Annual Report

Bayesian Modeling of the Rocky Mountain National Park Elk Population

N. T. Hobbs[1] and J. A. Hoeting[2]
Natural Resource Ecology Laboratory and Graduate Degree Program in Ecology [1]
Department of Statistics [2]
Colorado State University
Sunday, October 3, 2010

This report covers three areas of work completed during 2009-2010. During this reporting period, we developed a new data model to accommodate uncertainty in the estimation of population size in a single census as well as the uncertainty that arises from differences among censuses conducted on different dates. We examined the differences in classification data obtained on the ground and from the air. We updated model estimates and forecasts for abundance of elk on the winter range.

Modifications to the data model

There are two, fundamental sources of uncertainty in the annual estimate of the wintering elk population in Rocky Mountain National Park. The first arises from uncertainty in the estimates of the sightability model that is used to convert the raw counts into an estimate of the true population size at the time of census. We will refer to this source of uncertainty as estimation error and will abbreviate is as σ_e. The second source of uncertainty results because estimates made at different times have different values because the number of elk within the census area varies from day to day over the winter. We will call this source of uncertainty sampling error and will abbreviate it as σ_s. Before the winter of 2009-2010, only a single count was made, which made it impossible to estimate sampling error. However, during December - January 2009-2010, three replicate counts were made, thereby allowing an estimate of the population size reflecting the combined effects of estimation and sampling error.

In our initial analysis, we used a fully Bayesian hierarchical model using the 2009-2010 census data to estimate the mean population size as well as to estimate error on each count and sampling error among counts. The seeming advantage of this hierarchical approach is that individual counts that have the lowest estimation error have the greatest influence on the estimate of the mean across all counts, a property known as borrowing strength. The model was

$$\begin{aligned}
&\alpha_i = N_i^2 / s_i^2 \\
&\beta_i = N_i / s_i^2 \\
&P(\mathbf{N},\sigma_{obs} \mid \mathbf{y},\mathbf{s}) \propto \prod_{i=1,3} \text{gamma}(y_i \mid \alpha_i,\beta_i)\,\text{gamma}(N_i \mid r,n)\,\text{gamma}(r \mid .001,.001)\,\text{gamma}(n \mid .001,.001)
\end{aligned} \tag{1}$$

where **N** is a vector of the estimates of the population size, σ_{obs} is the estimate of the observation error including estimation error from the sightability model and sampling error, **y** is a vector of the replicated population estimates, **s** is the corresponding vector of the observed, estimation error for each census and *r* and *n* are shape parameters of the gamma distribution from which counts are drawn; the mean of that distribution is *r/n* . We assumed gamma probability distributions for the likelihoods to assure that all values were positive and chose uninformative gamma priors. This model allowed us to estimate the posterior distribution of individual counts (mean =r/n), that is, the probability that a single count would take on a given value, as well as the posterior distribution of the mean count (mean = $\bar{N}$) (Figure 1).

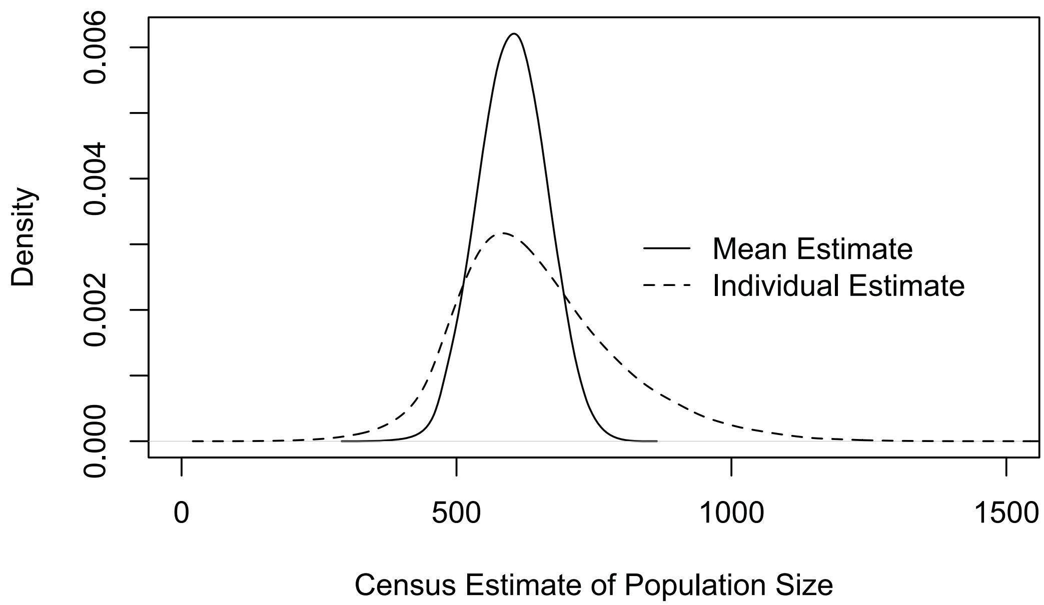

Figure 1. Posterior distribution of individual estimates of the number of elk in Rocky Mountain National Park and the estimate of the average number during winter 2009-2010 based on a Bayesian hierarchical model where the counts with the lowest estimation variance have the greatest influence on the overall mean estimate. These were not the estimates that were used in the final analysis.

The relatively high variance in the distribution of individual counts emphasizes the critical importance of replicating counts each year. These results also illustrate how a single estimate with low estimation error can exert a disproportionate effect on the overall mean count (=602, (95% CI = 483 - 721). For example the February count was 477 with a surprisingly low estimation standard deviation (σ_e=10) . Because σ_e was several fold greater greater for the December and January estimates than for the February estimate, the estimate of the overall mean was strongly influenced by the February census. This influence also accounts for the difference between the mean of the distribution of individual estimates and the mean of the distribution of the estimate of the mean (Figure 1).

However, after discussions with Park staff, it became clear that there were problems with the initial, hierarchical approach to the analysis of the census data because counts with low estimation error also tend to have lower mean values. This source of bias results because counts low estimation error occurs when there is uniform snow cover, which improves the precision of the estimates provided by the sightability data model. However, uniform snow cover also tends to cause animals to leave the park. Discussions with park staff motivated us to change the way we estimated the mean and standard deviation of the census estimates to prevent this bias. In the updated analysis reported here, we used:

$$\alpha_i = N_i^2 / s_i^2$$
$$\beta_i = N_i / s_i^2$$
$$P(\mathbf{N}, \sigma_{obs} \mid \mathbf{y}, \mathbf{s}) \propto \prod_{i=1,3} \text{gamma}(y_i \mid \alpha_i, \beta_i) \text{unif}(N_i \mid 0, 5000) \tag{2}$$
$$\bar{N} = \prod_{i=}^{3} N_i$$

In this model, there is no hierarchy in the estimates of the population size, and as a result, each individual census estimate exerts an equal influence over the estimate of the overall mean, $\bar{N}$. The final census estimates for 2010, using equation 2 are shown in Table 1.

Table 1. Estimates of the number of elk in Rocky Mountain National Park during winter, 2009-2010. Quantiles define the 95% credible interval, which is approximately the same as a frequentist confidence interval.

	Mean	SD	Median	.025% quantile	.975% quantile
December Count	768	191.8	781	326	1096
January Count	659	64.8	661	525	780
February Counts	477	10.0	477	457	496
Mean Count	634	67.5	639	482	753

The insight gained from this analysis motivated a change in the likelihood for the model's estimate of the total population size at time *t*, $N.total_t$. The first twenty five years (1969-1993) of census data have no associated estimates of uncertainty. In this case, the likelihood of $N.total_t$ is

$$y_t \sim \text{Normal}(N.total_t, \sigma_1).$$

where σ_1 is the unobserved observation uncertainty, which can be viewed as including both estimation error and sampling error. During years 1994-2009 the census data included observed estimation error at each time step ($s_{e,t}$) , but no sampling error. In this case, the likelihood becomes:

$$
\begin{aligned}
y_t &\sim \text{Normal}\left(N_t, s_{e,t}\right)\\
N_t &\sim \text{Normal}\left(N.total_t, \sigma_2\right),
\end{aligned}
$$

where σ_2 is the latent sampling error. Finally, during 2010 and in future years that include replicated population estimates, we have

$$
\begin{aligned}
y_{i,t} &\sim \text{Normal}\left(N_{i,t}, s_{e,t}\right)\\
\mu_t &= \text{mean}\left(N_{i,t}\right)\\
\mu_t &\sim \text{Normal}\left(N.total_t, \sigma_s\right),
\end{aligned}
$$

i indexes the individual counts and σ_s is the standard deviation of the distribution of the sample mean.

Analysis of effects of observation method on estimates of population composition

During January 2010, observers on the ground classified the sex and age of elk at approximately the same time as the elk were being classified from the air, permitting comparison of the two methods (Figure 2). Classification from the air tended to include more bulls and fewer calves than classification from the ground. We were particularly interested in the effects of observation method on the number of calves per cow because this ratio could potentially influence the model's estimate of recruitment and, hence, population growth rate. To examine these effects, we compared the proportion of calves in groups composed of cows and calves [i.e., number of calves / (number of calves + number of cows] observed from the ground and from the air (Figure 3).

Observers on the ground classified approximately 50% more calves in cow + calf groups than than were counted by aerial observers, although variance in the ground classifications was high (Figure 3). These observations suggested that downward bias in aerial classification of calves might influence model estimates of population size. To investigate this possibility, we examined model estimates of population size using the aerial classification data with model estimates assuming that the number of calves in classifications were 50% greater then were actually classified from the air.

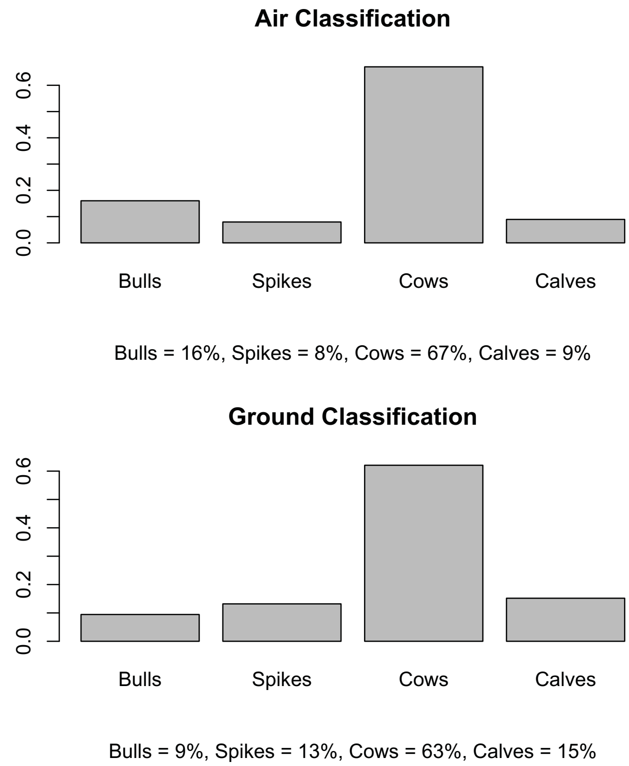

Figure 2. Classification of sex and age of elk from aerial and ground observations.

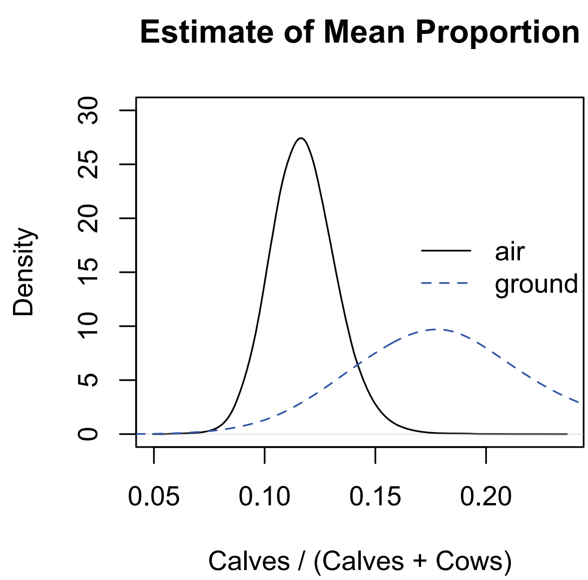

Figure 3. Posterior distributions of the proportion of calves in cow + calf groups classified by observers on the ground and in the air.

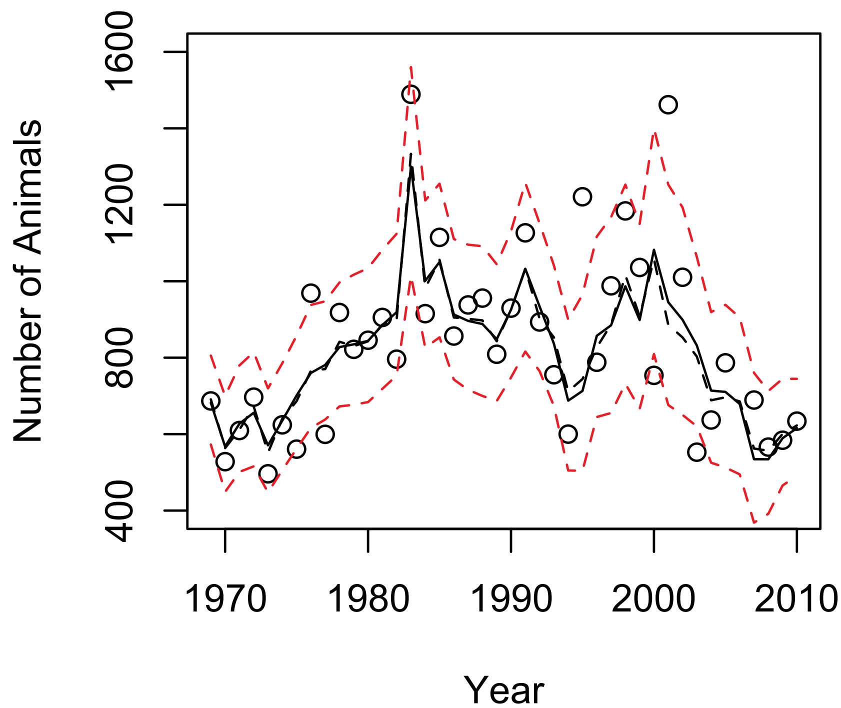

Figure 4. Comparison of model predictions of the elk population size including observed classification data and adjusted data where the number of calves in the sample was increased by 50%. Solid line gives the predictions based on the observed data, the dark dashed line slightly above the solid line shows results from the adjusted data. Light dashed lines are the 95% credible limits on the model predictions based on the observed classifications and points are data on the total population size.

Model predictions were largely insensitive to the change in calf ratios (Figure 4). This insensitivity likely resulted because other elements (cows and males) of the composition data remained unchanged and because the census data had an overriding influence on the model estimates. Thus, it appears the the model estimates of total population size are robust to variation in methods for classification.

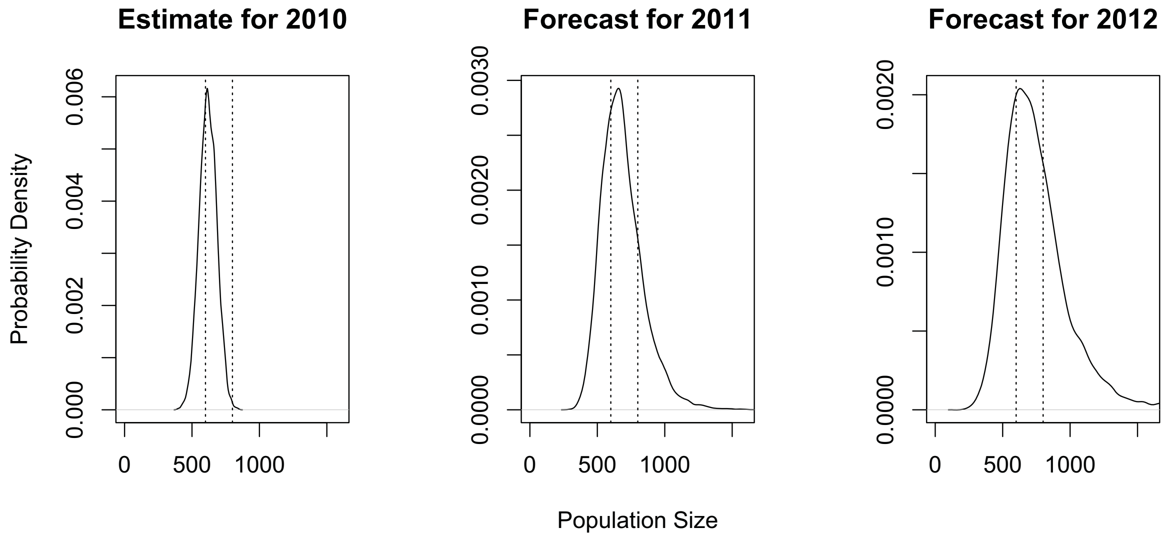

Figure 5. Posterior distribution of the estimated number of elk on low elevation winter range in Rocky Mountain National park during 2010 and posterior distributions of forecasts for 2011 and 2012.

Estimates of 2010 population size and forecasts for 2011, 2012

The average estimate of the number of elk in Rocky Mountain National Park during winter 2009-2010 was within objectives outlined in the elk-vegetation management plan (Figure 5); there was a 61% chance that the current population meets objectives, and <1% chance that it exceeds objectives. However, forecasts for 2011 and 2012 (Figure 5) indicated that in the absence of management, there is a growing probability that the population will exceed the upper limit of the target range in the future (.20 in 2011, increasing to .34 in 2012, Table 2).

Consequently, we compared four scenarios for culling the population: 0, 25, 50 and 75 adult females (Table 2) to support decisions on management needed to maintain the population within objectives. We assumed that culling would occur in 2010 and early 2011 and that no additional culling would occur after early 2011. We also assumed that future harvest could be projected to equal the harvest projection for 2010.

Conclusions

Model results indicate that management of elk in Rocky Mountain National Park is succeeding in meeting objectives specified in the Elk and Vegetation Management Plan. Model estimates of the current population size indicated that there is a 61% probability that the population in 2009-2010 was within the target range of 600-800 animals. However, the model results also implied that continued management will be needed to maintain the population at desired levels. In the absence of future management, we estimate the the probability that the population will exceed the upper limit for population size will increase from essentially 0 currently to 0.34 in 2012 (Table 2). Moderate levels of culling (50 animals in 2010) reduce that probability to .24

(Table 2). However, because harvest outside the park boundary plays a key role in shaping the population's dynamics, decisions on culling should be made with the knowledge that unforeseen shifts in harvest pressure could amplify or attenuate the influence of culling within the park.

Table 2. Estimates of population size and forecasts for effects of four scenarios for culling (0, 25, 50 and 75 adult females).

Year	Median	.025 CL[1]	.0975 CL[1]	P(N < 600)[2]	P(600 < N < 800)[3]	P(N > 800)[4]
2010	618	491	745	0.39	0.61	0.00
			Cull = 0			
2011	667	444	1061	0.31	0.49	0.20
2012	712	415	1427	0.27	0.38	0.34
			Cull = 25			
2011	639	424	1027	0.38	0.46	0.16
2012	679	388	1373	0.34	0.37	0.29
			Cull = 50			
2011	613	399	1010	0.46	0.41	0.13
2012	641	361	1321	0.41	0.35	0.24
			Cull = 75			
2011	585	377	970	0.55	0.36	0.09
2012	607	337	1284	0.49	0.32	0.20

[1]Upper and lower credible limits
[2]Probability that the population size will be below 600 animals.
[3] Probability that the population size will be within the target range of 600 - 800 animals.
[4] Probability that the population size will exceed 800 animals.