy and the track plate datasets. To achieve a *P* of >0.80 and >0.95 per sample unit, a 41 and 71 day survey duration would be required during the non-summer months, respectively. For both remote camera study areas, the SEs are biased low and the 95% confidence interval are overly precise due to the violation of independence between sample units, such that the same individuals can be detected at >1 sample unit, requiring the development of a variance inflation factor to adjust their estimates; This step, however is beyond the scope of this analysis.

Hair Snare: Mean LFD was 11.2 days representing 50.9% of the survey duration (Table 2). The 95% confidence interval for LFD was equal to 34-64% of the total survey duration. The mean number of inoperable visits was 0.17 (SE = 0.08), representing 0.02% of the survey duration.

The model with the lowest ΔAIC_c value (Model 1, Table 11) included the most simple parameters, a single group and a constant detection probability. Model 1 was highly supported as being the best model, with a ΔAIC difference of >6 compared to the second ranked model and receiving >95% of the Akaike weight (Table 11). Model 1 estimated a *P* of 0.813 (SE = 0.094, 95% CI = 0.62 to 1.0) per sample unit for the 22-day survey duration (Table10). To achieve a *P* of >0.80 and >0.95 per sample unit during the spring (March-May), a 22 and 38 day survey duration would be required, respectively.

Effects of Spatial and Temporal Changes in Survey Protocols on Detection Probability

Practitioners of survey protocols are often interested in how changes to the protocols will affect their ability to detect the target species. We investigated the effects of two types of changes to the track plate protocol, changing the number of stations or visits. The reduction of the number of stations and visits had similar effects on detection probability at low levels, such that the removal of a single station (from 6 to 5) or visit (from 8 to 7) resulted in <0.05 reduction in detection probability in the non-summer season (Figure 6a). However, the effect of reduction of either stations or visits from the 8-visit protocol results in a larger effect on *P* during the summer season, due to a lower overall *P*, with the reduction of a single station resulting in a > 0.05 decrease in *P* (Figure 6b). The effect of removing an increasing number of stations versus visits is not consistent. In general at high *P* levels (e.g., > 0.80) the effect of reduction in the number of stations is slightly more detrimental than the same amount of reduction of visits. However, at lower *P* levels (e.g., <0.65) the effect of the reduction in the number of stations is approximately twice more detrimental than the reduction in visits (Figure 6).

Station inoperabilities are commonly encountered when using detection devices, however they are rarely reported or explicitly considered when reporting survey results. The effects of inoperabilities on *P* are substantially larger for the 2-station remote camera protocol versus the 6-station track plate protocol (Figure 7, 8). If 20% of the survey duration is affected by inoperabilities, *P* decreases ≤ 0.05 for the track plate protocol versus >0.15 for the remote camera protocol. *P* decreases to <0.80 when >20% of the survey duration is affected by inoperabilities for the remote camera protocol, while it takes >60% of the track plate protocol survey duration during the non-summer months to be affected to cause *P* to decrease <0.80. While the increased number of stations in track plate sample units clearly play a role in this higher resiliency to inoperabilities, visits frequency does as well. Track plates are visited every 2 days, limiting the number of days a station can be inoperable to 2, while remote cameras are typically checked once a week and accordingly inoperabilites can be up to 7 consecutive days. Figure 9 demonstrates the reduced effect of inoperability as the number of days between visits decreases. By visiting stations every 3 or 4 days instead of every 7, the percent inoperability required to reduce *P* below 0.80 is doubled to >40%.

DISCUSSION

Post-hoc Studies

Previous authors using portions of the same track plate dataset we have analyzed here have reached alternative conclusions, largely because they failed to consider detection heterogeneity. Both Zielinski and Mori (2001) and Royle et al. (2008) used substantial portions (67% and 81%, respectively) of the same data, but used simpler models that assumed a constant *p* and did not account for detection heterogeneity. Their estimates of *P* (0.98) and *p* (0.40) were identical, but were only similar to our non-summer estimate (*P* = 0.99) and much higher than our summer estimate (*P* = 0.76). Royle et al.'s (2008) estimate of ψ (0.138) was more negatively biased compared to our estimate of ψ (0.207), which is expected when detection heterogeneity is not accounted for (MacKenzie et al. 2005). The estimate of ψ, using the top ranked model presented here, results in a 50% increase in the estimate of ψ when applied to the full dataset (n = 537 sample units) compared to prior estimates using 67-81% of the dataset and simpler models; this represents an estimated difference of 36 additional sample units that were presumed to be occupied, but where fishers failed to be detected. Accounting for, and modeling properly, detection probability makes a significant difference in estimating the number of occupied sample units.

Campbell (2004) used 33% of the data used here, but used a distinctly different analytical approach, treating each station individually for the analysis. In Campbell's (2004) analysis dataset, fishers were detected at an average of 2-3 stations per sample unit. We believe that this approach is biased due to the dependencies that exist between stations within each sample unit. At sample units where fishers are detected at more than a single station, disaggregation of the detection results into separate station-specific detection histories is a violation of the assumption of independence between detection histories (MacKenzie et al. 2005). In most cases, one or two individual fishers are detected at a sample unit (R. Truex pers. comm.). Unless the detection of these individuals can be separated (e.g., via sex discrimination or individual identity), the detection histories cannot be assumed to be independent at stations within a sample unit. Furthermore, because the likely effective area each sample unit covers (~314 ha) represents a small portion of a average fisher's home range in California or Oregon (8% for males and 23% for females; Truex et al 1998, Mazzoni 2002, Yeager 2005, Green Diamond Resource Company unpubl. data) and the distance between all stations can likely be traversed by fishers within a few hours, it is unlikely that stations are independent. The computational consequence of violating this assumption is to underestimate the variation in p, P, and ψ (e.g., SE and confidence intervals; MacKenzie et al. 2005).

Campbell (2004) also used her station-based estimates of p to propose a number of sample unit designs to meet multiple thresholds for P. The numbers of stations and visits Campbell (2004) proposed to meet estimates of $P > 0.80$ were far lower that the effort (number of stations and visits) we have proposed (Table 10). Beyond the issues mentioned above these differences are due in large part to distinctly different approaches to the respective analyses. Due to the spatial dependencies between the stations and the visit-specific detection probabilities, investigation of changing the sample unit design should ideally be accomplished by simulating the effects of incrementally removing stations and visit-specific detection probabilities on P. One could argue that if it can be demonstrated that specific stations within each 6-station sample unit routinely detect fishers above all others, an approach focusing on the detection histories from just those stations would have some validity. However, the systematic track plate data does not show any consistent pattern of detections with respect to station location (1-way ANOVA, df = 5, F = 0.28, p = 0.91). Thus an approach only using the detection histories from stations where fishers were detected to develop *sample unit design* recommendations will not be valid. Developing sample unit recommendations in this manner will result in over-optimistic projections of P for sample unit designs.

Despite having different protocols, involving different device types, survey durations, and visit intervals, most of the post-hoc studies reviewed had moderate to high estimated probabilities of detection, but most were only moderately precise for their respective protocols (Table 1). Tallancy's (2005) 5-station mixed device (camera, track plate, hair snare) sample unit had a high and precise P during the summer, but was lower and much less precise during the winter, demonstrating a seasonal effect but one that was opposite of the effect seen here. Gompper et al. (2006) found that remote camera and track plate sample units had similar estimates of P (considering the slight inequity in sampling effort, Table 11). However, it was not clear whether camera and track plate sampling occurred simultaneously in each sample unit, which could introduce dependency (e.g., due to learning) between the detection histories observed. Long et al. (2007) had the lowest overall sampling effort and understandably, the lowest estimate of P for a single camera station sample unit.

Slauson and Zielinski (2007) had the lowest estimate of P for a 6-station sample unit, but their estimate ($P = 0.65$, SE 0.08) is close to that estimated from using the visit-specific detection probabilities for 7-visits (14 days, $P = 0.71$, SE = 0.06) from the summer track plate surveys analyzed here. Thus, while the protocol was slightly modified (3-4 day visit intervals, 4-visits, 14 days total) from the standard 16-day, 8-visit protocol (Zielinski and Kucera 1995), the estimate of *P* was well approximated by the data analyzed here, and was consistent with the low summer pattern of estimates of *P*.

None of the post-hoc studies we reviewed investigated models with detection-related covariates, perhaps due to limited sample sizes for investigating more complex models or due to the recent developments in the techniques for modeling of *p*. Thus, our modeling is an improvement over what has been conducted previously.

Estimating Detection Probability and Evaluating the Factors that Affect It

Most of the environmental covariates we investigated had little effect on the probability of detection at track plates. The fact that habitat suitability, core versus population periphery, and geographic location did not significantly affect fisher visits, and thus their probability of detection, reinforces the success and consistency of the track plate protocol to detect fishers throughout their geographic range in California and Oregon. Although our sample was inadequate to investigate these same covariates for protocols for the cameras or hair snares, we do not have any reason to suspect that they would differ from track plates.

Several of these covariates (distance to coast, elevation, core versus periphery of distribution, and landscape habitat suitability) were hypothesized to be related to density but none of them had any effect on detection probability. Royle and Nichols (2003) identified density as an important source of detection heterogeneity in song bird point counts and Smith et al. (2007) found that the density of martens had a large effect on detection probability. Because we did not determine the density of fishers present in any sample units and test it as a covariate, we cannot rule out its potential effects on detection probability. However, because fishers have much larger home ranges than martens, and occur at much lower densities, the effect of variation in density is not likely to be significant. In the southern Sierra, 34 6-station track plate sample units, retrofitted with hair snares, to collect DNA for individual identification, have only detected a mean of 1.4 fishers (range 1-4) per sample unit (R. Truex pers. comm.). In the interior Klamath region of California, 20 4 mi^2 sample units with 2 hair snare stations per sample unit detected a mean of 1.65 fishers (range 1-3) per sample unit (Farber and Schwartz unpubl. data). Due to large home range size, relative to the area covered by a sample unit, and exclusion of same-sex conspecifics from home ranges, detection probabilities for fishers have a much lower potential to be significantly influenced by density than many other species. Thus, our findings for lack of significance for variables likely correlated with density and the relationship between fisher home range size and sample unit area suggest that density may not be a critical factor in determining detection probabilities for fishers using the protocols described here.

The assumption of independence between visits in the detection histories, that the status of any visit is independent of any other visit, does not appear valid for the track plate detection history data. Positive dependence was present in the data, such that the detection outcome for any visit was more likely to be the same as the prior visit than expected by chance. This likely relates to fisher movements and responses to a food resource. Because the sample units only occupied a small portion (e.g., <20%) of a fisher's home range, sequential 0's are more likely due to fishers being elsewhere in their home range. Then, once fishers detect the sample unit they are more likely to be detected there again (sequential 1s), because they are already in the vicinity and perhaps because they have been reinforced with food. The Markov-chain persistence factor was in the top models for the track plate and remote camera data. While not in the top model for the hair snare data, this should be revaluated with a larger dataset when available. Accounting for the lack of independence in the detection histories removes additional heterogeneity and results in more precise estimates of p (MacKenzie et al. 2005).

Our results underscore the need to test the assumption that p is constant across visits by including candidate models with visit-specific estimates of p. For the track plate dataset, estimating visit-specific detection probabilities led to the understanding that p increased over the survey duration. This understanding has important implications for making changes to survey protocols. For example, Zielinski and Mori (2001) assumed a constant p and concluded that, for a 6-station sample unit, the number of visits could be decreased from 8 to 4 and still maintain a $P > 0.85$. Using the new visit-specific estimates for p, a 4-visit protocol would result in a $P = 0.83$ during the non-summer season, but only $P = 0.44$ during the summer. While models with visit-specific p were not the top ranked models for the remote camera and hair snare datasets, these datasets were relatively small and models with visits-specific p should be evaluated with larger datasets that can accommodate the additional parameters. Use of the visit-specific estimates of p allows for a more accurate understanding of detection heterogeneity within survey protocols. More accurate projections using visit-specific estimates of p result in more precise estimates of the variation associated with P, and more precise estimates of occupancy (ψ).

The season in which surveys were conducted had an important effect on detection probability. Surveys conducted in the summer months (July, August, September) had lower detection probability than in the spring and fall. While it appears the winter-early spring season (December-April) also has a higher detection probability, we had too few data available to state this with confidence. However, this finding is consistent with both published and unpublished accounts of seasonal capture or detection differences for fishers and martens. Ivan (2000) observed that marten took 2-3 times longer to capture during the summer than winter, an observation supported by others (Hawley 1955, Raphael 1994). Bull et al. (1992) also found that detection rates for remote cameras and track plates were lower in the summer and recommended winter surveys for martens in Oregon. Peaks in capture success for fishers have been reported in the fall (September-November) and late winter through spring (February- April; Johnson 1984, Aurthur 1988, Weir 1997) although none of these studies trapped during the summer (June-August). However, Tallancy (2005) found just the opposite for fishers in the northeastern U.S., detection probability was higher during the summer (Table 1). This suggests that there may be regional variation in the pattern of this phenomena, but studies of fishers and martens in the western U.S. consistently support the non-summer seasons as having a higher detection or capture probability than the summer season.

The fall and winter periods are when the young of the year fishers are independent from their mothers, and are still learning to forage for themselves, and are dispersing. Thus, the fisher population is largest in the fall and this may continue into the winter until mortalities begin to reduce the population. Furthermore, early in the fall young of the year fishers are still learning to capture prey, and would be expected to repeatedly visit stations where bait is available. These two factors may also contribute to the higher probability of detection, especially during the month of October (Figure 3).

The summer months are the time of the year when food resources are most abundant and perhaps easiest to obtain (e.g., many young small mammals and birds present). Thus, fishers may not need to travel as far or frequently to acquire food or they may not be as interested in scavenging from the bait used with detection devices. Furthermore female fishers with kits use smaller portions of their home ranges during this period, reduce their movements in order to remain close to the den site, and may not select carrion for their own or their kits consumption (Leonard 1986). In addition, summer temperatures >30° C are common and likely result in fishers having to reduce their activity and behavioral patterns to minimize heat stress.

Anticipating detection heterogeneity and minimizing its effects, either in the study design phase where possible or by incorporating covariates, is essential for good performance of models for estimating detection probability and occupancy (MacKenzie et al. 2005). Indeed, models that did not account for detection heterogeneity produced negatively biased estimates of ψ from the track plate data. Specifically, the two-season and survey-specific covariates accounted for detection heterogeneity that represented a violation of the assumption of constant probability of detection. Violation of this or other assumptions results in biased estimators and inferences about factors that influence either occupancy or detection may be incorrect (MacKenzie et al. 2005). Use of the estimates from a simple model (constant p) resulted in the incorrect recommendations that the track plate protocol could be modified while maintaining P above a specific threshold (e.g., >0.80). Furthermore, not considering covariates that could account for detection heterogeneity present in the dataset led to the conclusion that a single protocol would yield a constant P when conducted any season of the year. Clearly, detection heterogeneity is an important factor to consider in the design and analysis of occupancy surveys for fishers. While the large systematic track plate dataset provided an opportunity to fully explore their effects, we do not believe these effects to be specific to this type of detection device and encourage investigation of factors causing detection heterogeneity in camera and hair snare surveys for fishers and related carnivores.

Evaluation of Detection Device Protocols

Track Plate--. Estimates for detection probability for the track plate protocol were both high and precise, resulting in an extremely precise estimate of ψ. During the non-summer months, the protocol achieves a near perfect level of detection. We caution however, that not all months during this period had large sample sizes (e.g., December-April) and these months should be considered carefully until future survey results confirm the relationship found here. Surveys conducted during the summer months had a lower and less precise estimate for detection probability compared to other months, however their estimate was still high and precise compared to the other protocols reviewed, and compared to estimates for the cameras and hair snares.

Remote Camera--. Estimates for the remote camera protocol were mixed; high and very precise in the Collins-Baldy study area, but moderate and low precision in the Ashland study area. The favorable estimate in Collins-Baldy is inflated for two reasons: (1) the surveys were truncated after a fisher was detected and (2) the spacing of the sample units violated the assumption of independence between sample units, requiring a variance inflation factor to adjust the estimates. The effect of ending surveys after the first detection of a fisher can be seen in the lower precision in the estimate of occupancy, despite a protocol with a high and very precise estimate of detection probability. If a precise estimate of occupancy is the goal of a survey, which is the goal for both single location and distributional surveys, the entire survey duration should be used. However, if the sample size is large, which is typical for distributional surveys, a proportion of sites with detections could be truncated after the first detection to increase efficiency.

The estimate of detection probability for the Ashland study area may reflect the low suitability of this study area to support fishers during the winter and perhaps year-round, resulting in a small number of fishers present in the study area. Detections occurring above 4,000 feet only occurred before heavy snowfall commenced (Farber and Criss 2006). This suggests that this study area may occupy the distributional limit for fishers in this region, due to the limiting effects that deep, soft snow can have on fishers (Krohn et al. 2004). This may explain the low detection probability and occupancy observed. Furthermore, the Ashland survey had a mean LFD 2-4 times higher than for any other study area and protocol. Unlike the other surveys, heavy snowfall events occurred during some of these surveys, likely affecting the mobility of fishers to traverse their home ranges and encounter camera stations. Survey durations were extended to compensate for this likelihood, resulting in detections after 28 days (Farber and Criss 2006). Thus, future winter surveys should remove the effects that winter storms may have on fisher detection by extending survey durations when such heavy snowfall events occur.

Hair Snare--. The estimate for probability of detection using the hair snare protocol was high, considering each sample unit consisted of only a single station. The precision of this estimate, however, was low. The precision, and likely the estimate of detection probability, will increase with the inclusion of additional hair snare stations to this protocol, especially when used during the non-summer months. The protocol proposed by Schwartz et al. (2006), consisting of 4 stations within 25 mi^2 sample units for 21 days, will undoubtedly exceed both the estimate and precision of detection probability reported here. However, if this protocol is used during the summer months, it's precision should increase but the estimate of detection probability may not change substantially. When new data are available, it should be used to calculate detection probability using the same methods as we have here.

The potential effects of using different baits at detection device stations has received some attention (Zielinski and Kucera 1995). Practitioners realized that a standardized bait was needed to make efforts comparable and chicken has been adopted for most field protocols. As important as the type of bait used, is the amount used. The three protocols we analyzed used substantially different quantities of bait, from small 3x1x1 in pieces, to 3 chicken drumsticks with 2 cans of cat food. Moreover, extreme examples include up to 10 pounds of chicken placed at each camera station (Lindstrand 2007). The importance of the amount of bait is the effect it may have on the revisitation of fishers to stations and thus the estimate of detection probability. Indeed, the protocol with the smallest bait reward (hair snare) did not show dependency in the detection histories, unlike the track plate and camera protocols that used larger baits. Ideally, both bait type and quantity would be standardized by practitioners of detection device protocols. We recommend that bait quantity be related to the number of days between checking stations, such that the quantity can be standardized and that baits used are described by types (e.g., chicken drumstick) and weight (e.g., 40 g). We recommend that 1,2 and 3 drumsticks be used when between-visit intervals are 2-3, 4-6, and 7-9 days, respectively. The benefit of such standardization will be the removal of the 'size of the reward' effect that bait quantity may be having on revisitation rates and facilitate more direct comparisons of detection devices and their protocols. Inclusion of the weight of bait used can also be used as a covariate to explain variation in detection probabilities or higher levels of revisitation after initial detection.

Effect of Spatial and Temporal Changes in Survey Protocols on Detection Probability

The remote camera protocol is inherently more sensitive to inoperabilities due to the reduced number of stations and comparably longer (e.g., 1-week) check intervals. Most inoperabilities to detection devices are from black bears, equipment failure (e.g., batteries run out), and field technician error. One solution to reducing the potential for inoperabilities due to bears is to survey during seasons when bears are inactive, the winter. However, winter is when juvenile dispersal and mortality rates are highest, with the greatest potential to violate the closure assumption. Because the average number of inoperabilities at track plate sample units where fishers were detected was low (1.1, 2%, or 2.2 days) we do not think inoperability due to bears is sufficient justification to shift to winter surveys.

We generally recommend that survey durations be extended when ≥10% (≥6 total survey days) of the 2-station 28-day remote camera and when ≥20% (≥19 total survey days) of the 6-station 16-day track plate protocol survey durations have been affected by inoperability. These threshold represent the point where there is a >5% reduction in P. Survey duration can be extended until the necessary duration has been reached. However, if a substantial period (e.g., ≥14 days for a remote camera survey) has passed with stations being inoperable, the entire survey duration should be repeated to remove any potential for changes in visitation behavior (e.g., that individuals may ignore stations due to lack of bait present).

These results also have more broad implications for carnivore survey design using baited detection devices. Sampling designs with more spatial replicates that are checked more frequently are more robust to inoperability problems than those with fewer spatial replicates checked at longer intervals.

FUTURE DIRECTIONS

Empirical Evaluation of Detection Probability

The occupancy estimation we employed relies on the post-hoc analysis of data from sample units where fishers were detected and, consequently, does not include explicit information on the number of sample units failing to detect fishers when they are present in that sample unit. Thus, a fundamental assumption of the occupancy modeling approach is that the sample units where fishers are detected represent a high proportion of the *true* number of sample units where fishers were present and available to be detected. If the sample units where fishers were detected do not represent a high proportion of the total number of sample units with fishers present in sample units, estimates of detection probability will be biased high. An alternative approach to estimating the probability of detection is to use an empirical method which can determine what percent of sample units with fishers present actually detect a fisher. The empirical approach is experimental in nature and we are not aware of any studies that have attempted this with fishers.

Ivan (2000) demonstrated, for martens, that a post-hoc approach to calculating probability of detection (98%) can over-estimate the *true* probability (70%) of detecting martens and underestimate ψ. Thus, in Ivan's study, the sample units where martens were detected (7 of 10) did not closely represent the number of sample units where martens were present and available for detection (10 of 10). However, a concern regarding Ivan's (2000) study design is the unknown behavioral effects of prior live-capture and radio collaring on each marten's subsequent likelihood to enter track plate enclosures. Track plates were deployed 1-12 days after martens were captured, the period of time where the aversive effects of capture are assumed to be the highest. However, a high percent of individuals that had been marked were recaptured in traps up to five times over $\leq$ 9 days, suggesting that device avoidance may also be low. Ivan (2000) conducted this work during the summer, the period of the year when prey is most plentiful, which may have contributed to lower detection rates. Fowler and Golightly (1994) conducted a similar experiment, but detected 5 of 5 marked martens >1 month after initial capture using detection devices placed in their home ranges. This experiment was conducted during the winter, when martens are more energetically stressed and prey is typically more limited than the summer, which may help explain the much higher detection rates for marked martens than in Ivan's (2000) work.

If the magnitude of difference between the empirical and post-hoc estimates of detection probability revealed by Ivan (2000) for martens also apply to fishers, it may represent a significant bias of detection devices. The fact that there have been only two studies to attempt to experimentally test the probability of detection devices on martens and none on fishers, is a concern. Martens have similar patterns of latency to first detection and overall detectability using detection devices as fishers (Zielinski et al. 1997). Because fishers occur at lower densities than martens, however, the magnitude of this effect has the potential to be greater for fishers than martens. Post-hoc approaches rely on the assumption that the animals actually being detected at detection devices represent all or a high proportion of the individuals present in a sample unit. In fact, this assumption represents two separate probabilistic events: (1) the probability that animals present find at least one station during the sampling period and (2) once ≥ 1 station has been found, the animals interacts with it such that a 'detection' occurs. The only way to estimate these probabilities is to have prior knowledge, via independent methods, of the presence and number of individuals within detection distance of a device. Until data can be presented to evaluate this underlying hypothesis for fishers, post-hoc methods to determine the probability of detection should be interpreted cautiously as being estimates, with the potential to be biased high.

Sex-Specific Detection Probability

Slauson and Zielinski (2007) found that male fishers had a significantly greater probability of detection than females; another potential source of detection heterogeneity. Females likely represent ~50% of the total number of individuals in most populations of fishers. If protocols have lower detection probabilities for females, this could result in an underestimate of true occupancy due to non-detections occurring where either only females are present during the survey duration. If either of these cases occurs frequently, occupancy estimates would be biased low. Furthermore it would be important to know whether significant differences between male and female detection probabilities varied by season. And, if the survey protocols are male-biased, habitat analyses conducted using these data could also be male-biased. The extent of the sex-specific bias in survey protocols was beyond the scope of this study, but could be evaluated by identifying gender from tracks, using the algorithm created by Slauson et al. (2008).

MANAGEMENT IMPLICATIONS

Limitations of this Analysis

The track plate dataset was both geographically and temporally extensive and included a large sample size. This allowed for a thorough analysis and investigation of factors that may influence the detection of fishers. The dataset is limited to inference to the native population of fishers found throughout California and southwestern Oregon. We suspect our findings should have general implications for surveys for fishers elsewhere in their geographic range. As mentioned previously, the number of sample units in the dataset from the winter and early spring months (December-April) are limited and thus until additional data are analyzed, we caution the interpretation of our results for these months.

In comparison to the track plate dataset, the remote camera dataset was extremely limited geographically and temporally. Direct comparison between track plate and remote camera protocols was not possible due to the different sampling scales, difference in the seasons for most data, and the alternative protocols used within the camera dataset. The remote camera results apply to the fall, winter, and spring seasons and are limited to the interior Klamath region of northwestern California. Additional analysis of remote camera data from a larger geographical area and from other portions of the year would increase our understanding of how protocols using this device vary.

The hair snare dataset was the smallest, but represented one of the first estimates of detection probability from hair snares. The protocol represented the minimal amount of effort per location (1 station), but still yielded a high P (>0.80) during the spring season. The future analysis of data from the 4-station hair snare sample units to survey for fishers in the Rocky mountains (Schwartz et al. 2006) should yield important insights into the ability of this more robust protocol to detection fishers using hair snares.

Survey Protocol Modifications

The objectives of any occupancy survey will drive the selection of the survey protocol to achieve the desired level of P per sample unit. For large scale distributional surveys (e.g., Zielinski et al. 2005) or monitoring efforts (Truex 2003), a lower level of P is tolerated due to the tradeoffs between maximizing the number of total locations that can be surveyed versus maximizing P for fishers at each location surveyed; this optimizes both the spatial extent and precision of the information to achieve goals of detecting a distributional change or developing predictive habitat models. For single-location or pre-project surveys, the objective is often to confirm the presence of fishers in an area where management activity may affect its habitat. In this case, the consequence of failing to detect a fisher when it's present is high, so the desired level of P should also be high (≥0.95). Thus practitioners of occupancy surveys should select the survey protocol that will achieve the desired level of P to meet their objectives.

We propose the following modifications to the Zielinski and Kucera (1995) occupancy survey protocols in order for these protocols to achieve the appropriate levels of detection probability for large-scale and small scale survey objectives. All recommendations are based on maintaining the 4 mi^2 sample unit, the same device-specific spacing recommendations (minimum of 0.8 km between track plate and 1.6 between camera stations), and the same general number of stations per sample unit (2-cameras, 6-track plates) as originally described in Zielinski and Kucera (1995). We make our recommendations in terms of the number of *functional survey days* needed to ensure moderate ($P \geq 0.80$) or high ($P \geq 0.95$) probabilities of detection. A survey day is a 24-hour period when a single station within a sample unit is functional. For example, a 6-station track plate unit run for 16 consecutive days has a total of $(6_{\text{stations}}) * (16_{\text{survey days}}) =$ 96 survey days (Long and Zielinski 2008). Thus, "survey days" can be a common currency among all protocols regardless of the detection device used. It effectively represents effort by integrating its temporal and spatial components. We make two assumptions at this stage: (1) track plates, remote cameras, and hair snares have nearly equal probabilities of detecting fishers when present, and (2) the number of survey days is directly related to the probability of detecting a fisher regardless of the detection device used. Finally, these recommendations apply to the use of sample units as independent measures of fisher occupancy. If practitioners seek to use multiple sample units as spatial replicates (e.g., for determining whether fishers occupy a watershed) and want know how much effort to put forth, alternative calculations for P and thus the effort required, will be required. We refer those interested in this topic to Long and Zielinski (2008).

Due to significantly lower detection probability, the summer season (July-September) will require additional effort to achieve either threshold for P. Protocols conducted during the summer months will have to achieve a minimum of 120 and 200 functional survey days to achieve detection probabilities of $\geq$0.80 and $\geq$0.95, respectively. Protocols conducted during the non-summer months (October-June) need a minimum of 50 and 60 functional survey days to achieve detection probabilities of $\geq$0.80 and $\geq$0.95, respectively. Applying these recommendations to the standard 6-station track plate protocol would yield a duration of 20 and 34 functional survey days during the summer season and of 8 and 10 functional survey days during the non-summer seasons to achieve protocols with $P \geq 0.80$ and $\geq$0.95, respectively. Applying them to the remote camera protocol would yield a duration of 60 and 100 functional survey days during the summer season and 25 and 30 functional survey days during the non-summer season. Any survey days lost must be compensated for by extending surveys until the target duration is met. Finally, future survey efforts should include the calculation and reporting of p and P, following the approach outlined in this report, as part of the standard reporting procedure for results.

Although we have made recommendations that will likely lead practitioners to conduct surveys during the portions of the year with the highest detection probabilities, some consideration should be given to both the underlying population being sampled and season-specific logistical challenges. Although the summer season had the lowest detection probability, this is the period of the year when detections most likely reflect the territorial adult and breeding segment of the fisher population. Furthermore, the summer season, as well as the late spring (May-June) and early fall (October-November), are the most logistically feasible seasons to conduct surveys due to weather conditions and day length. Detections during the spring season also largely reflect the territorial adult portion of the population and the juveniles that have survived through the winter. The spring is when females give birth, use small portions of their home ranges, restrict their movements in order to tend young kits (Leonard 1986), conversely making this the most potentially difficult season to detect reproductive females. Detections during the fall have the highest potential of including dispersing non-territorial juveniles. If detections are biased toward non-territorial juveniles, inference to population status or habitat characteristics may not be representative of the territorial adult portion of the population. Winter detections represent the territorial adult segment *and* those juveniles that have survived through the fall. Dispersal may still occur during this season, but it is not likely as significant as in the fall. The objectives of a survey should help guide the selection of the most appropriate survey season. Further research using techniques to identify the gender and individual detected will also help clarify the relative importance of these seasonal issues.

We recommend that single-location surveys or pre-project surveys be conducted during the fall and/or winter seasons (October-March) because this period is when fishers, particularly the reproductive female portion of the population, should be most detectable. Managers that are concerned about the occupancy of fishers when planning habitat alteration should rely on detections from the October-March period to reflect the occupancy status at those locations. Although dispersing juveniles may be of concern, the habitat they occupy represents dispersal habitat and may be important for maintaining habitat connectivity. Furthermore, repeat surveys in subsequent seasons or years can aid in determining the stability of the occupancy status, especially if gender or individual can be determined.

For large-scale distributional surveys, the objectives may be diverse and thus the design considerations diverse to match each sampling scenario. To cover the myriad of options and considerations based on the analysis presented here is beyond the scope of this document. We refer readers to Long and Zielinski (2008) for a more thorough treatment of large-scale survey design and recommendations. Our analysis and results are appropriate for making recommendations for a large-scale effort, the monitoring effort for fishers in the Sierra Nevada. We make the following recommendations as a case study for how our results may be applied to such an effort.

The findings from the analysis of the track plate dataset have important implications for the Sierra Nevada Framework fisher monitoring program. This program was developed by Zielinski and Mori (2001) and has been implemented as designed, but with the use of a 5-visit protocol (Truex 2003).

By using the more complex models that account for detection heterogeneity, our estimate of occupancy was actually higher than previously estimated by Zielinski and Mori (2003) $\psi = 0.45$ versus $\psi = 0.30$, respectively. By using an underestimate of ψ to design the monitoring program, the sample size recommended by Zielinski and Mori (2003) is larger than necessary to detect a 20% decline in the more precise estimate of ψ. However by using a 6-station, 5-visits track plate protocol during the summer period the precision of the estimate of ψ is likely to be lower than anticipated, which may increase the difficulty of detecting a true decline of near 20%.

The Sierra Nevada fisher monitoring program currently conducts occupancy surveys from May-November (R. Truex pers. comm.). Surveys conducted during the summer months (July-September) will have a lower P and if left unmodeled will result in an underestimate of ψ during this period. If the analysis of the new survey data results in a low but precise estimate of P for the 5-visit protocol, the summer month protocol will need to be adjusted to achieve the desired level and precision of P necessary to meet the monitoring program objectives for detecting population change. The desired level and precision of the estimate of P for summer month surveys can either be set to match that of the non-summer months or be further optimized with effort using simulations from the post-hoc power analysis methods recommended above.

We recommend that the analysis of the first 5 years of monitoring data should be conducted to investigate the extent to which these issues may affect the ability of the Sierra Nevada fisher monitoring program to detect a 20% decline,. During this analysis, specific attention should be given to estimating *P* using the top model presented here, and the resulting estimate and precision of ψ. In addition, dependency between occupancy status and detection probability for locations where fishers are detected on consecutive years should be investigated. Using these data, a post-hoc power analysis can be done to confirm the ability of the program to detect the desired level of change in occupancy originally specified (Zielinski and Mori 2001).

Minimally Valid Survey Effort

Our survey results have implications for those attempting to interpret the results of previous surveys for fishers in the Pacific States. Due to the inconsistent application of effort from previous survey protocols, these survey efforts often produced more reliable inferences about presence than absence. Fisher detections are undeniable, but non-detections (inferred absence) are probabilistic and based on the level of survey effort expended. A standard to judge their validity is needed to understand how well these previous surveys characterized the occupancy status of fishers. We recommend that the minimally valid survey effort be defined by the effort required to achieve a *P* of >0.95. Thus, a summer season survey would need a minimum of 200 functional days of survey effort per sample unit and a non-summer season survey a minimum of 60 functional days of survey effort. Survey effort lower than these thresholds should be reported, but are probably insufficient to conclude that fishers are absent.

ACKNOWLEDGEMENTS

We would like to thank Laura Finley and the Yreka Office of the Fish and Wildlife Service for funding this project. We would also like to thank Scott Yeager, Stu Farber, and Tom Franklin for sharing their respective fisher detection data sets. We would like to thank Keith Aubry and Kathy Raley for early assistance with data acquisition. We also thank Jan Werren for GIS assistance and developing figures and Rick Truex for review of the manuscript.

LITERATURE CITED

Arthur, S.M. 1988. An evaluation of techniques for capturing and radio-collaring fishers. Wildlife Society Bulletin 16: 417–421.

Aubry, K. B. and D. B. Houston (1992). "Distribution and status of the fisher *(Martes pennanti)* in Washington." Northwestern Naturalist 73: 69-79.

Aubry, K.B., Lewis, J.C. 2003. Extirpation and reintroduction of fishers (*Martes pennanti*) in Oregon: implications for their conservation in the Pacific states. Biological Conservation 114:79-90.

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

Bull, E.L., R.S. Holthausen, and L.R. Bright. 1992. Comparison of 3 Techniques to monitor marten. Wildlife Society Bulletin 20:406-410.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multi-model inference: A practical information-theoretic approach. 2nd Edition. Springer-Verlag, New York.

Campbell, L. 2004. Distribution and Habitat Associations of Mammalian Carnivores in the Central and Southern Sierra Nevada. Dissertation, University of California, Davis. 167 p.

Carroll, C, W. J. Zielinski and R. F. Noss. 1999. Using presence-absence data to build and test spatial habitat models for the fisher in the Klamath region, U.S.A. Conservation Biology 13:1344-1359.

Dark, S. J. 1997. A landscape-scale analysis of mammalian carnivore distribution and habitat use by the fisher. M. S. thesis, Humboldt State University, CA. 67 pp.

Davis, F. W., C. Seo, and W. J. Zielinski. 2008. Regional and landscape predictors of fisher (Martes pennanti) distribution in California. Ecological Applications 17: 2195-2213.

Farber, S. and T. Franklin. 2005. Presence-absence surveys for Pacific fisher (Martes pennanti) in the eastern Klamath province of interior Northern California. Timber Products Company, Unpubl. report. 35 p.

Farber, S. and S. Criss. 2006. Cooperative mesocarnivore surveys for the upper and west fork Beaver creek watersheds in interior Northern California. Timber Products Company, Unpubl. report. 25 p.

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

Gompper, M.E., R.W. Kays, J.C. Ray, S.D. LaPoint, D.A. Bogan, and J.A. Cryan In Press. A comparison of non-invasive techniques to survey carnivore communities in Northeastern North America. Wildlife Society Bulletin 34:1142-1151.

Hawley, V. D. 1955. The ecology of the marten in Glacier National Park. M.S. Thesis. Montana State Univ., Missoula. 131pp.

Ivan, J. S. 2000. Effectiveness of covered track plates for detecting American marten.

Thesis, University of Montana. 66 p.

Johnson, S.A. 1984. Home range, movements, and habitat use of fishers in Wisconsin. M.S. Thesis, University of Wisconsin, Stevens Point. 78pp.

Kays, R. and K. Slauson. 2008. Remote cameras. In: R. Long, J. Ray, and W. Zielinski (eds) Non-Invasive Survey Techniques for North American Carnivores. Island Press, Covelo, CA.

Klug, R. R. Jr. 1996. Occurrence of Pacific fisher (*Martes pennanti*) in the redwood zone of northern California and the habitat attributes associated with their detections. MS thesis, Humboldt State University, Arcata, CA

Krohn, W. B., C. Hoving, D. Harrison, D. Phillips, and H. Frost. 2004. Martes footloading and snowfall patterns in eastern North America: Implications to broadscale distributions and interactions of mesocarnivores. Pages 115-131 in: D. Harrison, A. Fuller, and G. Proulx (eds). Martens and fishers (Martes) in human altered landscapes: an international perspective. Springer Science+Buisness Media, Inc, New York, NY, USA.

Kucera, T.E. and R.H. Barrett. 1993. The Trailmaster camera system for detecting wildlife. Wildlife Society Bulletin 21:505-208.

Leonard, R.D. 1986. Aspects of the reproduction of the fisher, *Martes pennanti*, in Manitoba. Canadian Field-Naturalist 100:32-44.

Lindstrand, L. 2006. Detections of Pacific fisher around Shasta Lake in northern California. Transactions of the Western Section of the Wildlife Society 42:47-52.

Long, R. A., T. M. Donovan, P. MacKay, W. J. Zielinski, and J. S. Buzas. 2007(b). Comparing scat detection dogs, cameras, and hair snares for surveying carnivores. Journal of Wildlife Management 71:2018-2025.

Long, R. A. and W. J. Zielinski. 2008. Designing Effective Noninvasive Carnivore Surveys. In: R. Long, J. Ray, and W. Zielinski (eds) Non-Invasive Survey Techniques for North American Carnivores. Island Press, Covelo, CA.

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

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

Mackenzie, D.I. and L. R. Bailey. 2004. Assessing the Fit of Site-Occupancy Models Journal of Agricultural, Biological & Environmental Statistics, Volume 9: 300-318.

Mackenzie, D.I. 2005. What are the issues with presence-absence data for wildlife managers? Journal of Wildlife Management 69:849-860.

MacKenzie D.I., J.D. Nichols, J.A. Royle, K.H. Pollock, J.E. Hines, and L.L. Bailey. 2005. Occupancy estimation and modeling: inference methods for patterns and dynamics of species occupancy. Elsevier, San Diego, USA.

MacKenzie D.I., and J.A. Royle. 2005. Designing occupancy studies: general advice and allocating of survey effort. Journal of Applied Ecology 42: 1105-1114.

Mackenzie, D. L. and J. E. Hinze. 2006. Program Presence, version 2.0. http://www.mbr-pwrc.usgs.gov/software/doc/presence/presence.html

Mazzoni, A. K. 2002. Habitat use by fishers (*Martes pennanti*) in the southern Sierra Nevada. Thesis, California State University, Fresno, California, USA.

Powell, R. A. and W. J. Zielinski. 1994. Fisher. The scientific basis for conserving forest carnivores: American marten, fisher, lynx, and wolverine, pp. 38-73. L. F. Ruggiero et al. (eds.). USDA Forest Service General Technical Report RM-254.

Raphael, M. G. 1994. Techniques for monitoring populations of fishers and American martens. Martens, sables, and fishers. S. W. Buskirk, A. S. Harestad, M. G. Raphael and R. A. Powell. Ithaca, NY, Comstock Publishing Associates, Cornell University Press: 224-242.

Ray, J. and W. Zielinski. 2008. Track stations. In: R. Long, J. Ray, and W. Zielinski (eds) Non-Invasive Survey Techniques for North American Carnivores. Island Press, Covelo, CA.

Riddle, A.E., K. L. Pilgrim, L. S. Mills, K. S. McKelvey, L. F. Ruggiero. 2003. Identification of mustelids using mitochondrial DNA and non-invasive sampling. Conservation Genetics 4: 241-243.

Royle, J. A. and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology 84(3): 779-790.

Royle, J. A., T. R. Stanley, and P. M. Lucas. 2008. Statistical modeling and inference from carnivore survey data. In: R. Long, J. Ray, and W. Zielinski (eds) Non-invasive methods for studying North American carnivores. Island Press, Covelo, CA.

SAS 1999. SAS, Version 8.0. SAS Institute Inc., 100 SAS Campus Drive, Cary, NC, 27513-2414.

Schwartz, M. K., T. Ulizio, and B. Jiminez. 2006. U. S. Rocky Mountain Fisher Survey Protocol. USFS Rocky Mountain Research Station 800 E. Beckwith Missoula MT 59801. http://www.fs.fed.us/rm/wildlife/genetics/pdfs/Fisher_Survey_Protocol.pdf, checked 16 April 2007.

Slauson, K. M. and W. J. Zielinski. 2003. Distribution and habitat associations of the Humboldt marten and Pacific Fisher in Redwood National and State Parks. Final Report. USDAFS, Pacific Southwest Research Station, Arcata, CA. 29 p.

Slauson, K. M. and W. J. Zielinski. 2007. Strategic Surveys for *Martes* Populations In Northwestern California: Mendocino National Forest. July-September 2006. USFS, Pacific Southwest Research Station, Arcata. Unpublished Report, 22 p.

Slauson K. M., R. L. Truex, and W. J. Zielinski. In review. Determining the sex of martens at track plate stations. Northwest Science.

Smith, J. B., J. A. Jenks, et al. 2007. Evaluating Detection Probabilities for American Marten in the Black Hills, South Dakota. Journal of Wildlife Management 71(7): 2412-2416.

Tallancy, N. W. 2005. Effects of habitat fragmentation and landscape context on medium-sized mammalian predators in Northeastern National Parks. M. S. thesis, University of Rhode Island.

Truex, R. L., W. J. Zielinski, R. T. Golightly, R. H. Barrett, and S. M. Wisely. 1998. A meta-analysis of regional variation in fisher morphology, demography, and habitat ecology in California. Draft report submitted to California Department of Fish and Game. Arcata, California, USA: USDA Forest Service, Pacific Southwest Research Station.

Truex, R. L. 2003. Sierra Nevada Forest Plan Amendment Adaptive Management Strategy: Forest Carnivore Monitoring Protocol. Unpublished report. USDAFS, Pacific Southwest Station, Placereville, CA. 57 p.

USDI Fish and Wildlife Service. 2004. 12-month finding for a petition to list the west coast distinct population segment of the fisher (*Martes pennanti*). Federal Register 69(68):18770-18792.

Yaeger, J. S. 2005. Habitat at fisher resting sites in the Klamath Province of northern California. Thesis, Humboldt State University, Arcata, California, USA.

Weir, R. D. 1997. Inventory of fishers in the sub-boreal forests of North-Central British Columbia Phase 1: Capture and marking. Peace/Williston Fish and Wildlife Compensation Program, Report No. 130. 32pp plus appendices.

Zielinski, W.J. and T.E. Kucera. 1995. American marten, fisher, lynx and wolverine: survey methods for their detection. General Technical Report PSW-GTR-157. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture. 163pp.

Zielinski, W.J. and H.B. Stauffer. 1996. Monitoring *Martes* populations in California: survey design and power analysis. Ecological Applications 6:1254--1267.

Zielinski, W. J., and S. Mori. 2001. Adaptive management strategy, Sierra Nevada Framework Study Plan: What is the status and change in the geographic distribution and relative abundance of fishers? Redwood Sciences Laboratory, unplublished report. 41p.

Zielinski, W.J, R.L. Truex, F.V. Schlexer, L.A. Campbell and C. Carroll. 2005. Historical and contemporary distributions of carnivores in forest of the Sierra Nevada, California, U.S.A. Journal of Biogeography 32:1385-1407.

Table 1. Compilation of studies using occupancy surveys for fishers in North America. Details of each survey protocol are presented and their associated estimates of the single-visit (*p*) and entire protocol (*P*) detection probabilities. Detection devices include track plates (TP), remote cameras (RC), and hair snares (HS).

Study	Location	Season	Device Type TP	RC	HS	Bait	Lure	# Survey Days	#SUs*	Effort**	1-Visit *p* (SE)	Protocol *P* (SE)
Zielinski and Mori (2001)	California	Summer-Fall	6	--	--	Chicken	Gusto	16	357	96	0.40 (NA)	0.98 (NA)
Campbell (2004)	California	Summer	6	--	--	Chicken	Gusto	16	177	96	0.15 (0.06)	>0.99*** (0.06)
Royle et al. (2008)	California	Summer-Fall	6	--	--	Chicken	Gusto	16	464	96	0.40 (NA)	0.98 (NA)
Tallancy (2005)	NE U.S.	Summer	2	1	2	Chicken	Predator	14	~23	90	0.36 (0.05)	0.89 (0.03)
		Winter	2	1	2	Chicken	Predator	14	~23	90	0.23 (0.06)	0.73 (0.08)
Gompper et al. (2006)	New York	Summer	--	3	--	Chicken, Deer, Beaver	Gusto	30	54	90	0.10 (NA)	>0.90 (NA)
			6	--	--	Chicken, Deer, Beaver	Gusto	11-15	54	~78	0.13 (NA)	>0.80 (NA)
Slauson & Zielinski (2007)	California	Summer	6	--	--	Chicken	Gusto	14	35	84	0.23 (0.07)	0.65 (0.08)

Long et al. (2007)	Vermont	Summer	--	1	--	Chicken	Gusto	14	168	14	NA	0.28 (0.07)

*SU = sample unit. **Effort = (the number of stations * the total survey duration) for each protocol.
*** = p and P calculated using station-specific estimates, yielding estimates that are positively biased.

Table 2. Latency to first detection (LFD), mean of sample units detections, and mean number of station detections for track plate, remote camera, and hair snare sample units that detected fishers. Proportion of survey represents the proportion of the survey duration that had occurred up to the mean LFD. Sample unit abbreviated as SU.

Survey Method	Mean (SE)	95% C.I.	Range	Proportion of Survey
Track Plate				
LFD (days)*	6.9 (0.47)	5.9 to 7.8	2 to 16	43.1%
Detections/Sample Unit	3.1 (0.22)	2.7 to 3.6	1 to 8	
Total Station Det./SU	6.0 (0.75)	4.5 to 7.5	1 to 30	
Remote Camera				
Collins-Baldy Study Area				
LFD (days)	9.9 (1.6)	6.8 to 13.1	2 to 20	35.3%
Detections/Sample Unit	1.3 (0.7)	0 to 2.6	1 to 10	
Total Station Det./SU	2.3 (0.7)	0.94 to 3.6	1 to 11	
Mt. Ashland				
LFD (days)	25.2 (8.7)	7.8 to 42.6	3 to 44	90.0%
Visits/Sample Unit	3.0 (1.0)	0.9 to 5.1	1 to 6	
Total Station Visits/SU	2.5 (2.7)	1.2 to 3.8	1 to 5	
Hair Snare**				
LFD (days)	11.2 (1.6)	8.0 to 14.4	4 to 16	50.9%
Visits/Sample Unit	1.9 (0.5)	1 to 2.8	1 to 5	

*LFD calculations are aggregated across all stations within each sample unit, such that only the first visit to each sample unit is represented.
**Total Station Visits not applicable, as each sample unit consisted of a single station.

Table 3. Number of sample units surveyed and mean number of inoperable station check intervals by season for track plate sample units.

Season*	Sample Units	Mean Inoperable Intervals (SE)	Proportion of Survey that is Inoperable
Fall	194	4.49 (0.38)	9.4%
Spring	41	4.32 (0.67)	9.0%
Summer	278	5.46 (0.39)	11.4%
Winter	13	1.00 (0.45)	2.1%

*Fall = Oct-Nov, Spring = Mar-May, Summer = Jun-Sep, Winter = Dec-Feb.

Table 4. Descriptive detection statistics for sample units where fishers were detected in the cores versus on the peripheries of their distribution in California and southwestern Oregon.

	Edge (n = 31)	Core (n = 52)
LFD (days)		
Mean	6.35	7.27
SE	0.63	0.62
# Fisher Detections		
Mean	2.97	3.17
SE	0.34	0.63

Table 5. Expected counts of detection results (0,1) as proportions of the observed counts for two models estimating occupancy from systematic track plate surveys conducted in California and Oregon. For each visit transition the four cells a 2x2 frequency table represent the ratios of expected to observed counts; see the first row for example of the transitions that each cell represents. The observed diagonal counts ("00" and "11" transitions) for the 2-group, visit-specific *p* model are all lower than the expected (ratios <1), suggesting dependence. Addition of a simple Markov persistence factor to the model accounts for this dependence and brings the ratios much closer to 1.00.

Transition	**2-group Survey-specific p**		**2-group, Visit-specific *p*, Simple Markov Persistence Factor**	
Visit t to t+1	00	10	00	10
	01	11	01	11
Visit 1 to 2	0.96	1.1	1	1.01
	1.33	0.71	1.02	0.99
Visit 2 to 3	0.96	1.16	1.01	0.95
	1.1	0.84	0.94	1.12
Visit 3 to 4	0.95	1.16	0.97	1.07
	1.38	0.84	1.21	0.96
Visit 4 to 5	0.97	1.18	1	0.99
	1.12	0.9	1.04	0.99
Visit 5 to 6	0.96	1.06	0.99	1.03
	1.34	0.92	1.07	0.97
Visit 6 to 7	0.96	1.2	1	0.96
	1.07	0.94	1	1.01
Visit 7 to 8	0.94	1.23	1	1.07
	1.17	0.87	0.93	0.99

Table 6. Models describing detection probability of fishers at track plate sample units in California, ranked according to ΔAIC_c value. K represents the number of parameters in a model.

Model #	Model Description	AIC_c	ΔAIC_c	*w*	K
1	2 Seasons, Visit-specific *p* 1 Markov-chain	1244.5	0.00	0.89	11
2	2 Groups, 1 Markov-chain	1284.9	4.48	0.09	3
3	2 Groups, 2 Markov-chain	1253.8	9.42	0.01	4
4	2 Seasons , linear in logits	1254.9	10.50	0.00	3
5	2 Groups, Visit-specific *p*	1263.3	18.89	0.00	10
6	2 Groups, Constant *p*	1299.6	55.21	0.00	3
7	Visit number, 2 Seasons	1301.4	57.01	0.00	3
8	2 Seasons , Visit-specific *p*	1305.3	60.85	0.00	10
9	1 Groups, Visit-specific *p*	1323.3	78.89	0.00	9
10	2 Seasons	1325.4	80.99	0.00	2
11	Visit number	1327.5	83.08	0.00	2
12	Null	1351.1	106.66	0.00	1

Table 7. Normalized importance weights for all track plate detection probability variables.

Variable	
1-Markov chain persistence factor	0.60
Visit-specific *p*	0.27
2 Seasons	0.22
2 Groups	0.03
2-Markov chain persistence factors	0.01
Linear change in logits	0.00
Constant *p*	0.00
Visit #	0.00
1 Group	0.00

Table 8. Probability of detection (p), estimate of occupancy (ψ), and p from the top ranked model for fishers using track plates in California.

Visit #	Season Specific p (SE) Non-Summer	Season Specific p (SE) Summer	Cumulative P Non-Summer	Cumulative P Summer	ψ (SE)
1	0.181 (0.05)	0.092 (0.04)	0.181	0.092	0.207 (0.02)
2	0.476 (0.07)	0.121 (0.04)	0.570	0.203	
3	0.286 (0.06)	0.160 (0.06)	0.694	0.331	
4	0.447 (0.07)	0.157 (0.05)	0.831	0.436	
5	0.412 (0.07)	0.093 (0.04)	0.901	0.489	
6	0.615 (0.07)	0.271 (0.07)	0.962	0.627	
7	0.463 (0.07)	0.219 (0.06)	0.978	0.709	
8	0.583 (0.07)	0.171 (0.06)	0.991	0.758	

Table 9. Akaike's information criteria results for models describing remote camera detection probability for fishers in the Collins-Baldy and Mt. Ashland study areas in California.

Collins-Baldy

Model #	Model Description	AIC_c	ΔAIC_c	AIC_c wt*	#Parameters
1	1 Group, Markov-chain	164	0.0	1.00	3
2	1 Group, Constant p	193	29.6	0.00	3
3	1 Group, Visit-specific p	1953	1789.6	0.00	29
4	2 Groups, Visit -specific p	7072	6908.0	0.00	30
5	2 Groups, Visit-specific p Markov-Chain1	7285	7121.5	0.00	31

Ashland

Model #	Model Description	AIC_c	ΔAIC_c	AIC_c wt*	#Parameters
1	1 Group, Markov-chain	104	0.0	1.00	3
2	1 Group, Constant p	134	30.4	0.00	3
3	1 Group, Visit -specific p	7284	7179.8	0.00	59
4	2 Groups, Visit-specific p	28374	28270.4	0.00	60
5	2 Groups, Markov-Chain1	28848	28744.5	0.00	4
6	2 Groups, Markov-Chain1	29330	29226.6	0.00	4

Table 10. Probability of detection (*p*), estimate of occupancy (ψ), and *P* from the top ranked models for fishers using track plates and remote cameras in California and hair snares in Idaho. Estimates for single visit *p* are presented as ranges when visit-specific *p* was in the top model and as a single estimate when constant *p* was in the top model. Survey days to achieve desired P are effort required per sample unit.

Study Area	Single Visit *p* (SE)	ψ (SE)	Value of *P* for Protocol* (SE)	Survey Days for *P* ≥80%	≥ 95%
Track Plate		0.207 (0.02)			
Summer	0.09-0.27 (0.04-0.07)		0.607 (0.057)	30	58
Non-Summer	0.18-0.62 (0.04-0.07)		0.956 (0.003)	10	16
Remote Camera					
Collins-Baldy	0.181 (0.06)	0.889 (0.12)	0.930 (0.007)	17	32
Mt. Ashland	0.043 (0.02)	0.274 (0.13)	0.671 (0.155)	41	71
Hair Snare					
	0.141 (0.04)	0.534 (0.14)	0.813 (0.094)	22	38

*Track plate protocol: 6-stations, 8-visits, surveyed for 16 consecutive days; Remote Camera protocol: 2-stations, 1-8 visits, surveyed for 2-58 consecutive days; Hair snare: 1 station, 11 visits, surveyed for 22 consecutive days.

Table 11. Akaike's information criteria results for models describing hair snare detection probability for fishers in the Locsha corridor study area in Idaho.

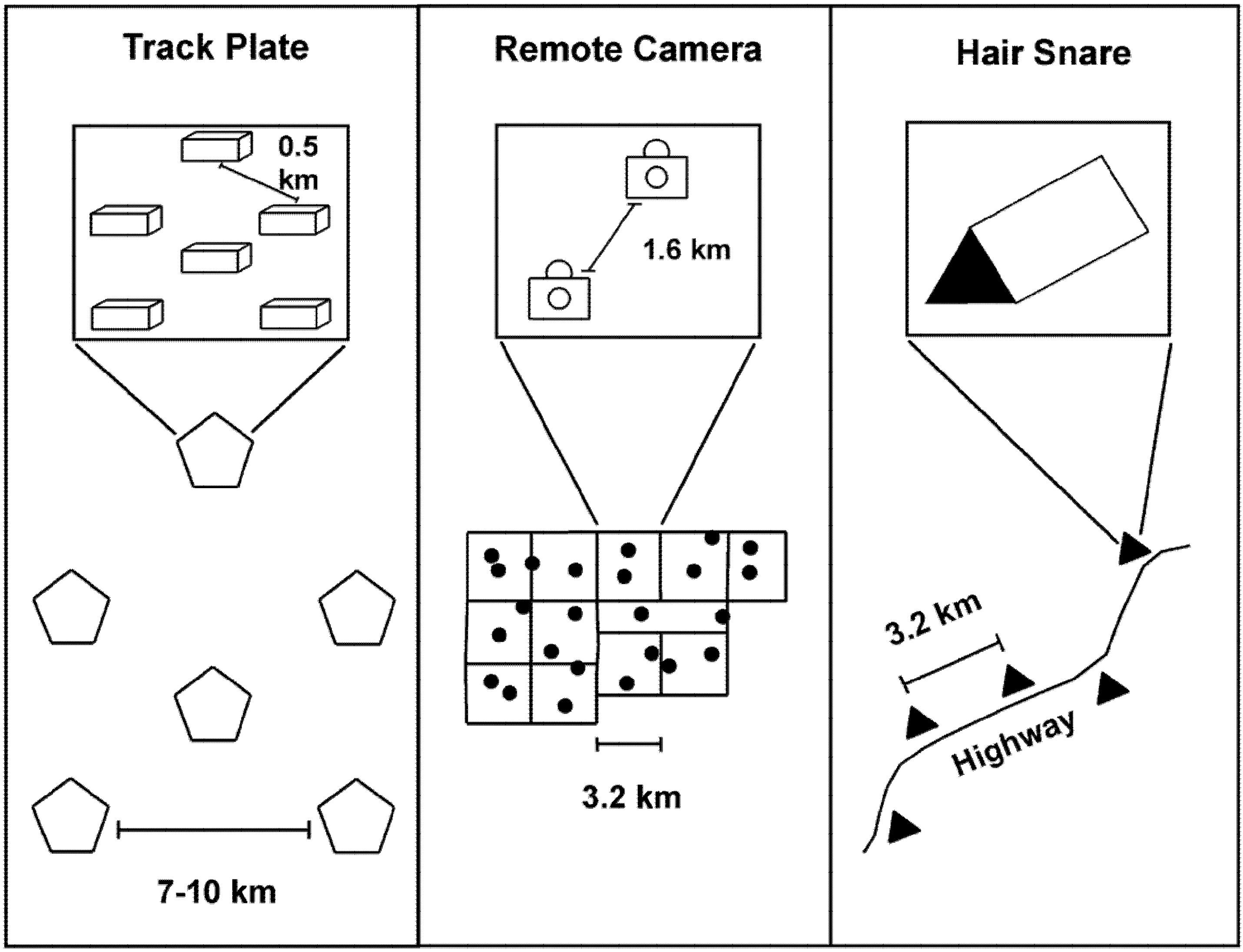
Model #	Model Description	AIC_c	ΔAIC_c	AIC_c wt*	#Parameters
1	1-group, constant *p*	133.7	0.0	0.95387	3
2	1-group, Markov-chain	139.7	6.0	0.04612	4
3	1-group, Visit-specific *p*	159.2	25.5	0.00000	13
4	2-group, Visit -specific *p*	1336.0	1202.2	0.00000	14
5	2-group, Markov-chain-1	1441.2	1307.4	0.00000	5
6	2-group, Markov-chain-2	1543.5	1409.7	0.00000	6

023387

Table 12. Number of functional survey days required for track plate, remote camera, and hair snare protocols to achieve 4 levels of detection probability. A functional survey day is defined a 24-hour period of continuous operability for a single detection device (e.g., track plate, remote camera, or hair snare station).

	Detection Probability							
	Non-Summer				Summer			
Survey Protocol	0.80	0.85	0.90	0.95	0.80	0.85	0.90	0.95
Track Plate: 6-Station	10	12	14	16	26	31	38	48
Camera: 2-Station								
Collins-Baldy	17	21	26	32	No Data			
Ashland	41	49	59	71	No Data			
Combined	27	33	41	51	No Data			
Hair Snare: 1-Station	22	26	30	38	No Data			

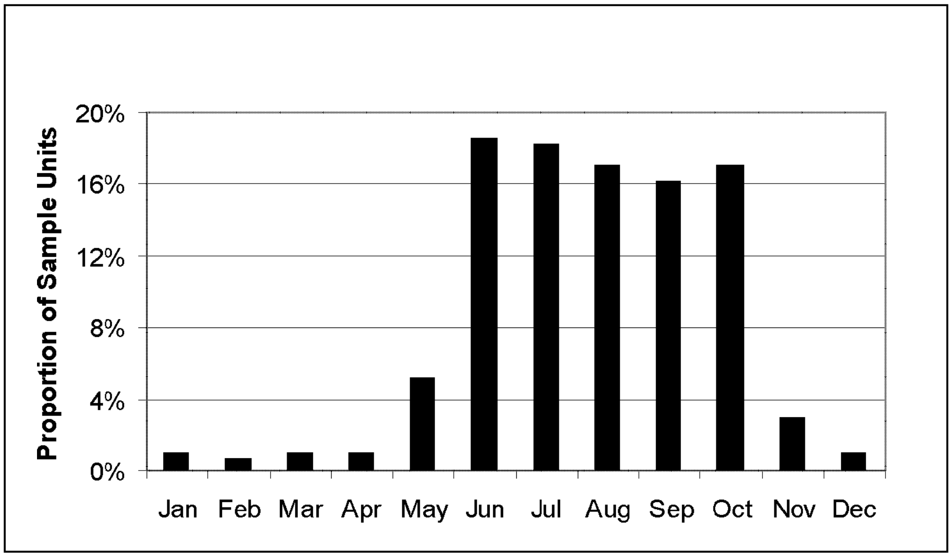
Figure 1. Schematic of the 3 different survey protocols using detection devices (track plate, remote camera, and hair snares) to detect fishers in the western United States. Each sub-figure illustrates the composition of each device sample unit (upper) and the distribution of sample units in each respective study area (lower).



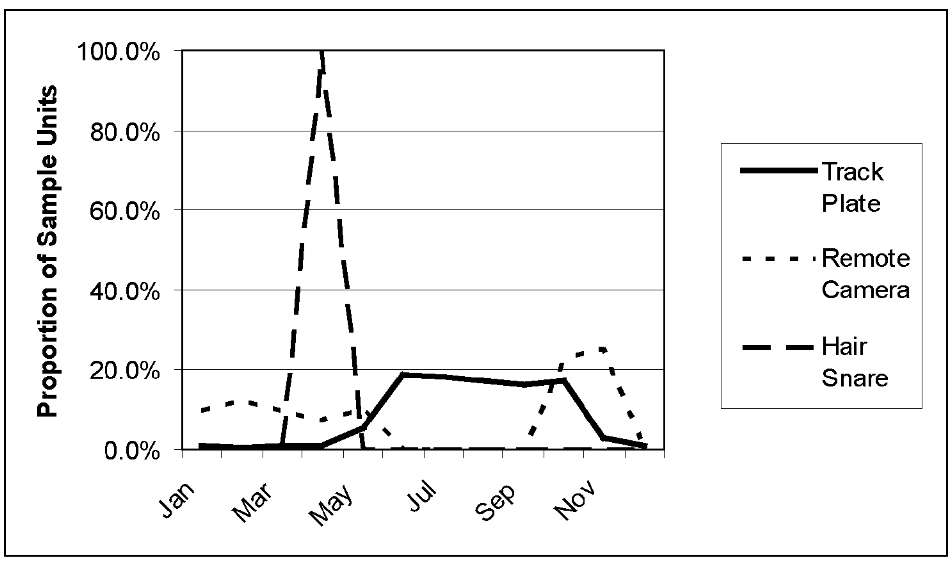
023389

Figure 2. Distribution of sampling effort by method during the months of the year.

2a. Proportion of track plate sample units surveyed by month.



2b. Proportions of sample units surveyed by month for each method.



023390

Figure 3. Monthly variation in mean latency to first detection (days), sample unit detections, and total station detections/sample unit for fishers at track plates sample units during the 6 months when the majority (>90%) of all surveys occurred.

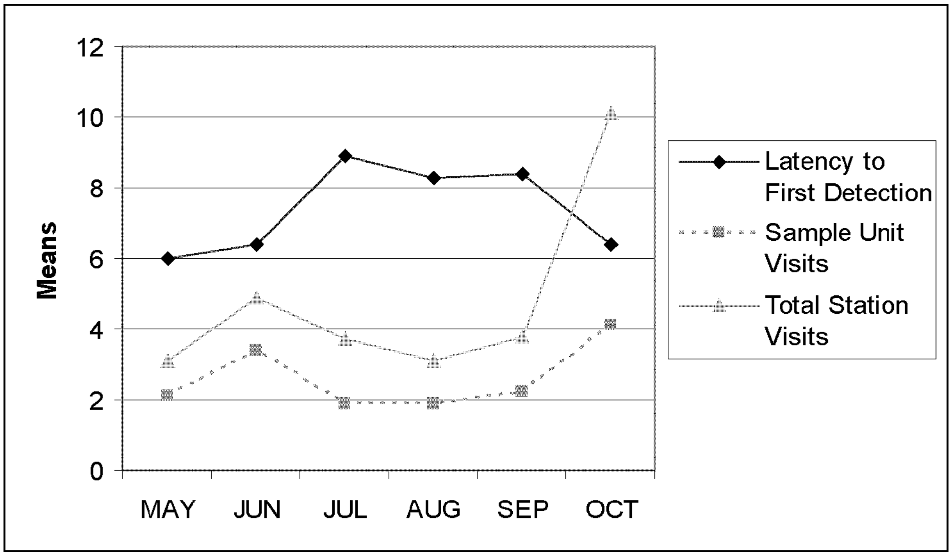
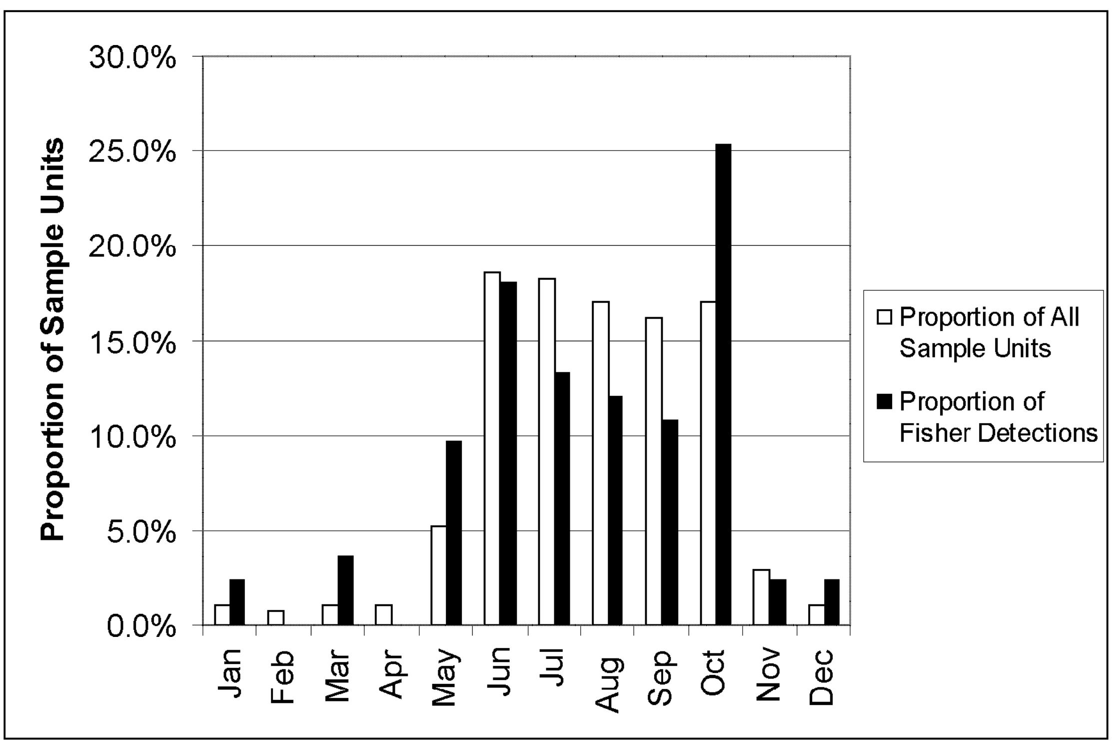
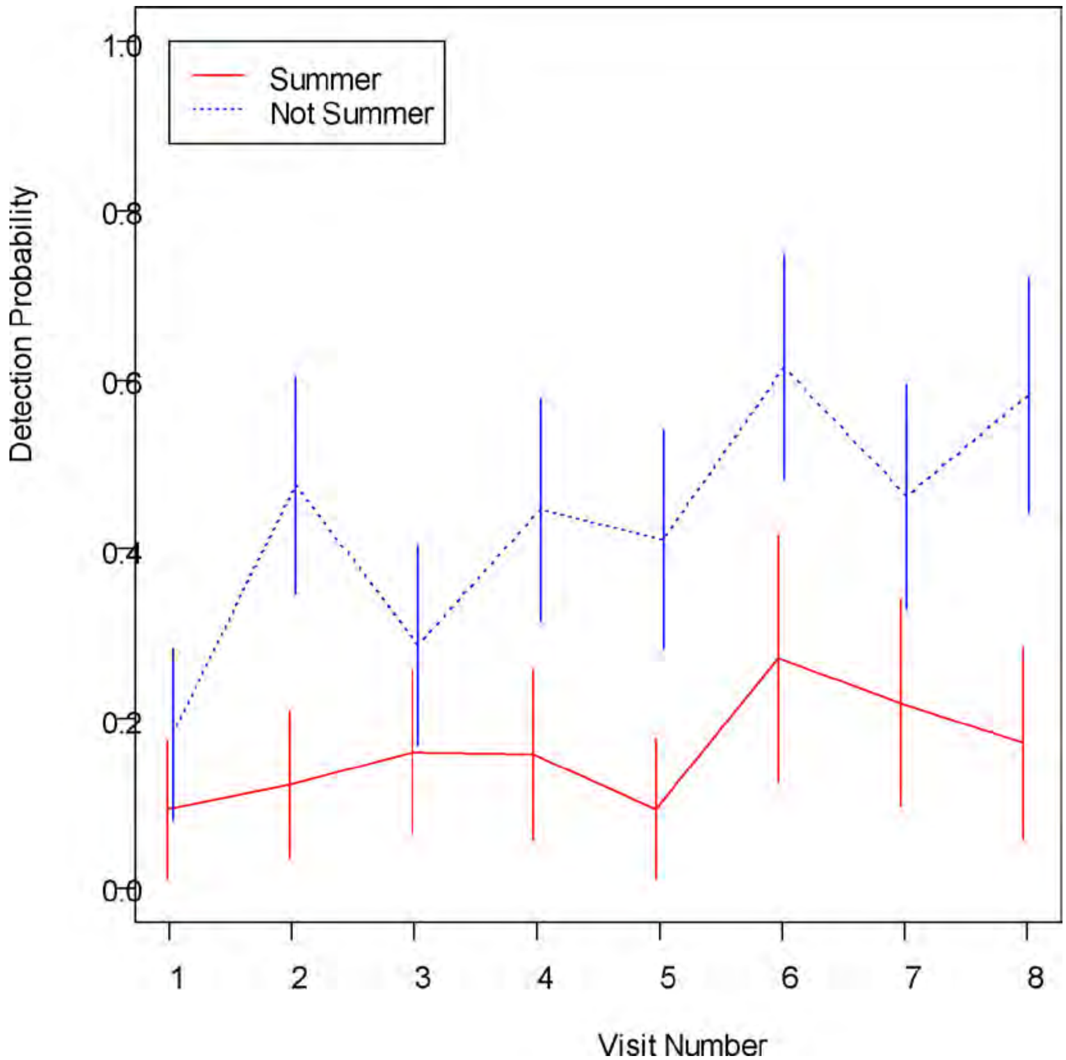


Figure 4. Proportions of systematic track plate sample units (n = 576) that detected fishers (n = 83) by month.



023391

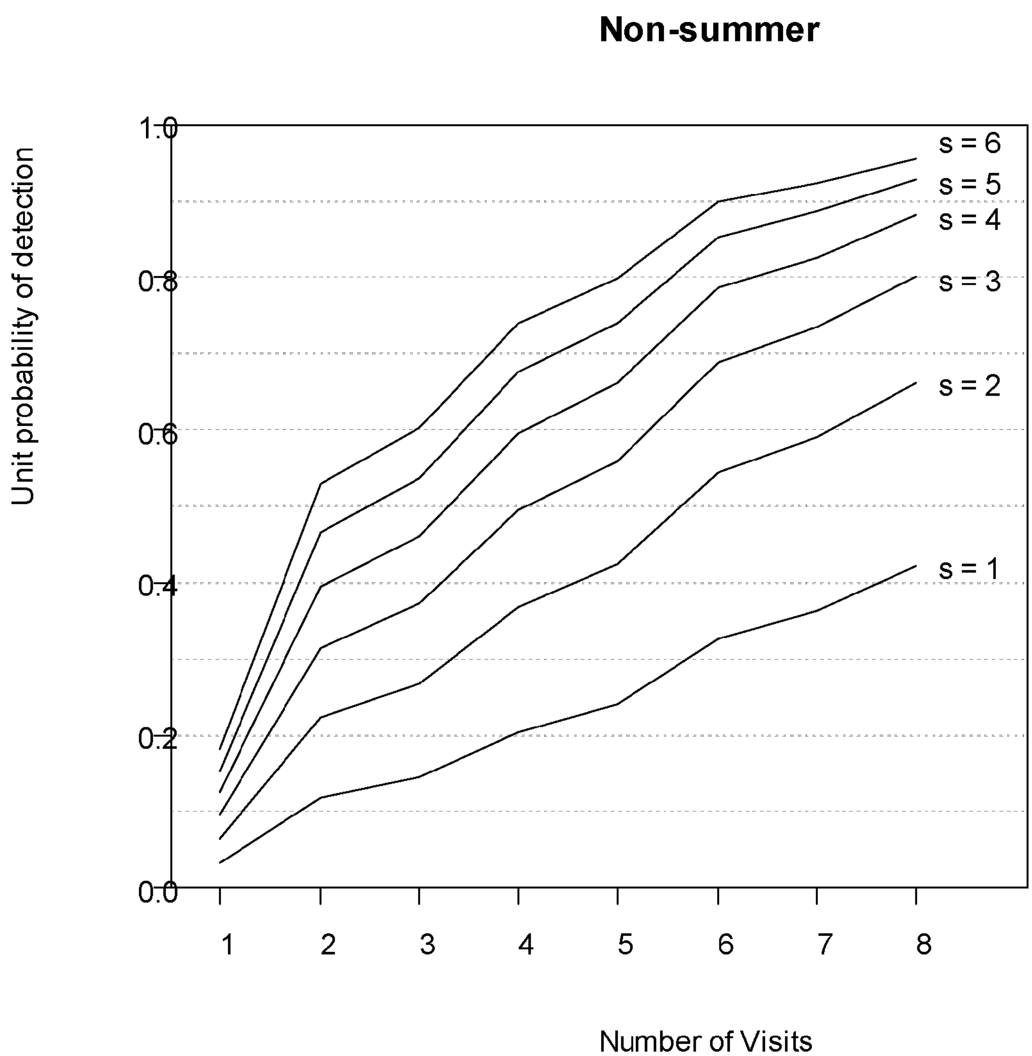
Figure 5. Approximate 95% confidence intervals for each of the survey specific detection probabilities in the two seasons. Summer includes data from the months of July-September and Non-summer includes data from October-June.



023392

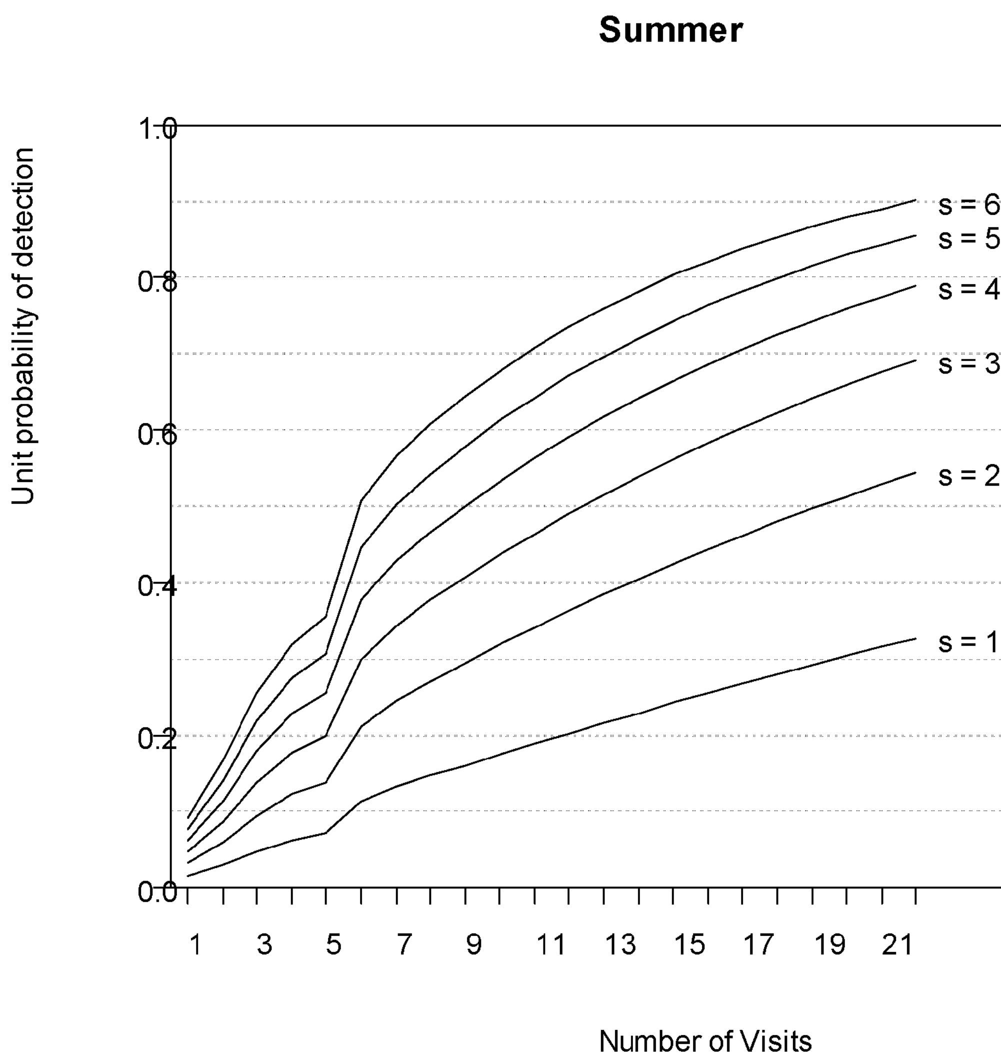
Figure 6. Probability of detection for track plate sample units with all combinations of stations (s; 1-6) and 2-day visits (1-8, non-summer; 1-22, summer) for surveys conducted during the non-summer (6a) and summer (6b) seasons. The top model was used to estimate all coefficients. Note that the probability of detection for an individual visit was estimated with the estimate for the 8-th visit for all visits numbers greater than 8.

6a.



023393

6b.


Summer
Unit probability of detection
0.0
0.2
0.4
0.6
0.8
1.0
1
3
5
7
9
11
13
15
17
19
21
Number of Visits
s = 6
s = 5
s = 4
s = 3
s = 2
s = 1

023394

Figure 7. Effect of station inoperability on different combination of survey effort.

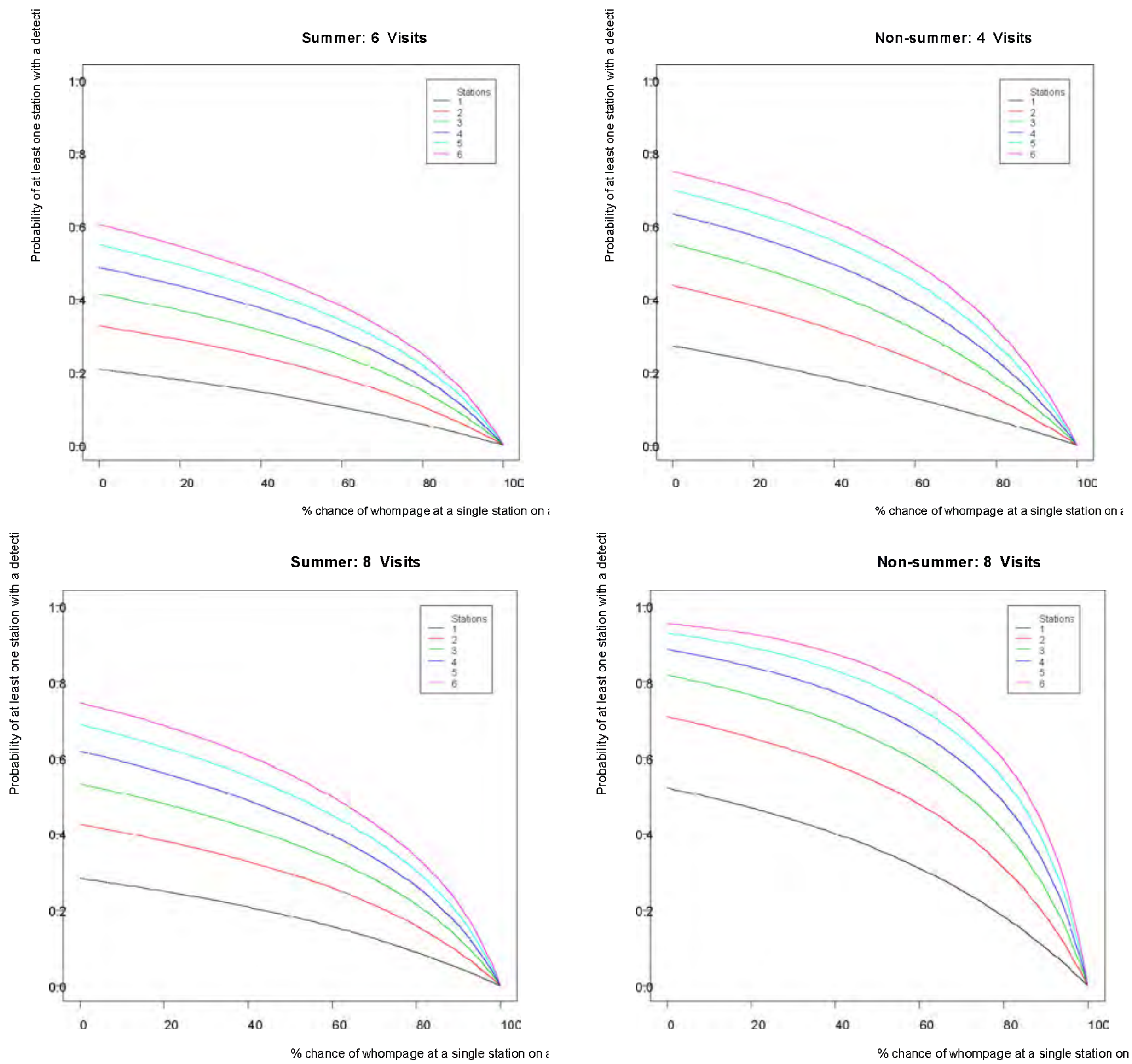
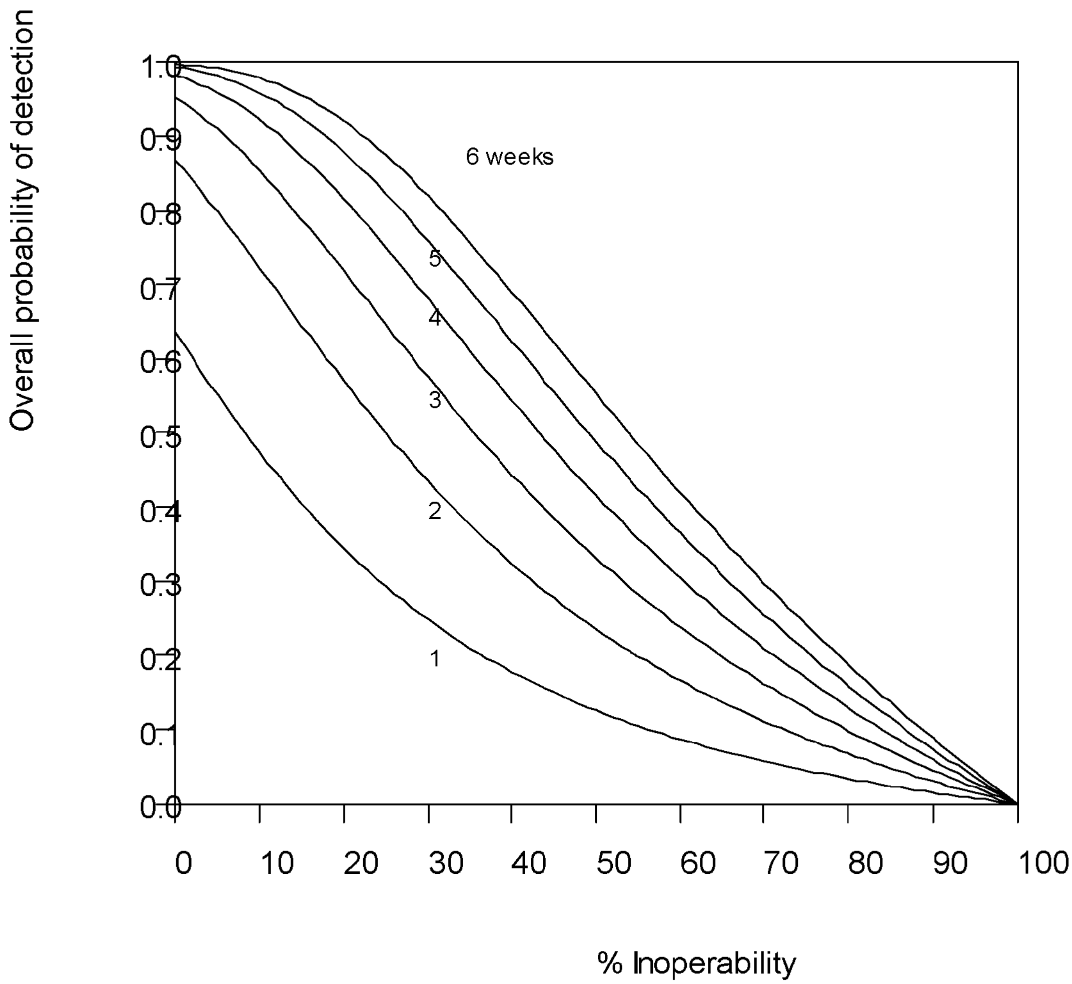
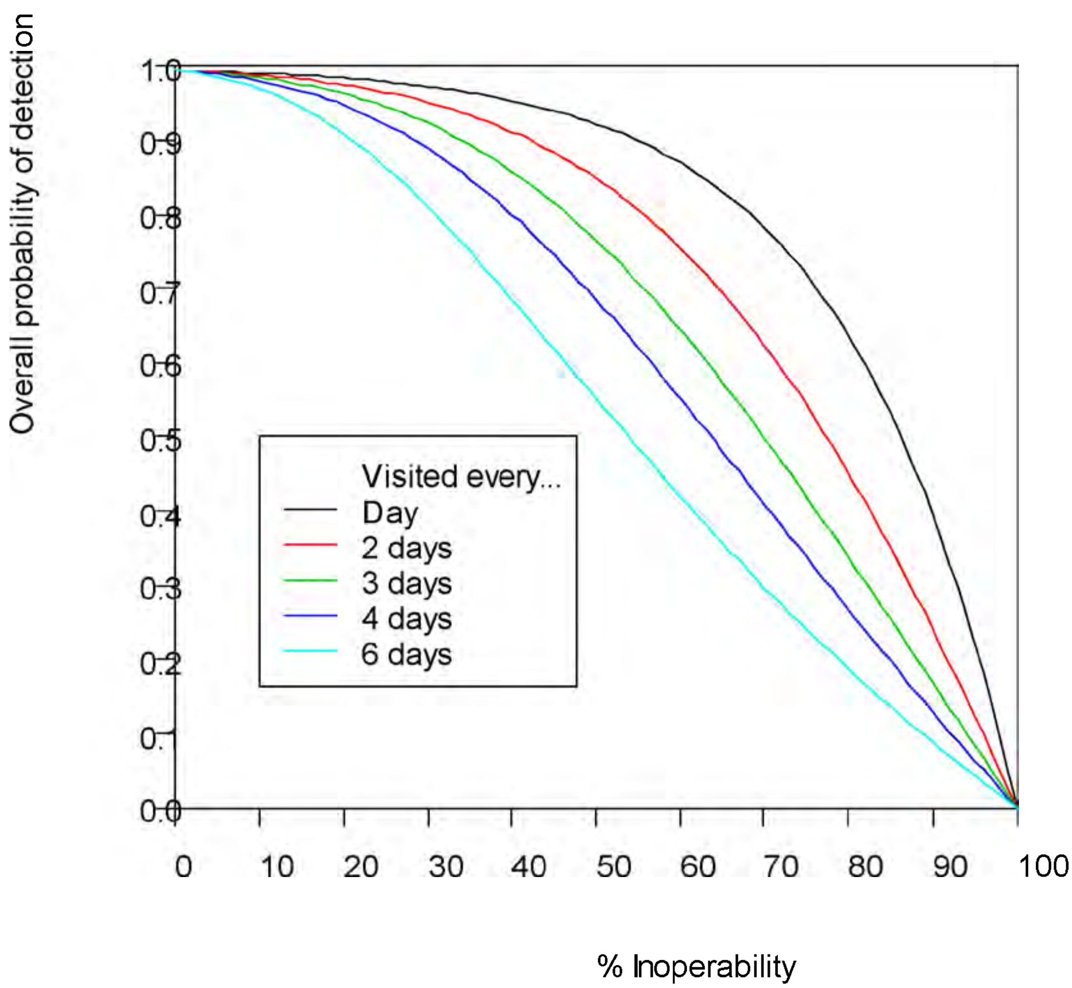


Figure 8. Effect of station inoperability on different combination of survey effort for remote cameras surveys.



023396

Figure 9. The relationship of changes in visit frequency to the effect that inoperability has on the sample unit probability of detection (*P)* for remote camera sample units. The survey duration simulated here is for 72 days of survey effort.



Appendix 1. This glossary defines terms that are frequently used throughout the manuscript. Definitions are tailored specifically to noninvasive survey methods, and may not represent the best definitions for other applications. Words in italics are defined elsewhere in the glossary. Adapted from Long and Zielinski (2008).

DESIGN—A term describing the overall layout of the *survey*, including the size and location of the *survey sites*, the number and arrangement of *devices* at *sites*, the *survey duration*, and the *protocol* used to detect species during each *sampling occasion*.

DETECTION—Evidence (i.e., track, hair, scat or photo) that confirms the occurrence of a target species at a *station* or *site* during a *sampling occasion*.

DETECTION HETEROGENEITY—Variation in detection histories caused by factors unrelated to the survey *design*. The consequences of not accounting for detection heterogeneity are biased parameter estimates and thus inferences from the detection histories.

DETECTION HISTORY—A string of 1s or 0s (e.g., 01001) that represents either the pattern of species' *detections-nondetections* at a *site* or the history of detections of an individual at a *station*, over a series of *sampling occasions*. Encounter histories are used in both *occupancy* estimation and *capture-recapture* approaches for estimating *population status*.

DETECTION PROBABILITY (or PROBABILITY OF DETECTION)—A parameter representing the probability of actually detecting a species or individual animal given that it (or its sign) is present at the survey *site*. Detectability can be estimated given an appropriate survey *design*, and is important for accurately estimating both *occupancy* and *abundance* via capture-recapture methods.

DEVICE—The specific entity (e.g., enclosed track plate, remote camera) that registers a *detection*. By this definition, a scat detection dog is considered a device. Some *methods* (e.g., natural sign surveys) do not use devices, but instead entail detections recorded by an observer.

DEVICE INOPERABILITY—Any factor that renders the detection device unable to record a detection. Typical causes of device inoperability include bear damaging stations, device malfunction (e.g., loss of power, flash not turned on in camera units), and snowfall that buries devices or alters the camera angle.

DISTRIBUTION—The actual area of species occurrence, typically expressed on a map as either "occupied" or "unoccupied" (or estimated to be in one of these states). Distribution can be displayed as a continuous surface (e.g., as with vector-based map elements) or as a surface divided into subunits (e.g., grid cells)—each indicating *presence* or (inferred) absence. In either case, a given location's state ("1" or "0") can be inferred based on binary detection-nondetection *surveys* at the actual location, or predicted via occurrence models in concert with a rule-based assignment method (e.g., all sites with predicted occupancy of >0.80 are assumed to be occupied).

FUNCTIONAL SURVEY DAY—Defined as a 24-hour period of continuous operability for a single detection device (e.g., track plate, remote camera, or hair snare station).

LATENCY TO FIRST DETECTION (LFD)—The number of survey days elapsed until the first detection of a fisher occurs at a sample unit.

MONITORING—Performing repeat *surveys* over time, with the goal of quantifying change in *population status* (i.e., *trends*). Monitoring should not be confused with repeat *sampling occasions* (sometimes referred to as "checks" or "visits"), which are conducted within a single *survey* and either allow the estimation of detection probabilities or provide an increase in overall *detectability* at the site.

OCCASION—See *Sampling Occasion*.

OCCUPANCY—A population state variable representing the proportion of *sites* estimated to be occupied (or in the case of wide-ranging species such as carnivores, the proportion used) by a species of interest. If an appropriate survey *design* (e.g., randomly chosen *sites*) is employed, occupancy is also considered an estimate of the proportion of the *survey area* occupied (or used) by the species. Occupancy is not estimated for an individual *site*, but only for *surveys* with multiple sites. Thus, it is differentiated from *presence* in that it is an estimated parameter whose value falls between 0 and 1.

OCCURRENCE—Typically a synonym for *occupancy*, occurrence is also synonymous with *presence*—as in the phrase *extent of occurrence*.

p—Small p s the single-visit probability of detecting a fisher at a sample unit, given there is at least one present in the sample unit.

P—Big P is the probability of detecting a fisher at a sample unit for the entire survey protocol, given there is at least one present in the sample unit.

ψ —'Psi' is known both as the proportion of sample units that are occupied by a fisher or the probability of occupancy of a sample unit.

PRESENCE—The state of a *sample unit* being occupied by a species (or in some cases, by the species' sign)—regardless of whether the species (or sign) is detected by a *survey*. Presence is evaluated at a single *sample unit* (e.g., at a *site*), and differs from *occupancy* or *occurrence*, which are typically parameters that can take on any value between 0 and 1 and are estimated with *detection-nondetection data* from a series of sample units.

PRE-PROJECT SURVEY—A *survey* aimed at the *detecting* or characterizing the *population status* of at least one member of a target species, and executed prior to the implementation of a natural resource management activity.

PROBABILITY OF DETECTION—See *Detection Probability*.

PROBABILITY OF OCCUPANCY/ PROPORTION OF SITES OCCUPIED—See ψ.

PROTOCOL—The specific actions taken to carry out a *sampling occasion*. Protocol differs from *design* in that the latter refers to the larger-scale considerations of *survey* layout, whereas protocol refers to detailed instructions that are repeated at the *sample unit*. Examples of protocols include baiting a track station with a single chicken leg, or testing a remote camera prior to departing a *station*.

SAMPLE UNIT—A statistical unit of analysis. For example, if a *site* comprising five survey *stations* is the *sample unit*, then a detection at any number of the stations results in a single detection recorded for the site. Elsewhere, the term sample unit has sometimes been used synonymously with site, and also in a nonprobability sense to refer to the subunits of a *survey* aimed at detecting a target species in an area of interest. We do not use it in this manner.

SAMPLING—Recording or measuring characteristics of a portion of the sample population in order to infer something about the entire population.

SAMPLING OCCASION—One sampling effort at a *site* or at a *station*. A sampling occasion (sometimes referred to as a "*visit*" or a "check") occurs when an observer returns to a site or survey station to record whether or not a *detection* has been registered at a detection *device*, or alternately when a non-station based approach is used in a single attempt to detect a species. Similarly, the recorded sampling history of devices capable of remotely tracking the dates of detections (e.g., remote cameras that imprint dates on photographs) can be divided into any number of sampling occasions between researcher visits. Sampling occasions can be repeated to increase the *probability of detection*, and the results of multiple sampling occasions can be used to estimate overall *detectability* if they are assumed to be independent (i.e., a detection during one occasion does not influence a detection during another occasion). The results of multiple sampling occasions are represented as an *encounter history*.

SITE—The statistical unit of analysis when using *binary data* or sign counts to assess *abundance*. The site is the area within which binary *detection-nondetection data* or *count data* are pooled for analysis. For instance, a single *sampling occasion* at a site may result in a *detection*, a nondetection, or a count (e.g., number of track stations with a detection, number of scats located). Sites are typically selected based on a statistical probability model and in many cases are considered independent, although this may not be the case for some applications.

STATION—A location within a *site* or *survey area* at which a *detection* attempt is made during a *sampling occasion*. Stations are typically assumed to be dependent (i.e., a detection at one station may affect detection at other stations within the site or survey area), and detections at multiple *stations* are often collapsed into *binary data* or *count data* at the level of the site for occupancy or relative abundance assessments. Alternately, stations are the locations at which individual animals are detected for *capture-recapture* approaches used to estimate population size.

SURVEY— One or more attempts (i.e., *sampling occasions*) to detect a species at either a single location or across many *sites* with the intention of making inferences about species *occurrence* or population size. Survey outcomes can include assessments of species *presence*, estimates of *occupancy*, predicted *distributions*, mean count per unit area or per survey time, or estimated population size.

SURVEY AREA—The area within which the *survey* results and resulting inferences are relevant. This is analogous to the statistical population. Survey *sites* should be distributed appropriately within the survey area and based on a statistical probability model if inferences gained from site data are to be unbiased.

SURVEY DURATION—The amount of time or the number of *sampling occasions* comprising the *survey*. The sampling duration affects *detection probability*, and should be chosen based on both the home range of the target species and the size of the survey *site*.

VARIANCE INFLATION FACTOR—A value used to adjust estimated variances when the samples used violate assumptions of sample independence.

VISIT—A synonym for *sampling occasion* that is often encountered in the carnivore literature. Visit is becoming a less accurate descriptor with the advent of noninvasive methods that permit *sampling durations* to be subjectively parsed after the fact (e.g., when remote cameras provide continuous sampling), and because a visit by the observer can be confused with a "visit" by the target species.

Appendix 2. Calculating the detection probability at a sample unit when number of visits, number of stations, and the amount of inoperability varies.

Introduction

After fitting a model which accounts for the probability of presence and the detection probability structures for a varying number of visits and a varying number of stations, we want to calculate the probability of detection for a specific number of stations and a specific number of visits.

If we label the status of visit *i* to station *j* as X_{ij} where $X_{ij} = 1$ if a detection was observed and $X_{ij} = 0$ if a detection was not observed, then the probability of detection at the unit (with *v* visits and *s* stations) can be written as follows:

$$\text{Probability of detection at unit} = 1 - \Pr(X_{11} = X_{12} = \cdots = X_{vs} = 0)$$

Single station model

Consider a unit with a single station. We then just use a single subscript which denotes the visit: $X_1, X_2, \ldots, X_v$. We then write the probability of detection at the unit as

$$1 - \Pr(X_1 = X_2 = \cdots = X_v = 0)$$

We can calculate the rightmost term as follows:

$$\Pr(X_1 = X_2 = \cdots = X_v = 0) = \Pr(X_1 = 0)\Pr(X_2 = 0 \mid X_1 = 0)\cdots\Pr(X_v = 0 \mid X_1 = 0, \ldots, X_{v-1} = 0)$$

The top model used a Markov chain structure which was characterized with probabilities of detection for each visit ($p_1, p_2, \ldots, p_v$) and a persistence factor ϕ that accounted for the dependence of status between visits. Specifically the probability of both X_i and X_{i+1} being 1 (*i.e.*, both visits having detections) is given by

$$\Pr(X_i = 1, X_{i+1} = 1) = p_i p_{i+1} + \phi(\min(p_i, p_{i+1}) - p_i p_{i+1})$$

when $\phi \geq 0$ (which represents positive persistence). (This structure can be interpreted as a special case of the extinction/colonization structure given by MacKenzie et. al. 2005) We also find from this probability structure that

$$\Pr(X_i = 0, X_{i+1} = 0) = (1 - p_i)(1 - p_{i+1}) + \phi(\min(p_i, p_{i+1}) - p_i p_{i+1})$$

Therefore the conditional probability of $X_{i+1} = 0$ given that $X_i = 0$ is

$$\Pr(X_{i+1}=0 \mid X_i=0)=((1-p_i)(1-p_{i+1})+\phi(\min(p_i,p_{i+1})-p_ip_{i+1}))/(1-p_i)$$

(We're almost there…) The probability of at least one detection at the unit is therefore given by the following:

$$Q_v = \Pr(\text{at least one visit with a detection}) =$$
$$1-(1-p_1)\prod_{i=1}^{v-1}((1-p_i)(1-p_{i+1})+\phi(\min(p_i,p_{i+1})-p_ip_{i+1}))/(1-p_i)$$

If the logit of p_i is a function of the visit number such as

$$\operatorname{logit}(p_i)=\log\left(\frac{p_i}{1-p_i}\right)=a+b\cdot i$$

then we can replace p_i with $1-1/(1+\exp(a+b\cdot i))$ and predict for any number of visits (as long as we don't extrapolate too far from the number of visits in the sample that were used to estimate the model parameters).

For ease of calculation with minimal and standard spreadsheet manipulations, note that we can calculate Q_v using the calculations that already went into Q_{v-1}:

$$Q_v = 1-(1-Q_{v-1})((1-p_{v-1})(1-p_v)+\phi(\min(p_{v-1},p_v)-p_{v-1}p_v))/(1-p_{v-1})$$

for $v>1$.

Also note that when $\phi=0$, we have a model of independence and

$$Q_v = 1-\prod_{i=1}^{v}(1-p_i)$$

Multiple station model

If stations can be considered independent, then the probability of at least one station out of *s* stations having a detection (at a unit with presence) during a single visit is given by

$$\Pr(\text{at least one station with a detection during visit } i)=p_i^{(s)}=1-(1-p_i^{(1)})^s$$

where $p_i^{(1)}$ is the probability of a detection at a single station during visit *i*. Then we plug in the values of $p_i=p_i^{(s)}$ in the previous equations to determine the probability of detection at a unit with *s* stations.

Now because the visit probabilities were estimated with 6 stations, we actually get an estimate of $p_i^{(6)}$ and need to determine $p_i^{(1)}$ so that we can make predictions for other numbers of stations. We use the equation $p_i^{(1)} = 1-(1-p_i^{(6)})^{1/6}$.