

## POWER OF SIGN SURVEYS TO MONITOR POPULATION TRENDS<sup>1</sup>

KATHERINE C. KENDALL

Science Center, Glacier National Park, West Glacier, Montana 59936 USA

LEE H. METZGAR

Division of Biological Sciences, University of Montana, Missoula, Montana 59812 USA

DAVID A. PATTERSON AND BRIAN M. STEELE

Department of Mathematical Sciences, University of Montana, Missoula, Montana 59812 USA

**Abstract.** The urgent need for an effective monitoring scheme for grizzly bear (*Ursus arctos*) populations led us to investigate the effort required to detect changes in populations of low-density dispersed animals, using sign (mainly scats and tracks) they leave on trails. We surveyed trails in Glacier National Park for bear tracks and scats during five consecutive years. Using these data, we modeled the occurrence of bear sign on trails, then estimated the power of various sampling schemes. Specifically, we explored the power of bear sign surveys to detect a 20% decline in sign occurrence. Realistic sampling schemes appear feasible if the density of sign is high enough, and we provide guidelines for designs with adequate replication to monitor long-term trends of dispersed populations using sign occurrences on trails.

**Key words:** black bear; Glacier National Park, Montana; grizzly bear; monitor; population; scat; survey; tracks; trend; *Ursus americanus*; *Ursus arctos*.

### INTRODUCTION

Our ability to monitor the status of secretive, threatened/endangered species cheaply and nonintrusively remains an elusive goal, in part because we rarely know the design requirements of possible survey techniques. Here, we examine the effort required to detect changes in populations of low-density, dispersed animals using sign they leave on trails. Our experience with grizzly bears (*Ursus arctos*), efforts to monitor their populations, and the difficulties of measuring their numbers inspired this project. We propose that our results possess broader applicability.

In their present, remnant range within the lower 48 states, grizzly bears occupy remote country at low density, wander widely, defy reliable observation, hold "threatened" status, and are hunted in some areas. The pressing need to monitor population trends of these bears contrasts with the formidable difficulties of doing so. Methods currently employed to monitor grizzly bear population trends involve either intensive radio-telemetry studies or the opportunistic observation of females with cubs (U.S. Fish and Wildlife Service 1990). The former is expensive and requires handling a large proportion of the population. The power of the latter is unknown and is highly sensitive to unquantified effort.

Because of these difficulties, workers have recognized the desirability of using sign surveys to monitor bear populations (Pelton 1972, Lindzey et al. 1977, Johnson and Pelton 1980). Unfortunately, the lack of

information about the power of such surveys bars the effective design of monitoring schemes.

In this report, we used data from Park bears to examine the ability of alternative survey designs to detect a change in sign deposition. We do not investigate the biology or population status of the bears of Glacier National Park. We do characterize our bear sign data and then use simulation to explore the power of various sampling schemes. Our results reveal the sample size (trail lengths and numbers) required to achieve a desired level of power when the data resemble those gathered for Glacier National Park bears and when the amount of sign deposited has decreased by a set amount. Generally, our guidelines apply to such data without regard to species or location, though our specific results may not.

### FIELD METHODS AND RESULTS

#### *Field methods*

Trails surveyed for bear sign were located throughout Glacier National Park (407 820 ha) in northwest Montana. The Park straddles the continental divide and, in this rugged, mountainous terrain, elevations range from 948 to 3190 m. Survey routes west of the divide were primarily forested, while those to the east were more open. Topography and vegetation generally encouraged use of trails by both animals and humans.

Surveys were conducted during late September and early October by three groups with three to five observers per group. The inexperienced observers received four hours of classroom and four to eight hours of field instruction on bear sign recognition. Training

<sup>1</sup> Manuscript received 8 July 1991; revised and accepted 31 January 1992.

TABLE 1. Bear sign survey summary statistics for mean number of sign per 1.6-km trail segment and proportion of segments with sign in Glacier National Park, 1984–1988. Values in parentheses are computed with an outlier count of 18 scats in one segment deleted.

	Year					Total
	1984	1985	1986	1987	1988	
No. trails	20	19	20	20	17	
No. segments	150	131	143	150	139	
No. segments per trail:						
Mean	7.5	6.9	7.1	7.5	8.2	
SD	4.3	5.6	5.4	5.2	4.8	
Range	1–17	1–21	1–21	1–21	3–21	
No. scats per trail segment:						
Mean	0.49	0.47	0.59 (0.47)	0.24	0.36	0.43 (0.40)
SD	0.85	0.94	1.68 (0.81)	0.54	0.76	1.03 (0.79)
Range	0–4	0–4	0–18 (0–4)	0–3	0–4	0–18 (0–4)
Proportion of trail segments with scat	0.30	0.27	0.34	0.20	0.25	0.27
No. grizzly bear tracks per segment:						
Mean	0.08	0.11	0.07	0.04	0.03	0.06
SD	0.32	0.33	0.28	0.23	0.21	0.27
Range	0–2	0–1	0–2	0–2	0–2	0–2
Proportion of segments with grizzly bear tracks	0.07	0.11	0.06	0.03	0.02	0.06
No. black bear tracks per segment:						
Mean	0.06	0.08	0.06	0.03	0.03	0.05
SD	0.29	0.30	0.27	0.27	0.17	0.26
Range	0–2	0–2	0–2	0–3	0–1	0–3
Proportion of segments with black bear tracks	0.05	0.08	0.06	0.02	0.03	0.04

emphasized species recognition for sightings and tracks, identification of digs, stripped trees, and other feeding sign, and evidence of marking behavior such as rubbing trees and bear trails (routes where bears have worn depressions or bare spots on the ground). While all types of bear sign were recorded, we encountered sightings and sign from feeding and marking behaviors so rarely that our analyses used only scats and tracks.

Trail selection was based on geographical representation and, in part, on logistical considerations. To cover as much ground as possible, we preferred long sections of trail that did not require backtracking. Of the  $\approx 325$  km (202 miles) surveyed on 27 trails, most were repeated each year, but some year-to-year variation could not be avoided. Detailed information on trails surveyed is available from the Science Center, Glacier National Park.

Observers recorded all bear scat encountered, and, because we could not distinguish between grizzly and black (*U. americanus*) bear scat, counts combined the two. This does not limit the utility of our investigation because, in this case, we analyze the power of sampling schemes rather than attempt to monitor bear populations. For the latter, scat identification is essential.

Observers recorded only front foot bear tracks since it is difficult to identify bear species from hind foot tracks. For each species of bear, we recorded separate track entries when:

1) the length or width of the front pad differed by  $> 1$  cm from a track previously recorded on that trail segment, or

2) tracks of similar size indicated travel in different directions, or

3) a track was seen at least 1.6 km (1 mile) from the last tracks of the same size.

For our analyses and subsequent recommendations, trail “segments” comprise the basic sampling units, and we converted field data to counts of sign on each segment. Based on our knowledge of bear behavior, we selected 1.6-km trail segments as appropriate for our study. We considered this length long enough to provide reasonably independent observations on adjacent segments and short enough to prevent unreasonable reduction of sample size.

#### Field results

We surveyed trails each autumn, 1984–1988. Scats were the most abundant type of bear sign encountered, with average counts of 0.43 scats per segment per year and 0.27 of the trail segments with at least one scat. Tracks were much rarer than scats, with means of 0.06 and 0.05 grizzly and black bear tracks per segment, respectively (Table 1). Except for their slightly lower abundance, black bear track counts and distributions resembled those of grizzly tracks and were not considered further.

## ANALYTICAL METHODS AND RESULTS

In this analysis, we first identified a distribution characteristic of our bear data. Second, we drew a hypothetical "sample" from that distribution to represent the results of a simulated monitoring effort. Next, we modified the distribution from which we drew samples to represent a population that leaves less sign (presumably having decreased) and drew a second sample. Finally, we applied a statistical test to our two simulated samples.

We repeated the process of drawing and testing pairs of samples many times for a specific combination of trail lengths and numbers. Knowing that our samples came from different populations, the frequency at which the test rejects the null hypothesis (no difference between the sampled populations) measures the power of that sampling scheme (combination of trail lengths and numbers) and statistical test.

We used a single change in amount of sign (20%) between the first and second samples in all simulations. We consider this a realistic amount of change for a manager to desire to detect: large enough to reflect a population change of concern and small enough to permit response prior to catastrophe. We also noted that smaller changes become increasingly difficult to detect with realistic samples sizes.

*Data treatment*

To test the hypothesis of no change in density of sign, we explored the merits of using both mean number of sign per trail segment and the proportion of segments with sign. We selected the latter for the following reasons:

- 1) Because these two statistics were closely correlated for all years (Table 1), absence/presence treatment discards little information included in mean counts.
- 2) Locally high scat depositions can arise when a single bear spends an extended period on a segment, and we considered it inappropriate to inflate the test statistic with such results. This behavior probably accounted for some of the extreme aggregation observed in our scat counts.
- 3) Absence/presence data are robust to differences in observer interpretation of sign occurrences.

Modeling is simplified if sign depositions on segments within a trail can be considered independent. We found no evidence of between-segment dependence in our data. We conducted runs tests for sequence randomness on the 0,1 data (0 = absence, 1 = presence) within each trail and year that had at least two segments with and two without scat (42 tests). We did not use track data here because they rarely met these criteria. Individually, the runs tests have low power for short trails. Therefore, we examined the distribution of the *P* values for all 42 tests; under independence, this distribution is uniform over the interval (0,1) and we

found no evidence to the contrary (chi-square goodness-of-fit test:  $\chi^2 = 3.7$ ,  $df = 7$ ,  $P = .81$ ). Even when attention was restricted to longer trails, the *P* values showed no significant departure from uniformity. We also used the same data to compute correlations between counts from adjacent segments for each trail. These coefficients were approximately symmetrically distributed around zero, again giving no evidence of dependence between segments. The assumption of independence between segments is used only in the simulation model and not in the statistical tests on the simulated data. Minor departures from this assumption in actual data should not significantly affect our results.

We desired to pool the data from all years to compute parameters for our simulation model on the assumption that there had been no major change in bear density over this 5-yr period. We performed two analyses of between-year differences, a loglinear model and Wilcoxon signed-ranks tests on all pairs of years. The difference between only one pair of years was significant and only on one of the tests. Therefore, we pooled the data from all years to compute model parameters. Including this variation in our parameter computations makes our results conservative.

Some of the observed variation between years may be due to reasons other than changes in sign density, such as short-term changes in bear behavior, food availability, or sign detectability due to weather. Averaging the sign counts from each trail segment over several years of effort evens out such differences.

Before fitting the model we had to establish the amount of between-year dependence of sign occurrence on individual trails. To investigate this we computed the proportion of each trail's segments that had scat for each pair of consecutive years. The resulting scatterplot of these pairs of proportions revealed no strong correlation (Pearson's  $r = 0.20$ ), indicating little dependence. However, because the model proved sensitive to this dependence and because we expected some data sets to show more correlation, we used two extreme levels of dependence in our simulations.

*Fitting the model*

We used the pooled absence/presence data from all years to estimate parameters for a beta-binomial model. In this model, each trail has a constant probability of sign occurrence on all segments. These probabilities vary among trails and follow a beta distribution. Within each trail, the presence or absence of sign on each segment is an independent Bernoulli event; therefore, the total number of segments with sign on a trail is a binomial variable. The beta distribution is a highly flexible, two-parameter model that can create variously shaped probability densities over the interval (0,1). It has been used to model a variety of data (e.g., Griffiths 1973, Paul 1979).

For a given year, let  $Y_{ij}$  be a 0,1 variable representing

absence/presence of sign on segment  $j$  of trail  $i$ ,  $i = 1, \dots, k$  and  $j = 1, \dots, n_i$ , where  $n_i$  is the number of segments in trail  $i$ . The beta-binomial model assumes the following:

1) The probability of sign on any segment of trail  $i$  is  $p_i$  where  $p_1, \dots, p_k$  are independent beta random variates. That is,  $p_i$  is a random variable with probability density function

$$f(p; a, b) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1 - p)^{b-1},$$

$$0 < p < 1, \tag{1}$$

where  $a >$  and  $b > 0$  are unknown parameters, and  $\Gamma(\cdot)$  is the gamma function.

2) The segment responses  $Y_{ij}$ ,  $j = 1, \dots, n_i$ , on the trail  $i$  are independent. Therefore, conditional on the value of  $p_i$ , the total number of segments with sign on trail  $i$ , say  $Z_i$ , is a binomial random variable with parameters  $n_i$  and  $p_i$ ; that is,

$$\Pr[Z_i = z \mid p_i] = \binom{n_i}{z} p_i^z (1 - p_i)^{n_i - z},$$

$$z = 0, 1, \dots, n_i. \tag{2}$$

It follows from Eqs. 1 and 2 that the marginal (unconditional) distribution of  $Z_i$  is beta-binomial:

$$g(z; a, b, n_i) = \Pr[Z_i = z]$$

$$= \binom{n_i}{z} \frac{\Gamma(z + a)\Gamma(n_i - z + b)\Gamma(a + b)}{\Gamma(n_i + a + b)\Gamma(a)\Gamma(b)},$$

$$z = 0, 1, \dots, n_i. \tag{3}$$

Given observations  $z_1, \dots, z_k$  of the number of segments with sign on trails of lengths  $n_1, \dots, n_k$ , respectively, maximum likelihood estimators of the unknown parameters  $a$  and  $b$  are obtained by maximizing the likelihood function

$$L(a, b) = \prod_{i=1}^k g(z_i; a, b, n_i) \tag{4}$$

with respect to  $a$  and  $b$ . We used a Newton-Raphson algorithm programmed in GAUSS (1988). We estimated  $a$  and  $b$  for both scat and grizzly tracks by pooling the data for all five years and considering them a sample of 96 separate trails (Table 2). The values of  $a$  and  $b$  reflect both the average rate of sign incidence on the trails and the variability of this rate across trails. The goodness-of-fit of each model was measured in the following manner. Eq. 3 was used to compute the probability distribution under the fitted model of the number of sign for each trail in the sample (the distribution depends on the length of the trail). These probabilities were summed over all trails and multiplied by the number of trails to give expected counts for the number of segments with sign. We considered these fits satisfactory and used the beta-binomial distribution to model bear sign on trail segments.

TABLE 2. Goodness-of-fit and parameter estimates for beta-binomial model fit to bear scat and grizzly bear track data. The top portion presents expected and observed numbers of trails on which there were 0, 1, 2, . . . segments with sign.

No. trail segments with sign	Bear scat		Grizzly bear track	
	Observed	Expected	Observed	Expected
0	23	25.0	70	69.5
1	22	23.8	18	18.0
2	23	16.5	4	5.5
3	10	11.1	2	1.8
4	7	7.2	≥ 2	≥ 1.1
5	4	4.6		
6	3	2.9		
7	4	1.8		
8	≥ 0	≥ 3.1		
Goodness-of-fit statistics				
$N$	96		96	
$\chi^2$	8.8		1.2	
df	6		2	
$P$	0.2		0.5	
Beta-binomial maximum likelihood estimates				
$a$	2.34		0.93	
$b$	5.92		15.84	
Mean	0.28		0.06	
SD	0.15		0.05	

*Simulation procedures*

The beta-binomial model, using the parameter estimates from all the data, was used to simulate absence/presence data for two different years for various combinations of numbers and lengths of trails. Three different aspects of the simulation were varied: the values of  $a$  and  $b$  in the beta distribution, the degree of dependence between counts on the same trail for the two years, and the amount of change in the bear sign. We used two pairs of values for  $a$  and  $b$ : one from the scat data, representing a relatively high probability of encountering sign, and one from the grizzly track data, representing a lower frequency of "hits." We used two degrees of between-year dependence: none and complete. We also included two levels of change in sign abundance between samples: none and a 20% decrease. We examined the "no change" situation only to evaluate our statistical tests.

Since the management implications of a Type I error (incorrectly declaring a decrease when none occurred) may be less severe than a Type II error (failure to detect a decrease), results for two relatively high significance levels, 1-sided  $\alpha = .1$  and  $.2$ , are reported here.

When simulating the sign occurrence for the 20% decrease in abundance, we reduced the probability of a hit uniformly across all trails. In nature, population declines (and lower sign deposition) may occur unevenly. In such cases, power will be lower than indicated by our results.

During the simulation for each of the eight combinations (scats, tracks; no dependence, complete dependence; no change, 20% decrease), we performed

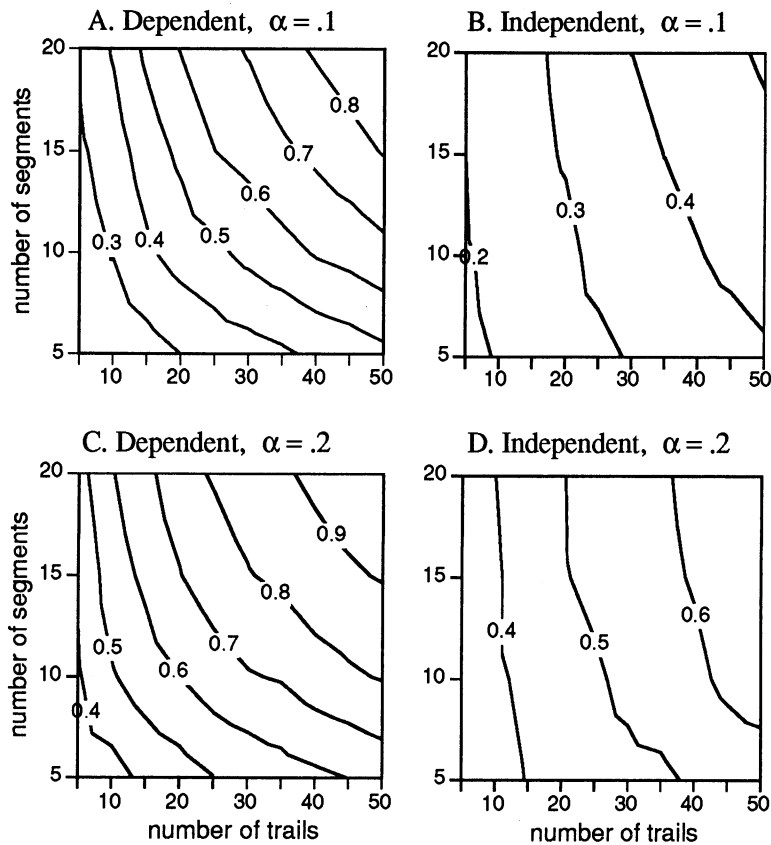


FIG. 1. Power isopleths for detecting a 20% decrease in bear scat density for the one-sided Wilcoxon signed-ranks test. Contours reveal the sample size (number of trails and 1.6-km trail segments) required to achieve a given probability of correctly rejecting the null hypothesis of no difference between years. Contour interval is 0.1 scats/trail.

1000 iterations for various combinations of trail number and length. For simplicity, all  $k$  trails had the same fixed length  $n$ , though this was not a necessity. The details of the simulation procedure appear in the Appendix.

The number of trails was varied from 5 to 50 in increments of 5 and the number of segments on each trail from 5 to 20, also in increments of 5. After completing the simulations for one combination of parameters, we estimated isopleths of equal power by linear interpolation across the grid consisting of number of trails on one axis and number of segments on the other.

We also examined the effect on power of averaging several years of data. To do this, we simulated three years of data at constant initial density, then simulated another three years of data at decreased density. We used the average of each set of three years of data in our analysis.

Both the Wilcoxon matched-pairs signed-ranks test (Daniel 1990) and the paired  $t$  test were computed in the simulations. These were applied to the between-year differences in the trail sums for the 0,1 absence/presence data. These tests do not assume that segments within trails are independent, an assumption made in the simulation but possibly violated in field data.

#### Simulation results

Simulations with no change in sign abundance showed that both the Wilcoxon signed-ranks test and the paired  $t$  test generally achieved the desired significance levels under the null hypothesis when the number of trails was 20 or more. Most of the observed rejection rates in these cases fell within .01 of the desired significance level. When the number of trails was  $<20$ , there was a tendency to exceed the desired significance levels, particularly with our low-density track counts, although most of the observed rejection rates still fell within .03 of the desired levels. While the estimated power levels for small samples may be somewhat inflated, these deviations were rare, and too small to affect our results or conclusions. The estimated powers of the two tests to detect a 20% decrease were very similar in most cases. Therefore, we report and discuss only the results for the Wilcoxon signed-rank test.

Higher correlation (dependence) of sign abundance between years produced tests with greater power (Figs. 1 and 2: A, C vs. B, D). Over the range of sample sizes tested (5–20 segments on 5–50 trails), the more abundant data (Fig. 1) achieved 50–90% power at the largest sample sizes. The less abundant data (Fig. 2) achieved

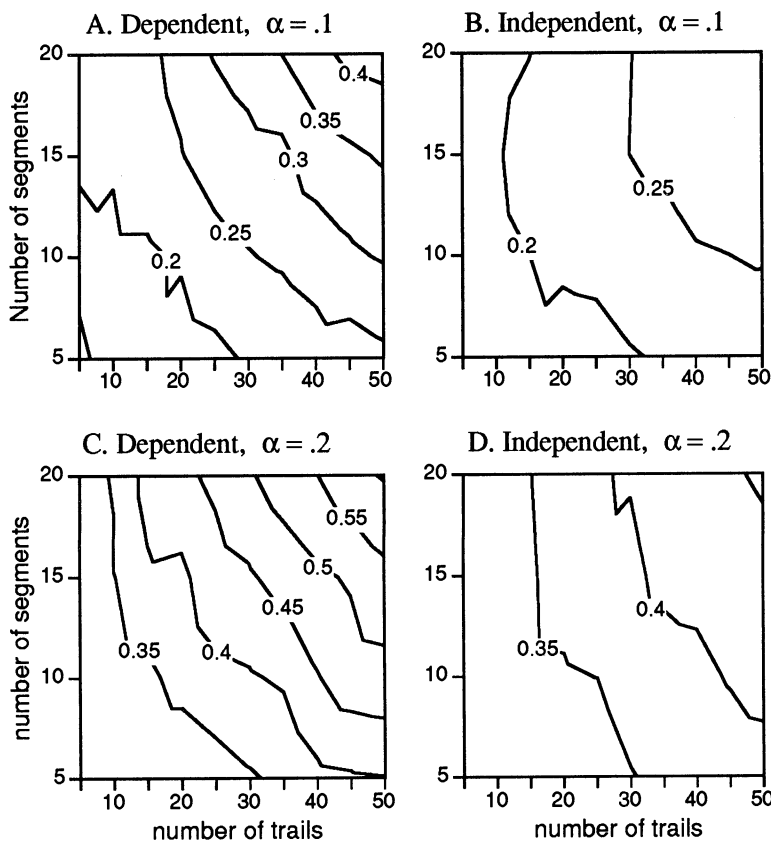


FIG. 2. Power isopleths for detecting a 20% decrease in grizzly bear track density. Levels of  $\alpha$  and dependence vary as in Fig. 1. Contour interval is 0.05 scats/trail.

≤60% power, even when 1000 segments (50 trails, 20 segments per trail) were surveyed (Fig. 2D).

Where the power isopleths appear nearly vertical (e.g., the left portion of Fig. 1D), little power is gained by increasing trail lengths (adding segments). Under independence (which most closely matches our data), increasing the total number of segments by increasing the number of trails is generally more effective than increasing the number of segments through more segments per trail (Table 3). For instance, Table 3 reveals that distributing 100 total segments as 20 trails with 5 segments per trail produces 1.24 (=0.42/0.34) times the power of using 5 trails with 20 segments per trail. This effect is not seen under complete dependence, where different allocations between number of trails and number of segments per trail produces little effect. Note how slowly power increases with increased effort (Table 3), particularly under independence. For example, under independence, a 16-fold increase in effort (from 50 to 800 segments) only doubled the power.

Averaging three consecutive years of data increased power by ≈30–50% (Table 3, 3-yr vs. 1-yr with both independence and dependence) in the high-density situation. Even with this increase, at least 200 segments must be used to obtain power of 0.70 under independence.

We also examined the effect of changing the variability between trails without changing the mean probability of sign. We repeated the simulations with the parameters  $a$  and  $b$  of the beta distribution doubled

TABLE 3. Estimated power of one-sided Wilcoxon signed-ranks test to detect 20% decrease in bear scat density for alternative sampling schemes. Results are shown for a 1-yr and a 3-yr sampling period under both the assumption of independence between years and complete dependence between years, with  $\alpha = .2$ .

Total no. of segments	No. of trails	No. of segments per trail	Independence		Dependence	
			1 yr	3 yr	1 yr	3 yr
50	5	10	.33	.41	.39	.55
	10	5	.36	.49	.36	.57
100	5	20	.34	.45	.46	.68
	10	10	.38	.56	.49	.70
	20	5	.42	.58	.46	.69
200	10	20	.40	.60	.59	.85
	20	10	.45	.68	.60	.85
	40	5	.51	.72	.58	.87
400	20	20	.49	.74	.76	.96
	40	10	.58	.84	.75	.97
800	40	20	.62	.89	.92	1.00
	50	16	.67	.92	.92	1.00

and halved from the original values, which approximately halves and doubles the variance, respectively. Under complete dependence, these differences produced little effect on power. However, under independence, power declined, sometimes substantially, with increasing variability.

#### DISCUSSION

Biologists who use sign surveys to monitor population trends enjoy a technique with relative low expense and ease of replication. While pellet-group and track counts are widely used indices of ungulate abundance, the success of such surveys is qualified (Dzieciolowski 1976, Kie 1988). Pellet-group abundance proved a successful index of penned white-tailed deer (*Odocoileus virginianus*) numbers (Eberhardt and Van Etten 1956) but failed to predict the size of a free-ranging deer population (Fuller 1991). The ability of track counts to measure population trends in deer depends, in part, on sign density and survey design. Among others, Mooty et al. (1984) and Tucker (1991) explored power and sample size requirements and found track surveys a feasible monitoring method.

With bears, several workers successfully used scat surveys to monitor the distribution and timing of activity (Klein 1959, Pelton 1972, Smith 1978). Unfortunately, and for a variety of reasons, sign counts may poorly measure their population changes (Spencer 1955, Edwards and Green 1959, Klein 1959). Carlock et al. (1983) compared black bear population estimates with indices of abundance and found that abundances of scats and fresh marks on "bear trees" correlated poorly with population trends.

Others have evaluated the power of population trend estimates. Gerrodette (1987) conducted an analysis of linear regression as a tool to detect trends over several years. Use of the analysis depends on the appropriateness of his parametric models. Harris (1986) analyzed trends derived from variable counts and showed that precision demands multiple counts within years and/or long monitoring periods (>12 yr). We found with Page's nonparametric test for monotone trend (Daniel 1990) that long monitoring periods (around 10 yr) were required to obtain more power than Wilcoxon's signed-ranks tests (on the first and last year) when there was a constant exponential decrease from the first year to the last. In addition, Page's test was substantially less powerful when the decrease was not monotone. Therefore, we believe using data from only the first and last year is a reasonable and robust procedure and will be a good indicator of power even for trend tests.

Despite the difficulties, sign surveys to monitor bear populations remain attractive and we desperately need an effective monitoring scheme for grizzly bear populations in the lower 48 states. That urgency led us to investigate the feasibility of a geographically extensive, reasonably priced, experimental design that would pos-

sess adequate power to monitor the population status of this threatened species.

To a large extent, our results conform with those of earlier studies. Sign surveys based on data like ours will not detect small, annual population fluctuations but may reveal long-term trends or impending disaster earlier and with less expense than other available tools. We emphasize that, at best, such data will reliably detect only substantial, potentially threatening declines, and then only with large sample sizes, relatively abundant sign, and the annoyance of false alarms.

For programs that monitor species of special concern, managers should lead the decision making process that sets the desired power, the  $\alpha$  level, and the maximum acceptable, undetected decline. The notorious difficulties of monitoring grizzly bears led us to select values that promote statistical power for our simulations (20% decline and relatively large  $\alpha$  values). We caution that these values may be inappropriate in many other cases and urge their careful consideration.

We anticipate that managers will request at least 80% power from any monitoring program for threatened or endangered species. Note that, with this power, they would reliably detect the specified decline only 80% of the time when such a decline actually occurred. Twenty percent of the declines could progress further until detected by subsequent monitoring efforts. Also, note that adopting  $\alpha = .2$  ensures that, in the absence of real change in the sign deposition rate, the statistical test will falsely signal a non-existent decline 20% of the time. If managers wish to detect either an increase or a decrease, the appropriate statistical tests is two-sided and the  $\alpha$  levels given here should be doubled.

Reasonable professionals might well consider the large declines simulated here (20%) to be catastrophic; they certainly should stimulate urgent concern. However, the large sample sizes required to detect even this level of change suggest that we are not likely to design geographically extensive, realistically priced monitoring programs of much greater sensitivity.

We offer the following suggestions for designing realistic sampling schemes that maximize power. In general, the best sample schemes will maximize the number of trails, even at the expense of trail length. This will help to boost power and better conform with the assumption that changes in sign deposition rate are uniform within a trail. It also will allow geographically extensive monitoring.

Power of the sample to detect change can be increased in several other ways. Pooling data from multiple years improves power and reduces the influence of random year-to-year fluctuations in both trail use by animals and the detectability of sign. Pooling data assumes negligible changes in population density during the period of the pool. Alternatively, power could be increased by making within-year replicate surveys, since the mean of several independent measurements

will have lower variance than the measurements themselves.

Greater sign density, as achieved by extending the period during which sign accumulates, will also increase power. Scats will persist longer than tracks, but scat use assumes that the species can be reliably distinguished. Survey time will also influence sign density. With grizzlies, spring surveys should find more sign because suitable habitat is restricted during this time. To confirm this, we replicated surveys on 76 km of trails in the northwest quarter of Glacier National Park during each of the first three weekends of May 1987. As expected, we found more sign (0.56 scats per segment and 0.14 grizzly tracks per segment) than in the fall (0.43 scats per segment and 0.06 grizzly tracks per segment).

It is not appropriate to increase sign density by placing survey routes in areas that receive the most animal use because survey routes must represent the entire area. Since population decline may be expressed last in preferred habitats, placing too many segments in such areas will seriously compromise the power of the design.

We believe that managers and biologists can devise a geographically extensive grizzly bear population trend-monitoring program in the Rocky Mountains based on scat occurrence. Application of our model will require species identification of scats. The cost of developing procedures to use DNA analysis to differentiate grizzly scats from those of black bears appears modest (E. Vyse, *personal communication*) and we recommend this as a research priority. Thin-layer chromatography constitutes another differentiation technique that holds promise and is currently being pursued (H. Picton, *unpublished data*).

The design and power of a feasible bear survey must approximate the parameters in Fig. 2C (e.g., survey 500–1000 segments or achieve a mean of about 0.5 sign per segment). We believe that this could be done by repeating surveys from spring through fall and pooling the data for each year. Also, this will reduce the influence of variation in sign density that results from seasonal fluctuations in food availability and/or behavior (Garshelis 1991).

The statistical inferences for the tests we describe apply only to the set of monitored trails, although more general inferences are possible if the sample represents a broader area. Inferences about population trends from trends in sign deposition on trails are a matter of biological judgement. Here, we assume that a significant decline in sign frequency would constitute a significant warning about population status.

Our simulation used equal lengths for sampled trails, but mixed trail lengths will be a practical necessity. Our results may be used as guidelines to estimate power if used with average trail length, but workers should recognize that power will probably decline slightly with unequal trail lengths.

The relationships that we found between sample size and power should apply to a variety of species and sign type if the sign density and variability are similar to ours. Track data reported here represent situations with low probability of encountering sign. Scats, on the other hand, were found relatively frequently and represent systems with more dense sign deposition. To apply this approach to other species, biological insight is required to ensure prudent assumptions are made. The shortest trail segment length that yields independence among segments should be selected. For some species, trail segment dependence may be a concern remedied by surveying every other segment. Obviously, segments should not be so long that the probability of sign approaches one. Power is also potentially affected by the variability between trails. Situations in which variability between trails is greater or less than we observed will result in decreased or increased power. If managers have no information on sign density in an area of question, a pilot survey should provide enough information to evaluate sample size requirements. The kinds of simulations we conducted may need to be repeated when sign deposition varies significantly from ours.

#### ACKNOWLEDGMENTS

The surveys were performed by Wildlands Research Institute (San Francisco State University) students and their instructor, U. I. Mattson. Modeling and analyses were supported by the U.S. National Park Service and the University of Montana. We thank K. A. Keating and B. N. McLellan for helpful reviews of the manuscript.

#### LITERATURE CITED

- Carlock, D. M., R. H. White, J. M. Collins, P. E. Hale, K. G. Johnson, A. S. Johnson, and M. R. Pelton. 1983. The tri-state black bear study. Tennessee Wildlife Resources Agency Technical Report 83-9.
- Daniel, W. W. 1990. Applied nonparametric statistics. Second edition. PWS-Kent, Boston, Massachusetts, USA.
- Dziociolowski, R. 1976. Estimating ungulate numbers in a forest by track counts. *Acta Theriologica* 21:217–222.
- Eberhardt, L., and R. C. Van Etten. 1956. Evaluation of the pellet group count as a deer census method. *Journal of Wildlife Management*. 20:70–74.
- Edwards, R. Y., and D. E. Green. 1959. The measurements of tracks to census grizzly bears. *Murrelet* 40:14–16.
- Fuller, T. K. 1991. Do pellet counts index white-tailed deer numbers and population change? *Journal of Wildlife Management* 55:393–396.
- GAUSS. 1988. The GAUSS system. Version 2.0. Aptech Systems, Kent, Washington, USA.
- Garshelis, D. L. 1991. Monitoring effects of harvest on black bear populations in North America: a review and evaluation of techniques. Proceedings of the Eastern Workshop on Black Bear Research and Management 10:120–144.
- Gerrodette, T. 1987. A power analysis for detecting trends. *Ecology* 68:1364–1372.
- Griffiths, D. A. 1973. Maximum likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of a disease. *Biometrics* 29:637–648.
- Harris, R. B. 1986. Reliability of trend lines obtained from variable counts. *Journal of Wildlife Management* 50:165–171.
- IMSL. 1989. International Mathematical and Statistical Li-



- brary, STAT/LIBRARY. Version 1.1. IMSL, Houston, Texas, USA.
- Johnson, K. G., and M. R. Pelton. 1980. Prebaiting and snaring techniques for black bears. *Wildlife Society Bulletin* 8:46-54.
- Kie, J. G. 1988. Performance in wild ungulates: measuring population density and condition of individuals. General Technical Report PSW-106. Pacific Southwest Forest and Range Experiment Station, USDA Forest Service, Berkeley, California, USA.
- Klein, D. R.. 1959. Track differentiation for censusing bear populations. *Journal of Wildlife Management* 23:361-363.
- Lindzey, F. G., S. K. Thompson, and J. I. Hodges. 1977. Scent station index of black bear abundance. *Journal of Wildlife Management* 41:151-153.
- Mooty, J. J., P. D. Karns, and D. M. Heisey. 1984. The relationship between white-tailed deer track counts and pellet-group surveys. *Journal of Wildlife Management* 48:275-279.
- Paul, S. 1979. A clumped beta-binomial model for the analysis of clustered attribute data. *Biometrics* 35:821-824.
- Pelton, M. R. 1972. Use of foot trail travellers in the Great Smoky Mountains National Park to estimate black bear (*Ursus americanus*) activity. Pages 36-42 in S. Herrero, editor. *Bears—their biology and management*. New Series Publications 23. International Union for Conservation of Nature and Natural Resources, Morges, Switzerland.
- Smith, B. L. 1978. Investigations into black and grizzly bear responses to coastal logging—1977. Thesis. Simon Fraser University, Burnaby, British Columbia, Canada.
- Spencer, H. E., Jr. 1955. The black bear and its status in Maine. Maine Department of Inland Fisheries and Game, Game Division Bulletin 4.
- Tucker, P. A. 1991. Evaluation of techniques to monitor white-tailed deer numbers in the North Fork of the Flathead River, Montana. Thesis. University of Montana, Missoula, Montana, USA.
- United States Fish and Wildlife Service. 1990. Grizzly bear recovery plan. U.S. Fish and Wildlife Service, Missoula, Montana, USA.

## APPENDIX

## SIMULATION DESCRIPTION

The simulation was written in FORTRAN and used IMSL (1989) subroutines RNBET, RNBIN, UVSTA, DSNRNK, and DFRDMN to generate random numbers and carry out the tests. The steps in the simulation were:

- 1) In all cases, generate  $k$  independent beta random variates,  $p_1, \dots, p_k$  from the appropriate beta distribution. The  $p_i$  represent the probabilities of detecting sign on the  $k$  trails in the first year.
- 2) Generate  $k$  independent binomial random variates,  $z_1, \dots, z_k$  with parameters  $n$  and  $p_1, \dots, p_k$ , respectively. These represent the numbers of segments (out of  $n$ ) on the  $k$  trails that had sign in the first year.
- 3) Calculate  $q_1, \dots, q_k$  in one of the following ways, depending on the degree of dependence desired between trails for the two different years:
  - a) Complete dependence: let  $q_i = p_i$ ,  $i = 1, \dots, k$ .
  - b) Complete independence: let  $q_1, \dots, q_k$  be  $k$  new independent beta variates generated from the same beta distribution as in step 1.
- 4) Calculate  $p_i^* = 1 - (1 - q_i)^{1+r}$ , where  $r$  is the proportional change in sign density ( $r < 0$ , decrease;  $r > 0$ , increase;  $r = 0$ , no change). The justification for this is given below.  $p_1^*, \dots, p_k^*$  represent the probabilities of detecting sign on the  $k$  trails in the second year.
- 5) Generate  $k$  independent binomial random variates,  $z_1^*, \dots, z_k^*$  with parameters  $n$  and  $p_1^*, \dots, p_k^*$ , respectively. These represent the numbers of the segments on the  $k$  trails with sign in the second year.
- 6) Perform the desired test of the hypothesis of no difference between years on the  $k$  pairs of values,  $(z_1, z_1^*), \dots, (z_k, z_k^*)$ .

7) Repeat steps 1 through 6 1000 times. For any fixed significance level, calculate the proportion of times the test rejects the null hypothesis of no change.

Generalizations of the above procedure to using three years of data at each sampling period and to generating several years of data for examining tests of trend are straightforward.

The possible dependence between years introduced in step 3 represents dependence between the probabilities  $p_i$  and  $p_i^*$  of sign on segments of a trail in the two different years. Given  $p_i$  and  $p_i^*$ , the numbers of segments with scat in the two years are independent.

The formula for computing  $p_i^*$  in step 4 is based on the following reasoning. Suppose the number of sign  $X$  on a trail segment is a Poisson random variable with parameter  $\lambda$  that is proportional to sign density. This would be the distribution if individual bear sign occurred randomly over the whole trail. Then the probability of detecting sign on a particular segment is

$$\pi_\lambda = \Pr[X \geq 1 | \lambda] = 1 - \exp(-\lambda). \quad (\text{A.1})$$

If the density changes to become proportional to  $(1+r)\lambda$ , then the probability of sign on a segment is

$$\pi_{(1+r)\lambda} = 1 - \exp[-(1+r)\lambda]. \quad (\text{A.2})$$

Therefore,

$$\pi_{(1+r)\lambda} = 1 - (1 - \pi_\lambda)^{1+r} \quad (\text{A.3})$$

The simulation model assumes that the change in sign is uniform across all trails, and that the probability of detecting sign remains constant from year to year on the same trail.