Evaluating Estimators of the Numbers of Females With Cubs-Of-The-Year in the Yellowstone Grizzly Bear Population

Steve CHERRY, Gary C. WHITE, Kim A. KEATING,
Mark A. HAROLDSON, and Charles C. SCHWARTZ

Current management of the grizzly bear (*Ursus arctos*) population in Yellowstone National Park and surrounding areas requires annual estimation of the number of adult female bears with cubs-of-the-year. We examined the performance of nine estimators of population size via simulation. Data were simulated using two methods for different combinations of population size, sample size, and coefficient of variation of individual sighting probabilities. We show that the coefficient of variation does not, by itself, adequately describe the effects of capture heterogeneity, because two different distributions of capture probabilities can have the same coefficient of variation. All estimators produced biased estimates of population size with bias decreasing as effort increased. Based on the simulation results we recommend the Chao estimator for model M_h be used to estimate the number of female bears with cubs of the year; however, the estimator of Chao and Shen may also be useful depending on the goals of the research.

Key Words: Beta distribution; Chao estimators; Closed population estimation; Individual heterogeneity; Model M_h; Negative binomial distribution; *Ursus arctos*.

1. INTRODUCTION

There is a rich literature on estimating closed population size using capture-recapture designs (Otis, Burnham, White, and Anderson 1978; Seber 1982). Initial attempts ignored capture heterogeneity but in recent years parametric and nonparametric methods have been proposed that account for heterogeneity over time and among individuals as well as heterogeneity due to behavioral changes caused by the capture process. Capture heterogeneity

---

Steve Cherry is Associate Professor of Statistics, Department of Mathematical Sciences, Montana State University, Bozeman, MT 59717 (E-mail: *cherry@math.montana.edu*). Gary C. White is Professor, Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Ft. Collins, CO 80523 (E-mail: *gwhite@cnr.colostate.edu*). Kim A. Keating is a Research Wildlife Biologist, U.S. Geological Survey, Northern Rocky Mountain Science Center, Forestry Sciences Laboratory, Montana State University, Bozeman, MT 59717 (E-mail: *kkeating@usgs.gov*). Mark A. Haroldson is a Supervisory Wildlife Biologist, U.S. Geological Survey, Northern Rocky Mountain Science Center, Forestry Sciences Laboratory, Montana State University, Bozeman, MT 59717 (E-mail: *Mark_Haroldson@usgs.gov*). Charles C. Schwartz is a Supervisory Research Wildlife Biologist, U.S. Geological Survey, Northern Rocky Mountain Science Center, Forestry Sciences Laboratory, Montana State University, Bozeman, MT 59717 (E-mail: *chuck_schwartz@usgs.gov*).

Journal of Agricultural, Biological, and Environmental Statistics, Volume 12, Number 2, Pages 195–215
DOI: 10.1198/108571107X198804

has been typically quantified by the coefficient of variation (CV) in probabilities or other parameters (e.g., Poisson means) that determine counts of individual animals in a sample.

Monte Carlo methods have been used to assess the statistical performance of estimators even when analytical methods exist for calculating bias and variability (Chao and Lee 1992). Such methods are useful for comparing the performance of different estimators (Ashbridge and Goudie 2000; Wilson and Collins 1992). Typically, authors simulate data from a random process with a known CV, then draw conclusions about estimator performance based on these simulations. This approach assumes that CV adequately quantifies capture heterogeneity—an assumption that may not be true, as we show below.

In our work we have attempted to estimate the number of adult female grizzly bears with cubs of the year (FCOY) in the Greater Yellowstone Ecosystem (GYE) based on the frequency of sightings of unique individuals. Knight, Blanchard, and Eberhardt (1995) developed a rule set to distinguish pairs of sightings of FCOY among those coming from two unique females or repeated observations of the same female. Tallies of unique females provided a minimum annual estimate of FCOY in the GYE grizzly bear population. These counts were then used to estimate minimum total population size and establish limits of annual allowable human caused mortality (USFWS 1993). Because tallies of unique females were used, the method returned a minimum rather than a total population estimate. Keating, Schwartz, Haroldson, and Moody (2002) evaluated several nonparametric estimators that use sighting frequencies to estimate the total number of FCOY. The CV of the probability of the ith individual being selected in the next sampling episode $\tilde{p}_i$ was used as a measure of capture heterogeneity. They recommended Chao's sample coverage estimator (Chao and Lee 1992) as a reasonable method of estimating total FCOY based on their simulations and results in Chao and Lee (1992). Lee and Chao (1994) also recommended the sample coverage estimator for model M_h under a multinomial sampling setting.

We identify two problems with the recommendations of Keating et al. (2002). First, the simulations on which their recommendations were based all assumed CVs < 1, but recent empirical data (Haroldson 2005) strongly suggest that CV sometimes exceeds 1. Second, further work has shown that CV is not adequate by itself to quantify capture heterogeneity. Specifically, two different distributions of $\tilde{p}_i$'s can yield identical CVs, yet be associated with dramatically different estimator performances. As we show below, the sample coverage estimator recommended by Keating et al. (2002) is not robust to this problem. Our concern is that, absent information about the true underlying distribution of the $\tilde{p}_i$ values, use of this estimator could lead to overestimating FCOY thereby setting annual mortality limits at unsustainably high levels.

We have two objectives in this article. First, we explore the effect of the particular method of data simulation on Monte Carlo based assessments of nine estimators of population size. Second, we update the recommendations of Keating et al. (2002) regarding estimation of the number of FCOY in the GYE.

2. METHODS

2.1 Notation

We use the notation of Keating et al. (2002). We draw a series of independent observations of n individual animals from a closed population of N animals, with $m \leq n$ of these animals being unique. The identity of each is recorded and we assume all animals are correctly identified. The probability of the ith individual being selected in the next sampling episode is denoted by $\tilde{p}_i, i = 1, \ldots, N$ with

$$\sum_{i=1}^{N} \tilde{p}_i = 1.$$

For convenience we will refer to the $\tilde{p}_i$'s as sighting probabilities. Let $n_i, i = 1, \ldots, N$ be the number of times the ith individual is seen. We let f_j denote the number of individuals seen exactly j times, $j = 0, \ldots, r \leq n$. Observable quantities are $n_i > 0, f_j, j = 1, \ldots, r$ and

$$m = \sum_{j=1}^{r} f_j.$$

Note that

$$n = \sum_{i=1}^{N} n_i = \sum_{j=1}^{r} j f_j.$$

We wish to estimate N or, equivalently, f_0.

2.2 Data Simulations

Nine estimators were compared using Monte Carlo simulation methods for population sizes of $N = 20, 40, 60$, and 80, following Keating et al. (2002). We generated data using two different procedures: a beta cumulative distribution (cdf) based procedure as per Keating et al. (2002) and a negative binomial based procedure as per Boyce et al. (2001). For both procedures, heterogeneity of captures was measured by the CV of the N individual sighting probabilities ($\tilde{p}_i$ values). We investigated values of CV $= 0$ to 1.75 in increments of 0.25 except that for the negative binomial model, a CV = 0.01 was used instead of CV = 0 which is impossible for this model. Details are in Appendix A.1 (p. 211).

One thousand simulated datasets were generated for each of 256 combinations of CV, n, and N for the simulations using the beta cumulative distribution function and CV, the mean of n, and N for the negative binomial based simulations. We used n/N as a measure of sampling effort. With each dataset, nine estimators of population size were computed from the resulting f_i statistics.

2.3 Estimators

We examined five nonparametric and four parametric estimators. The five nonparametric estimators included two versions of an estimator due to Chao (1984, 1987, 1989); an asymptotic version (Chao1) and a nonasymptotic version (Chao1Mod), a bias corrected version of Chao1 (Chao2) due to Chao (1989), a sample coverage estimator (Chao and Lee 1992; Lee and Chao 1994) hereafter referred to as SC2, and a modification of SC2 (SC2Mod) due to Chao and Shen (2004). The four parametric estimators were based on Poisson mixture models (Norris and Pollock 1998). Details are in Appendix A.2 (p. 212) along with additional references.

The nonparametric estimators were based on the previous work of Keating et al. (2002), who were looking for a nonparametric estimator that was robust over the range of conditions they believed held for the GYE grizzly bear population. We did not include other nonparametric estimators, either because previous work had ruled them out, for example, the jackknife estimators of Burnham and Overton (1978) (Ashbridge and Goudie 2000; Keating et al. 2002) or because they were not appropriate, for example, the additional sample coverage estimators of Ashbridge and Goudie (2000) which require fixed, defined detection occasions.

Our choice of parametric estimators was based on similar concerns. Recent work by Dorazio and Royle (2003, 2005) and Pledger (2005) also assumed fixed detection occasions. We used a Poisson-based mixture maximum likelihood approach because the usual Pledger (2000) or beta-binomial models do not work. Link (2003) showed theoretically that you cannot use the data to determine which of the underlying models might be appropriate, that is, the exact same data can be generated from two different underlying models (say beta-binomial and logit-normal), but with quite different population sizes. However, all model-based population estimators have an implicit assumption that the model is at least reasonable. We simulated data under different models and used those to estimate population size, and we have recommended the estimator that performs best over that range of simulated data.

2.4 SIMULATION SUMMARIES

Simulation results were summarized as the percent relative bias (PRB) for the number of simulations (l) being reported:

$$\text{PRB} = 100\left[\frac{1}{l}\sum_{i=1}^{l}\frac{\hat{N}_i - N}{N}\right]$$

and mean squared error (MSE) is computed for the number of simulations (l) being reported:

$$\text{MSE} = \frac{1}{l}\sum_{i=1}^{l}\left(\hat{N}_i - N\right)^2 .$$

2.5 COMPARISON OF EMPIRICAL DATA TO SIMULATED DATA

Simulation results ultimately form the basis for recommendations to applied settings. Ideally, such recommendations are valid when observed data are assumed to have been produced by the given simulation method(s). It seems likely, however, that different simulation methods could produce data consistent with observed data, but with differing results for estimator performance. Our comparison of empirical and simulated data was an attempt to show that the simulation results are consistent with our data providing confidence that the results can be used to help in the choice of a statistically valid estimator of FCOY. We attempted to determine this in two ways.

First we compared the sighting frequencies from the 1986–2004 data to the simulated data. The f_i statistics for the 19 years were summed across years, and these values standardized to sum to 1. The result was a mean proportion of animals observed $1, 2, \ldots$, times across the 19 years. Equivalent standardization was performed with the expected f_i of simulated data, and sums of squared errors were used to quantify the discrepancy between the observed and simulated datasets to provide a measure of agreement between the average observed data and the simulated data. Hence,

$$\mathrm{SS} = \sum_i \left(f_i^{\mathrm{sim}} - f_i^{\mathrm{observed}}\right)^2$$

was computed for each of the simulated scenarios to determine which sets of parameters generated data that most resembled the observed data.

Second, we attempted to assess the discrepancy between observed and simulated datasets for each of the 19 years separately. Comparisons were based on the $N/\mathrm{CV}/\mathrm{Effort}$ triple closest to the estimated triple for a given year. We computed the average proportions in each of the f_i categories. There were 132 such categories in the simulations for $N > 20$ and 80 for $N = 20$. We then determined the expected counts using the observed minimum count. The expected counts were computed assuming the observed data were indeed generated by our simulation choice. We computed

$$\chi^2 = \sum_{i=1}^{r} \frac{\left(\mathrm{Obs}_i - \mathrm{Exp}_i\right)^2}{\mathrm{Exp}_i},$$

where r is the number of nonzero expected counts. We carried out a randomization procedure by drawing m times from a multinomial distribution with probability vector equal to the vector of mean proportions computed above. We computed χ^2 for each of 1,000 random draws and generated a randomization distribution of χ^2's associated with our simulation triple and an observed sample size of m. We then determined an approximate randomization P value based on the observed χ^2. Large P values are taken as evidence that the observed data are consistent with the associated simulation $N/\mathrm{CV}/\mathrm{Effort}$ triple.

We also attempted a comparison of estimated sighting frequencies of FCOY in the GYE with the proportions seen in the simulated datasets. We estimated the number of times radioed FCOY were observed independently of radio telemetry relocation flights. To do this, we determined the number of females wearing a functional radio collar at den emergence that produced cubs. We then determined the frequency of sightings for these marked females that were made independent of the radio collar (they were seen without aid of telemetry relocation). We used this frequency distribution as an indication of average sightability.

3. RESULTS

3.1 Sighting Probabilities of Radio-Marked Bears

From 1986–2004, we monitored 85 radio-collared female bears that emerged from their dens with cubs of the year. During our observation period (den emergence to 31 August), 51.8% of these bears were seen independent of any radio-tracking flight. Treating this as a ratio estimate of the proportion yielded an approximate 95% confidence interval of (0.44, 0.59). Sighting frequencies included 22, 11, 7, and 1 observations of these collared bears 1, 2, 3, and ≥ 4 times. The proportion of the radio-marked population that was observed (51.8%) is bracketed by the simulated populations (Tables 1 and 2).

Observations of FCOY were also well distributed geographically throughout the GYE. Since 1997 the Interagency Grizzly Bear Study Team has been conducting observation flights in 37 bear observation areas that encompass 34,700 km^2 (Schwartz 1999). During 1997–2004, there were aerial observations of FCOY from 24 of 26 areas within the USFWS (1993) Grizzly Bear Recovery Zone (RZ), and 5 of 11 flight areas outside the designated RZ. If we include ground observations, there were sightings from 26 of 26 areas within the RZ and 7 of 11 areas outside the RZ. Only 2.7% of 914 FCOY sightings obtained during the period occurred outside the flight areas.

3.2 Comparison of Data From Simulation Methods

The higher the proportion of the population seen the better an estimator should perform. The average proportion seen in our simulations was a function of both effort and CV. As effort increased and CV decreased the average proportion seen increased (Table 1 and 2). In general, the expected proportion seen in the sample was higher with the simulations based on the beta cdf than with the negative binomial simulations. Note that the average proportion seen in the samples from the negative binomial simulations is the same across population sizes for a fixed CV and sampling effort (Table 2).

3.3 Comparison of Observed and Simulated f Statistics

The comparison of the observed and simulated f statistics suggest the beta model generated data most closely resembling the observed data when compared across all 19 years (Table 3). The smallest sum of squares was provided by the beta model with CV $= 1.75$, $N = 80$, and effort $= 2$ although these CV and N values are not similar to our observed data. On average we had an estimated effort of about 1.7 and an estimated CV of around 0.5. The estimator of CV we used is known to be biased low (Chao and Lee 1992) implying that CVs of around 0.75 are not unreasonable for our data on average. Four of the top 12 combinations have CV $= 0.75$ and effort of 1.5 for $N = 20$, 40, 60, and 80. The simulated data align closely with the standardized averages of the observed data (Figure 1), suggesting that our simulations are useful when making inferences about the process of estimating FCOY. Note, however, that the proportion of the population included in the sample for these simulated scenarios was around 70% (Table 1), compared to 51.8% for

Table 1. Parameters of the beta distribution used to simulate data under the various population size (N) and CV scenarios. The value of $\alpha = \beta$ for the U-shaped symmetrical beta distributions. The expected proportion of the population that appears in the sample 1 or more times (Proportion Sampled) is given as a function of the sampling effort, with an effort of 2 meaning that a sample of size $2N$ animals is observed.

						Proportion sampled with effort				
N	CV($\tilde{p}$)	α	$\overline{\tilde{p}}$	Min $\tilde{p}$	Max $\tilde{p}$	1	1.5	2	3	4
20	0.00	1.0000	0.0500	0.0500	0.0500	0.6415	0.7854	0.8715	0.9539	0.9835
40	0.00	1.0000	0.0250	0.0250	0.0250	0.6368	0.7811	0.8681	0.9521	0.9826
60	0.00	1.0000	0.0167	0.0167	0.0167	0.6352	0.7797	0.8669	0.9515	0.9823
80	0.00	1.0000	0.0125	0.0125	0.0125	0.6344	0.779	0.8664	0.9511	0.9821
20	0.25	0.7452	0.0500	0.0416	0.0845	0.6313	0.7723	0.8583	0.9443	0.9777
40	0.25	0.7601	0.0250	0.0211	0.0482	0.6271	0.7690	0.8559	0.9432	0.9773
60	0.25	0.7656	0.0167	0.0141	0.0348	0.6258	0.7679	0.8552	0.9429	0.9772
80	0.25	0.7686	0.0125	0.0106	0.0276	0.6252	0.7675	0.8549	0.9428	0.9772
20	0.5	0.5763	0.0500	0.0351	0.1213	0.6066	0.7430	0.8301	0.9244	0.9658
40	0.5	0.6069	0.0250	0.0182	0.0742	0.6056	0.7441	0.8324	0.9268	0.9675
60	0.5	0.6189	0.0167	0.0123	0.0558	0.6056	0.7449	0.8335	0.9277	0.9681
80	0.5	0.6257	0.0125	0.0093	0.0455	0.6058	0.7454	0.8341	0.9283	0.9685
20	0.75	0.4543	0.0500	0.0298	0.1591	0.5737	0.7058	0.7948	0.8983	0.9488
40	0.75	0.4989	0.0250	0.0159	0.1014	0.5791	0.7151	0.8052	0.9071	0.9549
60	0.75	0.5169	0.0167	0.0109	0.0779	0.5819	0.7190	0.8093	0.9103	0.9570
80	0.75	0.5272	0.0125	0.0083	0.0647	0.5837	0.7213	0.8116	0.9121	0.9582
20	1	0.3609	0.0500	0.0253	0.1971	0.5359	0.6636	0.7539	0.8660	0.9259
40	1	0.4172	0.0250	0.0140	0.1291	0.5509	0.6844	0.7760	0.8848	0.9396
60	1	0.4403	0.0167	0.0097	0.1006	0.5573	0.6925	0.7841	0.8912	0.9442
80	1	0.4538	0.0125	0.0074	0.0843	0.5611	0.6970	0.7885	0.8947	0.9466
20	1.25	0.2864	0.0500	0.0212	0.2352	0.4947	0.6169	0.7074	0.8266	0.8955
40	1.25	0.3523	0.0250	0.0124	0.1569	0.5218	0.6526	0.7450	0.8598	0.9214
60	1.25	0.3799	0.0167	0.0087	0.1235	0.5326	0.6656	0.7581	0.8706	0.9295
80	1.25	0.3961	0.0125	0.0067	0.1042	0.5388	0.6728	0.7652	0.8764	0.9337
20	1.5	0.2251	0.0500	0.0176	0.2733	0.4503	0.5656	0.6546	0.7783	0.8556
40	1.5	0.2990	0.0250	0.0110	0.1848	0.4920	0.6195	0.7122	0.8319	0.9000
60	1.5	0.3303	0.0167	0.0079	0.1465	0.5079	0.6384	0.7313	0.8485	0.9129
80	1.5	0.3488	0.0125	0.0061	0.1243	0.5168	0.6486	0.7415	0.8571	0.9195
20	1.75	0.1734	0.0500	0.0142	0.3113	0.4027	0.5089	0.5942	0.7193	0.8034
40	1.75	0.2539	0.0250	0.0097	0.2126	0.4615	0.5851	0.6773	0.8008	0.8749
60	1.75	0.2884	0.0167	0.0071	0.1696	0.4832	0.6108	0.7037	0.8246	0.8943
80	1.75	0.3090	0.0125	0.0056	0.1443	0.4950	0.6244	0.7174	0.8367	0.9038

Table 2. Parameters of the negative binomial distribution used to simulate data under the various CV and Effort scenarios. Values are the same for all population sizes. The proportion of the population appearing in the sample 1 or more times (Proportion Sampled) is given as a function of the sampling effort, with an effort of 2 meaning that a sample of size $2N$ animals is observed

		Effort									
		1		1.5		2		3		4	
CV($\tilde{p}$)	k	P	Proportion sampled	P	Proportion sampled	P	Proportion sampled	P	Proportion sampled	P	Proportion sampled
0.01	10000.	0.0001	0.6321	0.0002	0.7768	0.0002	0.8646	0.0003	0.9502	0.0004	0.9817
0.25	16.	0.0625	0.6209	0.0938	0.7616	0.1250	0.8481	0.1875	0.9360	0.2500	0.9719
0.5	4.	0.2500	0.5904	0.3750	0.7202	0.5000	0.8025	0.7500	0.8934	1.0000	0.9375
0.75	1.7778	0.5625	0.5477	0.8438	0.6630	1.1250	0.7382	1.6875	0.8275	2.2500	0.8770
1	1.0000	1.0000	0.5000	1.5000	0.6000	2.0000	0.6667	3.0000	0.7500	4.0000	0.8000
1.25	0.6400	1.5625	0.4524	2.3438	0.5382	3.1250	0.5962	4.6875	0.6713	6.2500	0.7186
1.5	0.4444	2.2500	0.4078	3.3750	0.4811	4.5000	0.5312	6.7500	0.5975	9.0000	0.6406
1.75	0.3265	3.0625	0.3673	4.5938	0.4300	6.1250	0.4733	9.1875	0.5314	12.2500	0.5699

Table 3. The level of agreement between the standardized mean observed f statistics and the top 20 simulation scenarios.

Data simulation model	CV($\tilde{p}$)	N	Effort	Sum of squares
Beta	1.75	80	2	0.0005351
Beta	1.5	60	2	0.0005736
Beta	1.25	40	2	0.0007240
Beta	1.5	80	2	0.0008256
Beta	1.5	40	2	0.0010979
Beta	0.75	20	1.5	0.0011624
Beta	1.75	60	2	0.0011970
Beta	0.75	40	1.5	0.0012626
Beta	1.75	40	2.5	0.0013654
Beta	0.75	60	1.5	0.0013832
Beta	1.25	60	2	0.0013849
Beta	0.5	20	1.5	0.0014395
Beta	0.5	40	1.5	0.0014497
Beta	0.75	80	1.5	0.0014529
Beta	0.5	80	1.5	0.0015270
Beta	0.5	60	1.5	0.0015759
Negative Binomial	0.5	80	1.5	0.0016076
Beta	1	20	2	0.0016446
Beta	1	40	2	0.0018139
Negative Binomial	0.5	20	1.5	0.0018322

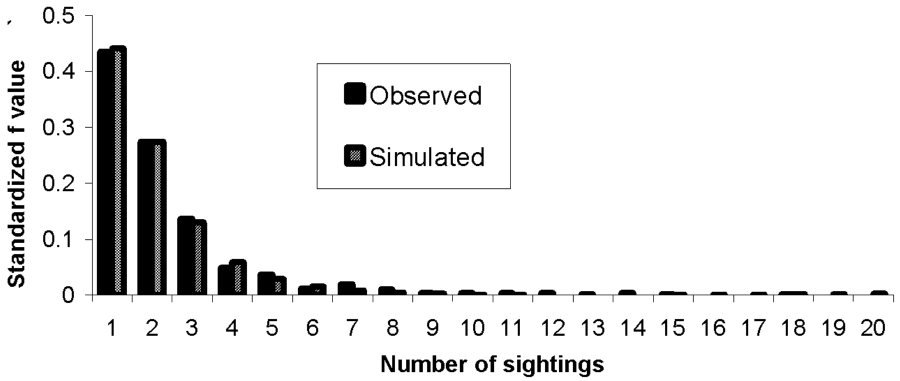

Figure 1. Comparison of the standardized f statistics for the data simulation from the beta model [CV($\tilde{p}$) = 1.75, $N = 80$, Effort = 2] with the means of the standardized observed sighting data 1986–2004.

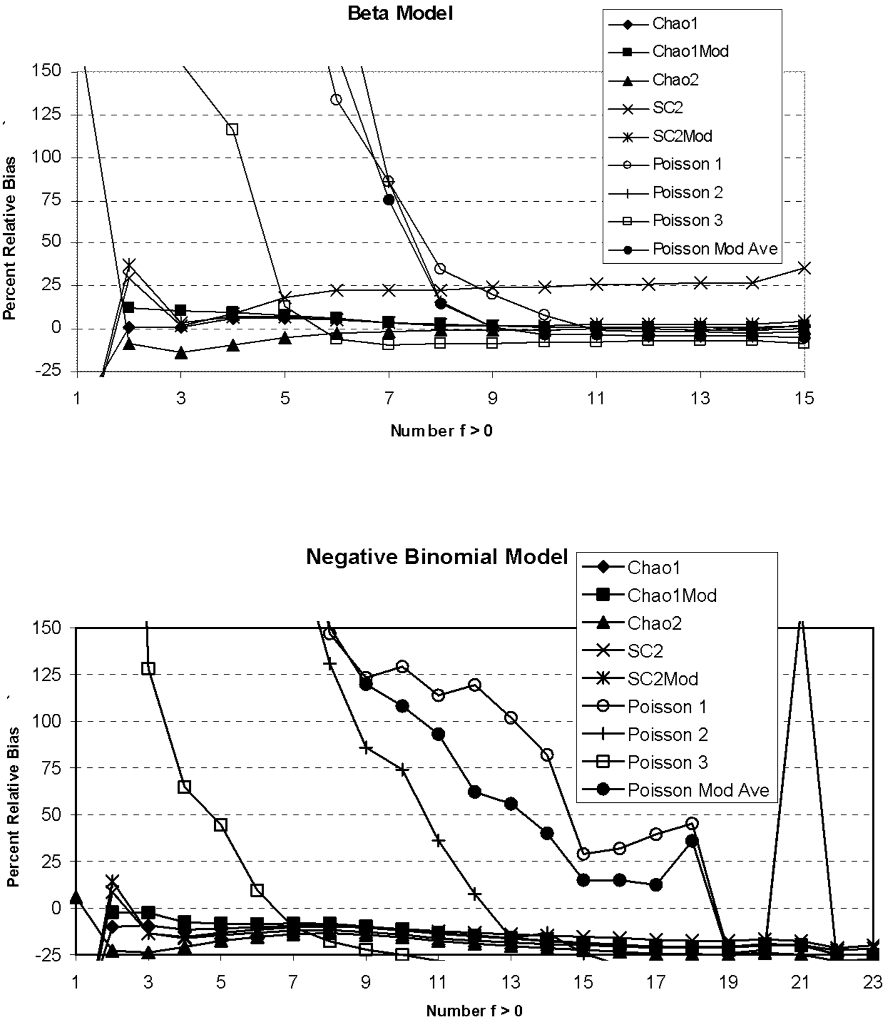

Figure 2. Percent relative bias of the nine estimators considered as a function of the number of the f_i statistics >0.

radio-marked bears. However, we believe the value of 51.8% should not have too much importance attached to it because it was based on a 19-year average with a wide range of estimated N/CV/Effort triples. The number of radio-marked bears seen in any one year is only four or five on average and there is a good deal of uncertainty as indicated by the confidence interval.

Comparisons based on individual years are consistent with at least one of the 256 simulation possibilities and with a simulation N/CV/Effort triple that is close to that actually observed (estimated). Randomization P values ranged from 0.11 to 0.92 for the beta CDF comparisons and 0.26 to 0.94 for the negative binomial based comparisons.

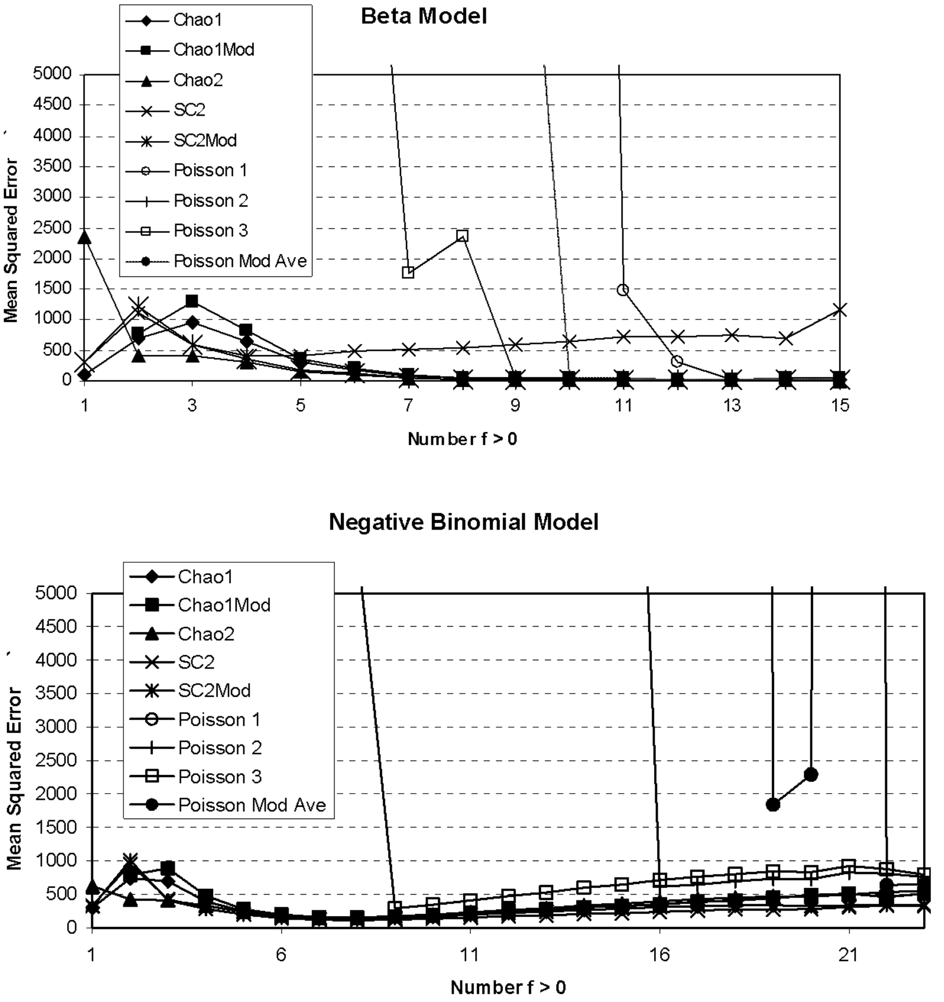

Figure 3. Mean squared error of the nine estimators as a function of the number of f_i statistics >0.

3.4 ESTIMATOR PERFORMANCE

Some simulated scenarios generated too few data to be useful (i.e., less than $3 f_i > 0$), particularly for low effort values. None of the estimators performed satisfactorily when only one f_i statistic was > 0. Most notably, the Poisson mixture estimators were undefined unless more than one f_i was > 0, and generated estimates of N of infinity, causing numerical optimization problems. Although some of the other estimators (e.g., Chao2) generated estimates of N when only a single f value was > 0, use of such estimates seems dubious. An ad hoc measure of the quantity and quality of data was the number of $f_i > 0$. As a first evaluation of the performance of the nine estimators, we evaluated the percent relative bias and mean squared error of each as a function of the number of positive f_i statistics (Figures 2–3). The poor performance of the Poisson mixture estimators for low numbers of $f_i > 0$

where $\hat{N} \rightarrow \infty$ was highlighted in these figures, in that the y-axis was scaled to exclude the extreme estimates from these estimators.

The number of $f_i > 0$ in the observed data ranged from 3 to 13 with mean 5.8 (SD = 2.4). Of the 19 years, only two years had the number of $f_i > 0 = 3$, and five years equal to 4. Because only 2 $f_i > 0$ is quite sparse data, we eliminated all of these simulated cases in the remaining analyses reported here.

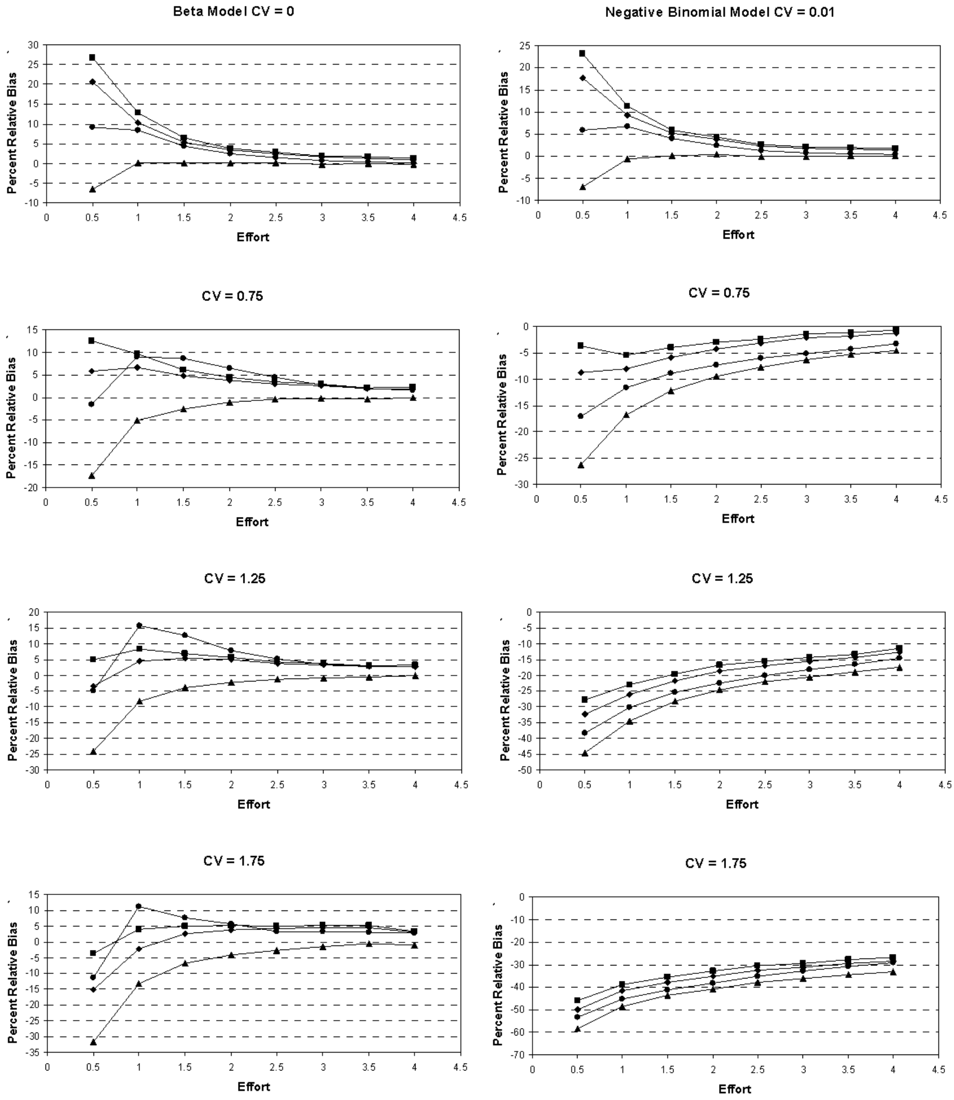

Figure 4. Percent relative bias of the Chao1 (diamond), Chao1Mod (square), Chao2 (triangle), and SC2Mod (circle) estimators as a function of sampling effort for four of eight CV($\tilde{p}$) values simulated for the beta (left column) and negative binomial (right column) simulation models. Note that y-axis scales vary.

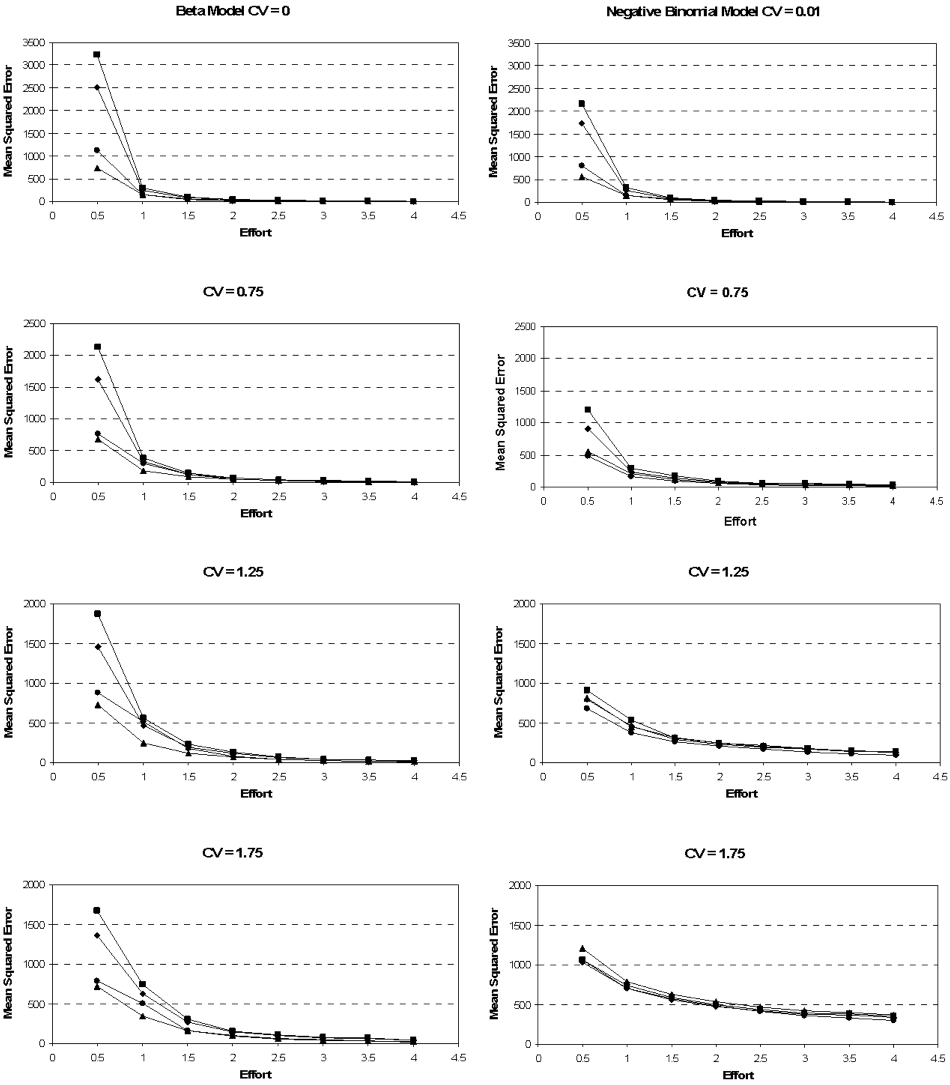

Figure 5. Mean Squared Error of the Chao1 (diamond), Chao1Mod (square), Chao2 (triangle), and SC2Mod (circle) estimators as a function of sampling effort for four of the eight CV($\tilde{p}$) values simulated for the beta (left column) and negative binomial (right column) simulation models. Note that y-axis scales vary.

The performance of the SC2 estimator and all of the Poisson mixture models was also poor compared to the remaining four models (Figures 2–3), so we have only summarized the performance of the Chao1, Chao1Mod, SC2Mod, and Chao2 estimators in the graphs of PRB and MSE as a function of effort for the eight CV($\tilde{p}$) values (Figures 4–5). These all performed in a roughly comparable way in terms of PRB (Figure 4). This was particularly true for higher levels of effort. With effort ≥ 2 these four estimators had PRB in the range of 0–10%. Chao2 was consistently biased low whereas the other three tended to be biased high. Other investigators have also noted the tendency of Chao2 to be biased low (Keating

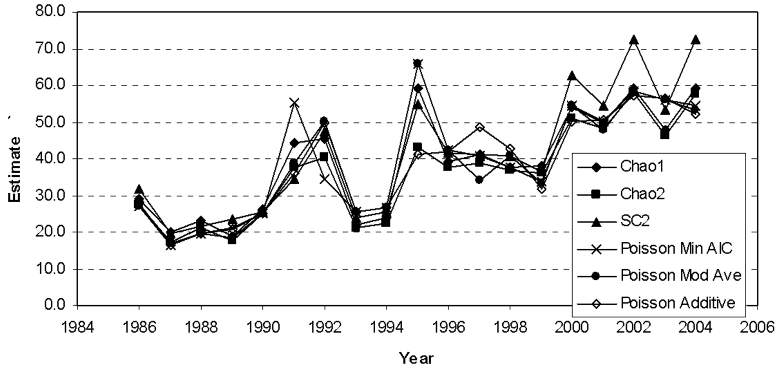

Figure 6. Comparison of estimates for GYE bear population. The Poisson additive model assumes a constant additive effect across years between the high and low sighting means for two mixtures.

et al. 2002; Wilson and Collins 1992).

Performance of the estimators was clearly different for the two different methods of simulating data. The Chao2 estimator demonstrates the smallest PRB, although somewhat negative for the simulated data from the beta model. Performance of the Chao2 estimator was poor for data simulated under the negative binomial model, particularly as heterogeneity increased. However, because the beta model was shown to most closely mimic the observed data, we recommend the use of the Chao2 estimator when data for only a single year are used to compute the population estimate.

Performance of the estimators with the observed bear data (Figure 6) does not suggest large differences between them, although the SC2 estimator does tend to produce larger estimates than the other estimators (Figure 6).

4. DISCUSSION

We believe that evaluation of estimators of this type with simulated data must be conducted more carefully than in the past. There has been an implicit assumption that a conclusion drawn about the performance of an estimator when CV = 1 is the same regardless of how the data were generated. Our results show that such an assumption may not be justified. Further, comparisons of results from different studies may be inappropriate if different simulation methods are used. At the least investigators who wish to apply such estimators need to assess whether or not the particular method of simulating data is relevant for their proposed application.

Although we argue that our simulation methodology produced data consistent with the empirical data we have on grizzly bears in Yellowstone, we have implicitly assumed that our population contains animals all of whom have a high probability of being seen one or more times during the summer field season. Our simulations were designed with this in mind.

Our "recapture rate" on collared bears was high (51.8%). Grizzly bears tend to be crepuscular (Schleyer 1983; Harding 1985), and tend to forage on both vegetable (Mattson, Blanchard, and Knight 1991a) and animal matter during spring and early summer (Green, Mattson, and Peek 1997; Mattson 1997) in meadows, open forests (Mattson et al. 1991a,b; Green et al. 1997; Mattson 1997), and alpine habitats (Mattson, Blanchard, and Knight 1991b) making them visible from both the ground and air. The Interagency Grizzly Bear Study Team flies a series of both radio-tracking and observation flights. Annually, from 1997–2004, two rounds of observation flights have been flown averaging 149 hours/year. The mean sighting rate of FCOY observed on these flights was 0.20 FCOY/hr (West 2005a). From 1997–2004, the average hours flown for telemetry relocation was 394 hr/year, with incidental observations of 0.02 FCOY/hr (West 2005b). Additionally, each year there were numerous aerial relocation, observation, survey, and other flights over the GYE associated with other species occurring over most of occupied grizzly bear range. Incidental sighting of FCOY are typically reported from these flights as well. Overall we believe there is a high likelihood that most if not all grizzly bears in the GYE have sighting probabilities high enough to yield a large proportion of animals sighted within a year. We cannot of course completely rule out the possibility that there are animals with low sightability or that there may be transient environmental conditions that render a good portion of the population effectively unsightable in a given year.

4.1 COMBINING DATA ACROSS YEARS

Likelihood-based estimators have a theoretical advantage over the nonparametric estimators considered here in that the data across years can be combined to model nuisance parameters with a reduced parameter space by assuming some parameters are constant across years (MacKenzie et al. 2005; White 2005). For example, parameters might be considered equal across years, or year-specific covariates could be used to model sighting probabilities. Another benefit from combining data across years is that estimator failures due to sparse data would occur less often, because information is borrowed across years to generate estimates for years where estimator performance would normally be poor. Modeling nuisance parameters across years will provide more precise estimates of the nuisance parameters and potentially generate more precise estimates of population size, although the risk incurred is some bias of the population estimates. Of the 9 estimators considered, $\hat{N}_{M_{h1}}$, $\hat{N}_{M_{h2}}$, $\hat{N}_{M_{h3}}$, and $\hat{N}_{\text{ModAve}}$ could benefit from this methodology. In contrast, none of the nonparametric estimators can use this approach.

The estimator Poisson Additive in Figure 6 demonstrates an application of this methodology for the observed bear data across 19 years. The additive model assumes a constant additive effect (i.e., a constant difference) in the high and low mean sighting probabilities for a two-mixture model, so reduces the number of parameters by 18.

A negative aspect of combining data across years to estimate nuisance parameters is that adding new information each year changes population estimates for previous years. As an example, suppose that an estimate of population size puts allowable mortality in year t at just over observed mortality. However, with addition of data for year $t + 2$, the population estimate in year t is reduced, and now the mortality threshold is exceeded. Such behavior is likely in our application of interest, as the bear population approaches carrying capacity, and mortality equals recruitment complicating the decisions managers reach. A referee pointed out, however, that in practice data would only be aggregated over a limited time span determined in part by consideration of information on grizzly bear life history. Such aggregation might mitigate the problem of varying mortality thresholds.

Surprisingly, the likelihood-based methods did not perform well in either of our simulation scenarios. We attribute this poor performance to (1) low number of $f_i > 0$, and (2) general lack of numerical stability for even moderate numbers of $f_i > 0$. The optimization of likelihoods for mixture distributions is difficult because of multiple optima. In simulation studies such as reported here, user intervention to assess whether a reported solution is the global maximum is precluded. We suspect that at least some of the solutions used in these simulations are not global maxima, and thus affect the reported results on bias and MSE. Although a rule to discard extreme likelihood estimates was considered, this approach was not used because of the subjectivity in defining such a rule.

5. CONCLUSIONS

We view our simulation results as applicable to estimation of population size where there is considerable individual heterogeneity of detection probabilities and where overestimation is to be guarded against. Management of endangered species often is conservative because the penalty for underestimation is to increase management efforts, whereas the penalty for overestimation could lead to extinction if a decline in population size is not detected in time. Individual heterogeneity of detection probabilities is likely high in our example because of the various ways bears are resighted. Our simulations likely apply to other "small" populations where various methods of detecting individuals introduce heterogeneity.

The four nonparametric estimators (Chao1, Chao1Mod, SC2Mod, and Chao2) all performed in a roughly comparable way in terms of PRB (Figure 4, p. 206). This was particularly true for higher levels of effort. If effort was ≥ 2, then all the estimators had PRB in the range of 0 to 10% for the beta cumulative distribution function based simulations, and we believe that these simulations provided results consistent with the empirical results produced by the varying range of conditions in the GYE. Chao2 was consistently biased low whereas the other three estimators tended to be biased high. Given our current state of knowledge, we recommend managers do everything possible to get effort ≥ 1.5. In general, with large enough effort it appears that both SC2Mod and Chao2 perform well on average with Chao2 providing a lower bound. For our specific focus we recommend the use of Chao2 because it is important to not overestimate the number of FCOY in the GYE.

Keating et al. (2002) determined standard errors using the bootstrap. An approximate variance formula for the Chao2 estimator also exists (Chao and Shen 2006). An approximate variance formula for SC2Mod is given in Chao and Shen (2004).

Additional work may produce improved estimation techniques. The modeling approach clearly has great potential in this regard. The simpler nonparametric SC2Mod may also be improved by further work on the choice of a cutoff value κ, although such a cutoff would appear to be a function of factors not under the control of or even knowable by managers.

APPENDIX

A.1 DATA SIMULATIONS

Chao and Lee (1992) showed that for fixed n the CV of the $\tilde{p}_i$ is related to the sighting frequencies (f_i) as

$$(\mathrm{CV})^2 = N\sum_{i=1}^{N}\tilde{p}_i^2 - 1 = \left(N\sum_{i=1}^{n} i(i-1)\mathrm{E}(f_i)/[n(n-1)]\right) - 1,$$

where n is the total number of captures, $n = \sum_i i f_i$.

The vector of capture frequencies (f_i) for the beta distribution simulations was simulated using the methods of Keating et al. (2002). We calculated $\tilde{p}_i$ as the integral of a standard beta distribution over the interval $(i-1)/N$ to i/N as

$$\tilde{p}_i = I_{i/N}(\alpha,\beta) - I_{(i-1)/N}(\alpha,\beta),$$

where $I_x(\alpha,\beta)$ is the incomplete beta function ratio with parameters α and β. There are an infinite number of (α,β) pairs that yield distributions of $\tilde{p}_i$'s with the same CV. For example, the two (α,β) pairs (0.4172, 0.4172) and (6.1386, 6.1386) both produce distributions of $\tilde{p}_i$'s with CV $= 1$ for a population of $N = 40$, but the expected proportion of animals seen in a sample of size $n = 80$ (effort of 2) is 0.776 (Table 1, p. 201) for $\alpha = \beta = 0.4172$ and 0.568 for $\alpha = \beta = 6.13864$. Obviously, estimator performance will differ for these two scenarios. For convenience we limited ourselves to beta distributions with $\alpha = \beta < 1$. This restriction still allowed for simulation of data with a wide variety of distributions of $\tilde{p}_i$'s.

We used the Solver routine in Microsoft Excel to select $\alpha = \beta < 1$ to give the desired CV of the $\tilde{p}_i$'s with the additional constraint that the minimum $\tilde{p}_i$ was maximized so that all animals in the population had nonzero $\tilde{p}_i$ (Table 1). Using the resulting $\tilde{p}_i$ values, we randomly drew n sightings from the simulated population so that the number of sightings per individual in the population (n/N) ranging from 0.5 to 4.0 in equal increments of 0.5 (with this variable designated as sampling effort). After each sighting, the identity of the individual was recorded. Results were then used to tabulate the vector of sighting frequencies, f. Data could have been generated from any cdf but we chose the beta family because it is flexible and because it was used by Keating et al. (2002).

Data were also generated by sampling from a negative binomial distribution although we did not draw directly from that distribution. We assumed that the number of times an individual animal was seen followed a Poisson process with parameter λ_i, $i = 1, \ldots, N$. The λ_i's were assumed to be a random sample from a gamma distribution. Following Boyce et al. (2001) we considered the following parameterization for the negative binomial model:

$$\Pr(X = x) = \frac{\Gamma(k+x)}{\Gamma(k)x!}P^x(1+P)^{-(k+x)},$$

for $x = 0, 1, \ldots$ where k, and $P > 0$. The random variable X is the number of times an animal is sighted and has mean kP and variance $kP(1+P)$, where k and P are the shape and scale parameters of the gamma distribution used to generate the Poisson means. To obtain the appropriate CV we set $k = 1/\mathrm{CV}^2$, and $p = (n/N)/k$ to provide an average of the Poisson means as n/N. Here CV is a measure of the heterogeneity in the Poisson means; however, this is equivalent to the CV for the $\tilde{p}_i$'s. Note that in simulations using the beta model, n (and hence effort) is a fixed constant, whereas for the negative binomial model, n is a random quantity varying with each simulation. To simulate the number of captures of the ith animal for the negative binomial model, we generated a random variable from the gamma distribution with shape parameter k, multiplied the result by P, and used this result as the Poisson mean to generate the observed number of captures of that animal.

A.2 ESTIMATORS

Nine estimators (five nonparametric and four parametric) of population size (N) were compared. We first examined Chao's (1984, 1987, 1989) estimator (hereafter referred to as Chao1):

$$\hat{N}_{\mathrm{Chao1}} = m + \frac{f_1^2}{2f_2},$$

where m is the number of individual animals captured. This estimator assumes model M_t (Otis et al. 1978). Using $\hat{N}_{\mathrm{Chao1}}$, the statistical expectation for the estimate, $E(\hat{N})$, equals N only when sighting probabilities are the same for all animals; that is, when CV = 0. When CV > 0, $E(\hat{N}) < N$ (Chao 1984). This does not ensure $\hat{N}_{\mathrm{Chao1}} \leq N$ in all cases, but it does suggest that $\hat{N}_{\mathrm{Chao1}}$ might provide an inherently conservative approach to estimating N.

The estimator $\hat{N}_{\mathrm{Chao1}}$ is an asymptotic version of the nonasymptotic estimator:

$$\hat{N}_{\mathrm{Chao1Mod}} = m + \frac{f_1}{(n-m_1)^2 + m_2 - m_1^2}\left[\frac{(n-m_1)^3}{nm_1 - m_2} + \frac{m_2 - m_1^2}{n}\right],$$

where $m_1 = \frac{2f_2}{f_1}$ and $m_2 = \frac{6f_3}{f_1}$ (Chao 1984).

We also considered an estimator developed by Chao (1989) for model M_h. When the sample unit is the individual animal, this estimator (Chao2) is given by (Wilson and Collins 1992):

$$\hat{N}_{\mathrm{Chao2}} = m + \frac{f_1^2 - f_1}{2(f_2 + 1)}.$$

Chao and Lee (1992) proposed an estimator based on sample coverage (C), where C is the sum of the $\tilde{p}_i$ values for the m individuals actually observed in the sample. Lee and Chao (1994) offered two estimators of C (SC1 and SC2) that, in the notation of our sampling model, are given by

$$\hat{C}_1 = 1 - \frac{f_1}{n},$$

and

$$\hat{C}_2 = 1 - \frac{f_1 - 2f_2/(n-1)}{n}.$$

For model M_h of Otis et al. (1978), Lee and Chao (1994) then estimated N as

$$\hat{N}_{\mathrm{SC}i} = \frac{m}{\hat{C}_i} + \frac{f_1}{\hat{C}_i}\mathrm{CV}^2,$$

where $i = 1$ or 2, and CV is a measure of the coefficient of variation of the p_i's. Where the sample unit is the sighting of an individual animal, CV^2 is calculated as (Chao and Lee 1992),

$$\mathrm{CV}^2 = \max\left\{\hat{N}\sum_{i=1}^{n}\frac{i(i-1)f_i}{n(n-1)} - 1, 0\right\}.$$

Calculation of CV^2 requires an initial estimate of N. Following Chao and Lee (1992), we used Darroch and Ratcliff's (1980) estimator, $\hat{N}_{\mathrm{DR}} = m/\hat{C}_i$ which assumes equal sightability among all animals in the population. We only evaluated the $\hat{N}_{\mathrm{SC2}}$ estimator here.

A modification of $\hat{N}_{\mathrm{SC2}}$ (SC2Mod) was given by Chao and Shen (2004). Individuals are partitioned into two groups, those with high probabilities of appearing in the sample (high $\tilde{p}_i$ values) and those with low probabilities (low $\tilde{p}_i$ values). Individuals commonly seen are those observed more than κ times. Define $S_{\mathrm{rare}} = \sum_{i=1}^{\kappa} f_i$ and define $\tilde{C}_{\mathrm{rare}} = 1 - f_1/\sum_{i=1}^{\kappa} if_i$. The number of "rare" individuals not seen during sampling is estimated by

$$\hat{f}_0 = \frac{S_{\mathrm{rare}}}{\tilde{C}_{\mathrm{rare}}} + \frac{f_1}{\tilde{C}_{\mathrm{rare}}}\hat{\gamma}^2 - S_{\mathrm{rare}}.$$

The squared coefficient of variation $\hat{\gamma}^2$ is estimated by

$$\hat{\gamma}^2 = \max\left\{\frac{S_{\mathrm{rare}}}{\tilde{C}_{\mathrm{rare}}}\frac{\sum_{i=1}^{\kappa} i(i-1)f_i}{\left(\sum_{i=1}^{\kappa} if_i\right)^2} - 1, 0\right\}$$

if the sample size n is fixed and

$$\hat{\gamma}^2 = \max\left\{\frac{S_{\mathrm{rare}}}{\tilde{C}_{\mathrm{rare}}}\frac{\sum_{i=1}^{\kappa} i(i-1)f_i}{\tilde{C}_{\mathrm{rare}}\left(\sum_{i=1}^{\kappa} if_i\right)\left(\sum_{i=1}^{\kappa} if_i - 1\right)} - 1, 0\right\}$$

if the sample size varies. The population size is estimated by

$$\hat{N}_{\mathrm{SC2Mod}} = m + \hat{f}_0.$$

It is not clear that there is a single cutoff value for κ that is appropriate for general use. Chao and Shen (2004) applied their estimator to simulated data using κ values of 5, 10, and 15 and found that it made little difference under conditions of homogeneity but was more important for heterogeneous populations. Setting κ equal to 10 seemed to yield estimates closest to the truth in their simulations and we also chose that as the cutoff for our simulations.

Keating et al. (2002) used a bootstrapping method to estimate standard errors. They determined that bootstrapped standard errors were comparable to the standard deviations of repeated simulations. We do not further evaluate that method here. Standard deviations presented for the five nonparametric estimators are determined from the simulation results.

The last four estimators are based on the maximum likelihood estimator of N computed from Poisson mixture distributions. Norris and Pollock (1998) developed a mixture model of the number of animals captured i times, $i = 0, 1, \ldots,$ based on the Poisson distribution with likelihood

$$L(N, \pi_a, \lambda_a | A, f_j, m) = \frac{N!}{(N-m)! \prod_{j=1}^{\max(j)} f_j!} \prod_{j=0}^{\max(j)} \left[\left(\sum_{a=1}^{A} \pi_a \frac{e^{-\lambda_a} \lambda_a^j}{j!} \right)^{f_j} \right],$$

for a fixed value of A. A is the number of mixtures for the mean of the Poisson parameter, λ, with $\sum_{a=1}^{A} \pi_a = 1$, where π_a is the probability of inclusion in mixture a. Estimator $\hat{N}_{\text{Poisson}A}$ was computed for $A = 1, 2$, and 3, giving 2, 4, and 6 parameters estimated for the three models. Optimization of these likelihoods was conducted with the NLP procedure of SAS Institute (2003) to obtain maximum likelihood estimates and SEs from the variance-covariance matrix obtained by inverting the negative of the information matrix. Note that the $\hat{N}_{\text{Poisson}A}$ models can be derived from the estimator proposed by Pledger (2000) by replacing the binomial likelihood with the Poisson likelihood.

Model averaging as described by Burnham and Anderson (2002) was applied using the AICc values from these three likelihood-based models to obtain $\hat{N}_{\text{Poisson ModAve}}$. SEs were computed using the formulas provided by Burnham and Anderson (2002).

[Received October 2005. Revised December 2006.]

REFERENCES

Ashbridge, J., and Goudie, I. B. J. (2000), "Coverage-Adjusted Estimators for Mark-Recapture in Heterogeneous Populations," *Communications in Statistics B—Simulation and Computation*, 29, 1215–1237.

Boyce, M. S., MacKenzie, D. I., Manly, B. F. J., Haroldson, M. A., and Moody, D. (2001), "Negative Binomial Models for Abundance Estimation of Multiple Closed Populations," *Journal of Wildlife Management*, 65, 498–509.

Burnham, K. P., and Anderson, D. R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.), New York: Springer-Verlag.

Burnham, K. P., and Overton, W. S. (1978), "Robust Estimation of Population Size when Capture Probabilities Vary Among Animals," *Ecology*, 60, 927–936.

Chao, A. (1984), "Nonparametric Estimation of the Number of Classes in a Population," *Scandinavian Journal of Statistics*, 11, 265–270.

——— (1987), "Estimating the Population Size for Capture-Recapture Data with Unequal Catchability," *Biometrics*, 43, 783–791.

——— (1989), "Estimating Population Size for Sparse Data in Capture-Recapture Experiments," *Biometrics*, 45, 427–438.

Chao, A., and Lee, S.-M. (1992), "Estimating the Number of Classes via Sample Coverage," *Journal of the American Statistical Association*, 87, 210–217.

Chao, A., and Shen, T.-J. (2004), "Nonparametric Prediction in Species Sampling," *Journal of Agricultural, Biological, and Environmental Statistics*, 9, 253–269.

——— (2006), *User's Guide for Program SPADE (Species Prediction and Diversity Estimation)*, (*http://chao.stat.nthu.edu.tw/SPADE_UserGuide(11_12_2006).pdf*).

Darroch, J. N., and Ratcliff, D. (1980). "A Note on Capture-Recapture Estimation," *Biometrics*, 36, 149–153.

Dorazio, R. M., and Royle, J. A. (2003), "Mixture Models for Estimating the Size of a Closed Population when Capture Rates Vary Among Individuals," *Biometrics*, 59, 351–364.

——— (2005), "The Performance of Mixture Models in Heterogeneous Closed Population Capture-Recapture" (Reply), *Biometrics*, 61, 874–876.

Green, G. I., Mattson, D. J., and Peek, J. M. (1997), "Spring Feeding on Ungulate Carcasses by Grizzly Bears in Yellowstone National Park," *Journal of Wildlife Management*, 61, 1040–1055.

Harting, Jr., A. L. (1985), "Relationships Between Activity Patterns and Foraging Strategies of Yellowstone Grizzly Bears," unpublished M.S. Thesis, Montana State University, Bozeman, Montana.

Haroldson, M. A. (2005), "Unduplicated Females" in *Yellowstone Grizzly Bear Investigations, Annual Report of the Interagency Grizzly Bear Study Team*, 2004, eds. C. C. Schwartz and M. A. Haroldson. U.S. Geological Survey, Bozeman, Montana, (*http://www.nrmsc.usgs.gov/products/IGBST/2004report.pdf*).

Keating, K. A., Schwartz, C. C., Haroldson, M. A., and Moody, D. (2002), "Estimating Numbers of Females with Cubs-of-the-Year in the Yellowstone Grizzly Bear Population," *Ursus*, 13, 161–174.

Knight, R. R., Blanchard, B. M., and Eberhardt, L. L. (1995), "Appraising Status of the Yellowstone Grizzly Bear Population by Counting Females with Cubs-of-the-Year," *Wildlife Society Bulletin*, 23, 245–248.

Lee, S.-M., and Chao, A. (1994), "Estimating Population Size Via Sample Coverage for Closed Capture-Recapture Models," *Biometrics*, 50, 88–97.

Link, W. A. (2003), "Nonidentifiability of Population Size from Capture-Recapture Data with Heterogeneous Detection Probabilities," *Biometrics*, 59, 1123–1130.

Mattson, D. J. (1997), "Use of Ungulates by Yellowstone Grizzly Bears," *Biological Conservation*, 81, 161–177.

Mattson, D. J., Blanchard, B. M., and Knight, R. R. (1991a), "Food Habits of Yellowstone Grizzly Bears, 1977–1987," *Canadian Journal of Zoology*, 69, 1619–1629.

——— (1991b), "Bear Feeding Activity at Alpine Insect Aggregation Sites in the Yellowstone Ecosystem," *Canadian Journal of Zoology*, 69, 2430–2435.

MacKenzie, D. I., Nichols, J. D., Sutton, N., Kawanishi, K., and Bailey. L. L. (2005), "Improving Inferences in Population Studies of Rare Species that are Detected Imperfectly," *Ecology*, 86, 1101–1113.

Norris, J. L., and Pollock, K. H. (1998), "Non-parametric MLE for Poisson Species Abundance Models Allowing for Heterogeneity Between Species," *Environmental and Ecological Statistics*, 5, 391–402.

Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. (1978), "Statistical Inference from Capture Data on Closed Animal Populations," *Wildlife Monographs*, 62, 1–135.

Pledger, S. (2000), "Unified Maximum Likelihood Estimates for Closed Capture-Recapture Models using Mixtures," *Biometrics*, 56, 434–442.

——— (2005), "The Performance of Mixture Models in Heterogeneous Closed Population Capture-Recapture," *Biometrics*, 61, 868–873.

SAS Institute (2003), Online documentation for SAS/Base and SAS/OR, version 9.1, Cary, North Carolina: SAS Institute.

Schleyer, B. (1983), "Activity Patterns of Grizzly Bears in the Yellowstone Ecosystem and their Reproductive Behavior, Predation and the Use of Carrion," unpublished M.S. Thesis, Montana State University, Bozeman, Montana.

Schwartz, C. C. (1999), "Evaluation of a Capture-Mark-Recapture Estimator to Determine Grizzly Bear numbers and Density in the Greater Yellowstone Area," in *Yellowstone Grizzly Bear Investigations: Annual Report of the Interagency Grizzly Bear Study Team*, 1998, eds. C. C. Schwartz and M. A. Haroldson, U.S. Geological Survey, Bozeman, Montana. Available online at *http://www.nrmsc.usgs.gov/products/IGBST/1998report.pdf*.

Seber, G. A. F. (1982), *The Estimation of Animal Abundance and Related Parameters* (2nd ed.), New York: Macmillan Press.

U.S. Fish and Wildlife Service (1993), *Grizzly Bear Recovery Plan*, Missoula, Montana, USA. Available online at *http://ecos.fws.gov/docs/recovery_plans/2006/060224.pdf*.

West, K. (2005a), "Observation Flights," in *Yellowstone Grizzly Bear Investigations: Annual Report of the Interagency Grizzly Bear Study Team*, 2004, eds. C. C. Schwartz, M. A. Haroldson, and K. West, U.S. Geological Survey, Bozeman, Montana. Available online at *http://www.nrmsc.usgs.gov/products/IGBST/2004report.pdf*.

——— (2005b), "Telemetry Relocation Flights," in *Yellowstone Grizzly Bear Investigations: Annual Report of the Interagency Grizzly Bear Study Team*, 2004, eds. C. C. Schwartz, M. A. Haroldson, and K. West, U.S. Geological Survey, Bozeman, Montana. Available online at *http://www.nrmsc.usgs.gov/products/IGBST/2004report.pdf*.

White, G. C. (2005), "Correcting Wildlife Counts with Detection Probabilities," *Wildlife Research*, 32, 211–216.

Wilson, R. M., and Collins, M. F. (1992), "Capture-Recapture Estimation with Samples of Size One using Frequency Data," *Biometrika*, 79, 543–553.