USDA FOREST SERVICE REGION 2

MONITORING PROTOCOL FOR AMERICAN BEAVER (*Castor canadensis*): EXAMPLES FROM THE BIGHORN AND BLACK HILLS NATIONAL FORESTS

(Photo from Nature of New England Website – http://www.nenature.com/Beaver.htm)

Protocol Prepared by:

Jeffrey L. Beck
Department of Renewable Resources, University of Wyoming, Laramie, WY

Daniel C. Dauwalter
Department of Zoology and Physiology, University of Wyoming, Laramie, WY

Dennis M. Staley
Renewable Resources, Rocky Mountain Region, USDA Forest Service, Golden, CO

Steven R. Hirtzel
Black Hills National Forest, USDA Forest Service, Custer, SD

Revised May 2008

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOTE ON REVISED PROTOCOL

The final draft version of this protocol was completed in November 2005, and it was first implemented in fall 2007 by the Black Hills National Forest. Data collected during the first season of implementation and experiences by the biologists involved in data collected resulted in this revision of the original draft protocol. The revised protocol has changed from the original in a number of ways. The revision discusses the importance of detection probabilities of beaver food caches when they are counted during aerial surveys, the factors that influence detection probability, and how to control detection probability during aerial surveys (Section 5.3). The revision also explains how detection probability can be measured and used to adjust counts of food caches (Appendix B). Also added to the revision is analysis of data from the first monitoring season on the Black Hills National Forest (Appendix C). The analyses compare beaver food cache densities and presence in watersheds between the habitat suitability strata used in monitoring. Also reported are the sample sizes needed to estimate mean food cache density and proportion of watersheds with caches present at various levels of precision (Appendix C). Results from this analysis were also incorporated into sections explaining the methods for monitoring beaver abundance (Section 5.2) and distribution (Section 6.2). Another added component was detailed analytical methods for determining trends in food cache density (Section 5.4) and trends in presence of food caches in watersheds (Section 6.4). Finally, some organizational changes were also made to integrate the newly added material. Despite these changes, the protocol can still be used as a template by other Forests, beyond the Bighorn and Black Hills National Forests, when designing a long-term monitoring program for beaver populations.

I. INTRODUCTION

In this document we present a revised monitoring protocol for American beaver (*Castor canadensis*) on the Bighorn and Black Hills National Forests. The document begins with natural history and background to consider when monitoring American beaver (hereafter beaver). We then present sections on objectives and sampling design for monitoring abundance and distribution of beaver on the Bighorn and Black Hills National Forests. We describe monitoring structured around a stratified random sample of 6th level Hydrologic Unit Codes (HUC) watersheds. Hydrologic Unit Codes are stratified by a combination of vegetation and physical characteristics important to beaver. We describe separate monitoring approaches to examine trend in abundance and trend in distribution within each monitoring framework and detail the methods one could follow to conduct a monitoring survey.

In this protocol, both abundance and geographic distribution are sampled to achieve inference to forest-wide trends. Evaluating abundance provides information on changes in the size of a beaver population, while evaluating change in distribution provides an understanding of expansion or contraction of beaver and the environments they are influencing. Our approach examines trends in an index to beaver abundance (food cache density) that is sampled every 3 years. Trend in geographic distribution (or range) is examined through sampling every 3 (concurrent with abundance monitoring) to 6 years. To be most efficient, trends in abundance and geographic distribution can be monitored simultaneously during field sampling. Although much of this document includes information specific to the Bighorn and Black Hills National Forests, the underlying purpose is to provide pertinent information for biologists designing a beaver monitoring program on other Forests and Grasslands of Region 2.

1.0. Natural History and Habitat Characteristics

Certain characteristics of beaver, their distribution, and behavior influence the design of monitoring. Beaver occur in aquatic systems throughout much of the Rocky Mountain Region where they serve as important ecological “engineers” by creating and maintaining dams that raise water tables, providing habitat for fish and wildlife, and promoting establishment of woody riparian vegetation. Beaver live in family groups of 3.2 to 8.2 individuals (Novak 1987). The largest groups in North America occur in the midportions of their range (Novak 1987). Beavers are monogamous, and reproduction is typically focused on one female in a family group (Jenkins and Busher 1979). Across North America, the density of beaver family groups ranges from 0.35 to 1.25 / km (Novak 1987).

Food caches are created by beaver in fall to provide a source of food under the cover of ice in winter. The presence of food caches marks the location of family groups (Jenkins and Busher 1979) because only one cache is found per group (Hay 1958). Onset of construction of food caches begins with the first heavy frost and caches are most visible following leaf fall (Novak 1987). Food caches are placed near lodges in deeper water to facilitate access under ice in winter (Grasse and Putnam 1955).

When building a food cache, beaver often form a “raft” of inedible material such as mud, peeled sticks, or conifer branches and logs. This “raft” or “cap” eventually becomes waterlogged and submerges edible foods such as aspen (*Populus tremuloides*) and willow (*Salix* spp.) placed under the raft below the level of winter ice (Slough 1978). Branches and logs forming the consumable portion of a food cache are eaten throughout winter; however, beaver will continue to cut fresh trees as long as they can break through the ice at the edges of ponds (Jenkins and

Busher 1979). The composition of food caches shifts from fall to spring to higher proportions of less preferred forage species such as alder (Alnus spp.; Slough 1978).

Due to an association between relative abundance of beaver and their food caches (Novak 1987), monitoring the density and distribution of food caches can be used to determine beaver population abundance and distribution. Caches can be enumerated using aerial surveys. Aerial monitoring of food caches is most effective in late October and early November after leaf-off when food caches are most visible. Monitoring food caches after ice freezing is difficult if not impossible because only a few sticks may remain above ice (Olson and Hubert 1994). Conflicts with aerial cache surveys include early ice formation in late fall and disturbance to hunters (via aircraft noise) during early to mid-fall hunting seasons. Thus, surveys should be planned to maximize counts while minimizing conflicts.

Beaver are typically associated with specific features of the landscape. Several factors are commonly identified as important: food availability, water conditions, and topography (Table 1). Clearly the most fundamental consideration for suitable beaver habitat is a source of perennial water (Novak 1987). Beaver depend on woody foods during winter in the northerly areas of their distribution (Jenkins and Busher 1979, Allen 1983, Novak 1987). Consequently, availability of winter foods may limit beaver populations (Allen 1983). Beaver eat leaves, twigs, and bark of most woody species found in riparian areas, but they demonstrate a preference for salicaceous trees and shrubs such as aspen, willow, and cottonwood. In addition to providing food, these woody species are also used to construct dams and lodges (Jenkins and Busher 1979, Novak 1987). Thus, relative availability of important shrubs and trees in riparian areas is an important component in selecting areas most suitable for beaver. In addition, beaver select habitats with low gradients (0-15%; Retzer et al. 1956) and wide valley bottoms to provide suitable locations for dams. Narrow canyons with little or no riparian vegetation and steep channels are not suitable for beaver; flat floodplains allow beavers to construct lateral canals to access food supplies (Olson and Hubert 1994).

It is important to recognize potential geographic variation in beaver habitat associations as beavers inhabit a wide-range of aquatic systems including lakes, ponds, creeks, swamps, and rivers across North America (Jenkins and Busher 1979, Novak 1987). Robel et al. (1993) reported poor correlations between beaver colonies/km in Kansas with habitat suitability index values computed from the model presented by Allen (1983). This finding suggests it is crucial to evaluate important predictors of habitat suitability relative to habitat components available on a respective Forest. In other words, we should not expect the same criteria to function in cottonwood (*Populus* spp.) gallery forests of the National Grasslands as in high elevation glaciated streams in a conifer landscape on National Forests.

At a minimum, any analysis to delineate areas of high beaver density should consider water availability, gradient of stream or river, and availability of preferred woody species. Based on our review it seems that a classification process should include identifying: (1) sites with perennial water sources, (2) gradient of lotic systems, (3) width of valley bottoms around lotic systems, and (4) composition of woody riparian vegetation especially hardwoods. These variables were used to model habitat suitability across the Bighorn and Black Hills National Forests, and 6th level HUCs were stratified according to their composite habitat suitability (Appendix A). These stratifications provide a way to monitor beaver within homogeneous sampling units, resulting in more precise estimates of trend and savings in time and money.

1.1. Background

American beaver have been selected as a management indicator species on the Bighorn and Black Hills National Forests. Other Forests or Grasslands may be interested in monitoring populations of beaver to evaluate certain aspects of riparian and stream restoration or management. However, a rigorous, complete monitoring protocol has not been developed for individual Forests or Grasslands within Forest Service Region 2. This document represents a protocol for monitoring beaver on the Bighorn and Black Hills National Forests. It also provides a template for other Forests or Grasslands in Region 2 to design a program to monitor trends in the abundance and distribution of beaver populations.

The design of a monitoring program depends on the goals and objectives for monitoring. Therefore, sampling designs may differ among Forests to meet the goals and objectives of the monitoring program of individual National Forests. Critical choices in designing the monitoring program include defining sampling units, determining what parameter(s) to measure on each sampling unit, and establishing a process for selecting sampling units. Sampling units for broad-scale monitoring of beaver may be defined in a variety of ways. For instance, stream segments, whole stream courses, watershed units, or blocks of land not conforming to watershed units could be established.

To develop an efficient sampling scheme, streams or watersheds (the sampling units) with similar densities of beaver could be grouped into strata. Doing so provides the foundation for a "stratified sampling scheme" that, if stratification is successful, will lead to reduced variance. Reducing variance results in a more efficient estimates of abundance and distribution (reduced cost per unit information) and improves the ability to detect trend over time (Thompson et al. 1998). Defining effective strata requires identifying selection criteria based on environmental factors associated with beaver density. Identifying a small number of strata (2 or 3) should provide the best estimates of beaver abundance considering the low numbers of streams available as well as the sampling cost.

Estimating abundance of beaver is difficult due to their behavior. The literature strongly suggests recording the density of beaver food caches as the preferred observation on sampling units (Novak 1987). Food caches, then, represent an index of the abundance of family groups. Aerial monitoring of food caches in late October and early November provides the best index of abundance because only one winter food cache is established annually by each family group (Hay 1958) whereas one to several lodges may be used by each family group.

Recent discussions in the wildlife literature point to the problems and pitfalls in using indices to evaluate relative abundances of wildlife (Anderson 2001, Anderson 2003, Engeman 2003). Nevertheless, given the challenge of meeting assumptions for direct measures of abundance for some species, indices of abundance, such as counts of beaver food caches per length of stream, provide measures of relative population size and are useful to track year to year changes in populations (Caughley and Sinclair 1994:215, Hayward et al. 2002). The fundamental concern with indices is this—when used to represent relative abundance, index values are suspect because they assume equal detection probabilities across time, habitat types, and observers (Anderson 2001, Anderson 2003). Beaver food caches should have high detection probabilities given that counts are conducted from the air and food caches tend to be in open water and are highly visible. Indeed, mean detection probability of caches with helicopters was 0.89 (i.e., 89%) on 2 prairie rivers in Montana (Swenson et al. 1983). Detection probabilities can be compromised by timing of aerial flights, obstructing ice or vegetation, and observer bias (Payne 1981), but these problems can be alleviated by following a specific field sampling protocol that controls for these factors. Efforts to evaluate detection probabilities should be implemented to ensure that they are high, or estimates of detection probability can be used to adjust counts of food caches and obtain unbiased estimates of cache abundance (Appendix B). This would result in the most reliable estimates of trend in food cache abundance. Another potential problem is that counting food caches may not be a feasible means to monitor abundance of beaver in low elevation areas because beaver may not construct food caches in areas where ice does not form in winter (Collins 1976, Olson and Hubert 1994).

II. OBJECTIVES

The Bighorn and Black Hills National Forests seek to monitor trends in abundance and geographic distribution of beaver at the scale of the National Forest. Monitoring will be done to sufficiently identify a 5% annual decline in beaver food cache abundance over a 9-year period ($1 - [1 - 0.05]^{9\text{years}} = 0.37$, equivalent to a 37% absolute decline) and a 2-fold change (increase or decrease) in the odds (odds = p / [1 – p]; p = proportion of watersheds with caches) of watersheds having beaver food caches after 12 years. For example, the initial monitoring in 2007 for the Black Hills National Forest found caches to be present in 20 of 52 watersheds (p = 0.38), indicating that the odds of a watershed having food caches is: (0.38 / [1 – 0.38]) = 0.61. A two-fold increase in odds after 12 years results in the odds of a watershed having caches present is 1.23 (0.61 × 2 = 1.23); this is equivalent to a proportion (p) = 0.55 of watersheds with food caches present after 12 years. A two-fold decrease after 12 years results in the odds of a watershed having caches present is 0.31 (0.61 × 0.5 = 0.31); this is equivalent to a proportion (p) = 0.23 of watersheds with food caches present after 12 years.

Distribution is expressed as a change in the odds of watersheds having food caches present because proportions can only range from 0 to 1 and they often can be expected change non-linearly in response to factors such as time. For example, if the initial proportion of watersheds with caches present was 0.2, then it might be reasonable to expect the proportion to increase by 150% after 12 years resulting in 0.5 (or 50%) of watersheds to have caches present ($0.2 \times [1 + 1.5] = 0.5$). However, it would be unreasonable to expect a 150% increase in the proportion of watersheds with food caches if the original proportion occupied was 0.85 because it would result in a proportion of 2.13 ($0.85 \times [1 + 1.5]$), an impossible result that suggests that over 100% of watersheds would be occupied (i.e., 213%). Using the change in odds of watersheds having food caches present not only limits the expected proportion to range from 0 to 1, but it allows the expected amount of change to be realistically expressed relative to the odds [and subsequently proportion] of watersheds with food caches present during the initial monitoring year. Finally, the natural logarithm of odds (referred to as the logit transformation, $\log(p/[1-p])$) has a linear relationship with variables such as time. This allows the change in odds per unit time to be estimated over multiple time periods using regression methods (specifically, logistic regression). The parameter estimate for time $(\hat{b}_{year})$ indicates that the logit increases by $(\hat{b}_{year})$ every year, and the $\exp(\hat{b}_{year})$ represents the change in odds per year.

The Forests seek to detect these changes in abundance and distribution with 80% power at α (statistical Type I error rate) of 0.2. This balances the two statistical errors, making it just as likely to fail to detect real changes (Type II error) versus saying change is occurring when in fact it is not (Type I error). If monitoring suggests that the abundance or distribution of beaver is changing by these amounts, the Forest will evaluate and address potential drivers for the change. Our selection of these target levels of change is based on balancing conservation with variability in beaver abundance in Rocky Mountain systems and identifying a level of decline that is detectable given the number of sampling units and duration of the monitoring period. There is no evidence that beaver populations are cyclic (Novak 1987); therefore, we assume that increases and decreases in beaver abundance occur at relatively constant rates that are not interrupted by periodic oscillations. Though beaver abundance or distribution may be influenced positively by transplantation or negatively by their removal to nuisance complaints or fur harvest, these factors, if known, should be qualitatively considered when evaluating the long-term trend related to food cache abundance or distribution.

III. PLANNING AND DESIGN

3.0. Sampling Design

Monitoring beaver on the Bighorn and Black Hills National Forests will be accomplished by employing a sampling design to monitor abundance and distribution. Abundance will be monitored as the change in the density of food caches (caches / km) in a stratified sample of watersheds on each Forest. Food cache density represents an index of beaver family group use of watersheds. Monitoring change in food cache density on streams will be used to determine trends in beaver abundance. Monitoring changes in the presence of food caches in watersheds will be used to determine changes in distribution.

Establishing a sampling design requires attention to the definition of sampling units, definition of the population of interest, and important decisions regarding selection of elements from the population (e.g., simple random, stratified; Thompson et al. 1998). The sampling design to collect observations and estimate abundance of beaver on the Bighorn and Black Hills National Forests includes *sampling elements*, which are defined as beaver food caches; *sampling* *units*, which are 6th level HUC watersheds with suitable beaver habitat; the s*ampling frame* is the list of all 6th level HUC watersheds with suitable beaver habitat that could be sampled on each Forest; the *sample* is the list of randomly selected sampling units (6th level HUCs) on each Forest to be sampled for beaver food caches; the *sampled population* is beaver food caches in the sampling frame of 6th level HUC watersheds on each Forest; and, the *target population* for each Forest is all 6th level HUC watersheds with suitable habitat in late October or early November following aspen and willow leaf drop. Abundance can be monitored on any National Forest in Region 2 using beaver food cache density. As mentioned previously, monitoring beaver food cache density on Grasslands where waterbodies do not freeze may be problematic and likely requires that other techniques to evaluate abundance.

3.1. Selection Criteria to Delineate Suitable Beaver Habitat

Developing an efficient beaver monitoring program on any Forest or Grassland will require classification of beaver habitat suitability. Habitat suitability is used to identify non-habitat to exclude it from the sampling frame. It can also used to stratify sampling units based on habitat quality. We reviewed beaver habitat associations in the western United States and used important habitat characteristics to classify and identify homogenous sampling units (watersheds) for use in monitoring trends in a beaver food cache density.

To classify beaver habitat suitability, it is useful to compare habitat features used by beaver to features not used by beaver. This information can then be used to stratify sampling units according to beaver habitat suitability. This is the process we used on the Bighorn and Black Hills National Forests (Appendix A).

3.2. Sample Size Calculation

Central to any monitoring program is determining the number of sampling units to monitor (Thompson 1992). This number, or sample size, is dependent on the objectives for monitoring, the precision with which the population trend is to be estimated, and acceptable rates of error when determining whether or not a population trend exists.

Pilot studies provide a means to obtain baseline information relative to time and financial costs to monitor beaver. They also allow estimates to be made regarding the variation among sample units that will likely be observed during monitoring. Pilot studies typically involve less sampling than will be done during actual monitoring.

A pilot study for beaver might involve sampling a subset (5 or 10) of the sampling units selected for monitoring among the different strata. A helicopter survey would then be carried out to search the entire length of perennial streams, rivers, and water bodies in each randomly selected HUC.

Costs are a very important consideration when conducting monitoring programs. A helicopter was used to initiate beaver monitoring on Black Hills National Forest in October 2007 according to the original protocol. It cost $36,875 to survey 55 6th-level HUC watersheds (including 3 additional watersheds not selected for sampling totaling 15 flight minutes and $182.50). Helicopter and pilot rental cost $24,875. This included $24,000 for 32.8 hours of flight time at $730 per hour (including 8.9 hours transport time from Durango, CO) and $875 pilot per diem for 7 days. An additional $5,000 was spent on Helitack and geographic information system support, and another $7,000 for biologist time (Grade series 11). Within the sampled watersheds, 804 miles of stream and 14 reservoirs (<1 to 852 ha) were surveyed during 23.9 hours of flight time. In all, 1,775 miles were flown across the Forest, which included shuttle time between watersheds and refueling at Custer or Spearfish, SD airports. Including shuttle time between watersheds, the cost of monitoring averaged $317 per watershed or $21 per perennial stream mile (26 minutes per watershed, or 1.8 minutes per perennial stream mile). All perennial streams were flown in both abundance and distribution watersheds during the 2007 survey. Thus, these costs apply to monitoring abundance watersheds; monitoring distribution watersheds could cost considerably less depending on how fast caches are detected in watersheds where they are present.

Sample size for abundance monitoring

Detecting trends in food cache density across a National Forest is dependent on several factors. The mean food cache density must be estimated precisely for each monitoring year to estimate Forest-wide densities. The variability in trend estimates among watersheds directly influences detection of Forest-wide trends. Thus, less variation in trend estimates for individual watersheds and less variation among watersheds both lead to more precise estimates of Forest-wide trends. The mean cache density within a monitoring year and mean trend estimate across the Forest can both be estimated more precisely by increasing sample size. Considering cost in sample size estimators can increase monitoring efficiency. Methods for considering cost are not reported here, but they can be found in statistical texts that discuss sample size estimators (Thompson 1992, Scheaffer et al. 1996).

Data from a pilot study can be used to determine the sample size (n) needed to estimate the mean food cache density with a certain level of precision for a given monitoring year as:

$$n = \frac{\sum_{i=1}^{L} N_i^2 \sigma_i^2 / w_i}{N^2 D + \sum_{i=1}^{L} N_i \sigma_i^2}$$

where N = number of sampling units in the sampling frame, L = the number of strata, w_i = the fraction of samples allocated to stratum i (n_i / n), $D = \frac{B^2}{z^2_{\alpha/2}}$ where B = the bound on the error of estimation (B caches / km within the true mean), and z = the normal deviate at a specified α ($z_{0.05/2}$ = 1.96; $z_{0.20/2}$ = 1.28). σ_i^2 = the population variance in cache density for stratum i, computed as (Thompson 1992):

$$\sigma_i^2 = \frac{\sum_{i=1}^{N}(y_i - \mu)^2}{N-1}$$

where μ = the unknown, true population mean (the sample mean, $\bar{y}$, can be substituted here), and N = the size of the sampling frame.

More importantly, data from a pilot study covering multiple years can be used to determine the sample size (n) needed to estimate the mean trend in food cache density with a certain level of precision using the same formula:

$$n = \frac{\sum_{i=1}^{L} N_i^2 \sigma_i^2 / w_i}{N^2 D + \sum_{i=1}^{L} N_i \sigma_i^2}$$

Where N = number of sampling units in the sampling frame, L = the number of strata, w_i = the fraction of samples allocated to stratum i (n_i / n), σ_i^2 = the population variance in trend estimates (see above) for stratum i, and $D = \frac{B^2}{z^2_{\alpha/2}}$ where B = the bound on the error of estimation (B percent annual decline within the true mean) and z = the normal deviate at a specified α ($z_{0.05/2}$ = 1.96; $z_{0.20/2}$ = 1.28).

Sample size for distribution monitoring

When data from a pilot study are available, one may estimate the sample size (n) required to estimate the proportion (p) of sampling units with beaver food caches present for each monitoring year within a certain level of precision as (Scheaffer et al. 1996:99):

$$n = \frac{Npq}{(N-1)D + pq}$$

where, n = the number of samples needed, N = the total number of sampling units in the sampling frame, $q = 1 - p$, and $D = \frac{B^2}{z^2_{\alpha/2}}$ where B = the bound on the error of estimation (B units within the true proportion) and z = the normal deviate at a specified α ($z_{0.05/2} = 1.96$; $z_{0.20/2} = 1.28$).

Methods to estimate sample sizes (n) to determine trend in the proportion of sampling units with caches are more complex. As discussed later in Section 6.4, trend in the proportion of sampling units with caches is estimated using logistic regression methods. Sample size estimators for logistic regression are given by Hosmer and Lemeshow (2000; p339). These methods are complex and there is not agreement on which methods to use (Hosmer and Lemeshow 2000). Consequently, they are not reported here. Consult a statistician when attempting to use pilot data to determine sufficient sample sizes to detect trend in beaver distribution.

Other sample size approaches

When working with pilot data, sequential sampling is a useful procedure to assess sample sizes. Moving averages of mean values (mean food cache density or mean trend) and standard deviations of mean values from pilot data are plotted (y axis) against sample sizes (x axis). Bias is evaluated through change in the mean values and precision evaluated through change in the standard deviation as sample size is increased (Elzinga et al. 1998). Sampling beyond the number of sampling units where the plots smooth out indicates a negligible amount of improvement in accuracy per unit effort. In general, we recommend consultation with a statistician when a Forest reaches the stage in planning and implementation of a monitoring protocol where sample sizes are considered.

3.4. Prospective Power Analysis

When pilot data from multiple years are available, a statistician employed by the Forest Service could use these preliminary results to evaluate the statistical power to detect population change (the probability of saying change has occurred when, in fact, it has). For instance, pilot data could be entered into a statistical program such as Program Monitor (Gibbs 1995) to evaluate the number of samples needed to achieve a desired level of power to detect declines in beaver abundance and changes in distribution. Program Monitor requires inputs including mean food cache density, variation in density over space and time, number of years that monitoring will occur, α (statistical Type I error rate), and rate of decline in cache density (trigger point; i.e., 5% annual decline) to assess power to detect change (Gibbs 1995). Power results are very useful to assist biologists in determining sample sizes needed to detect trend within the bounds of their monitoring objectives. At present, there is no information on the temporal variation in food cache density and a prospective power analysis cannot be completed.

IV. TRIGGER POINTS AND EVALUATION

Effective monitoring provides managers with information to evaluate Forest Plans and leads to improved decision-making and management. Monitoring beaver, such as when they have been designated as a management indicator species, requires managers to obtain and evaluate data related to habitats and populations. Consequently, adjustments in land management activities or priorities may need to be made if habitat or population objectives are not met.

Disease, disturbance or forest succession changing cover of preferred woody species, changes in predator communities, intensity of trapping or control efforts, reintroductions, drought or high water years are all potential factors that may lead to changes in the abundance or geographic range of beaver. We suggest an approach where monitoring beaver abundance (food cache density) occurs every 3 years and monitoring geographic distribution (or range) is examined every 3 (concurrent with abundance monitoring) to 6 years. To be most efficient, trends in abundance and geographic distribution can be monitored simultaneously during every or every other 3-year survey for abundance. A 5% annual decline in abundance (i.e., $1 - [1 - 0.05]^{9\text{years}} = 0.37$, or 37% after 9 years) is the trigger point in abundance suggested to motivate changes in Forest management. A 2-fold change (increase or decrease) in the odds of a watershed having food caches present after 12 years is the trigger point for distribution. This is equivalent to the slope from a logistic regression between proportion of watersheds with caches present and year as $\hat{b}_{year} = \pm 0.058$ ($\exp[0.058 \times 12\text{years}] = 2$, $\exp[-0.058 \times 12\text{years}] = 0.5$).

Trend analyses for beaver abundance could be conducted to provide over 80% power to detect a 5% annual decline over 9 years in beaver food cache density with a 20% chance of Type I errors ($\alpha = 0.20$). This same example is also applicable to trend analysis for beaver distribution to provide over 80% power to detect a 2-fold increase or decrease in the odds of watersheds having food caches present after 12 years with a 20% chance of Type I error ($\alpha = 0.20$). However, management context should determine the values used for each Forest.

Monitoring changes in habitat alone would be a poor, unsupportable estimate of the change in beaver populations. The reason for this is because there are many places on National Forests or Grasslands where suitable beaver habitat exists but beaver are not present. The habitat models we created for the Bighorn and Black Hills National Forests are based on habitat use; however, we use that information to stratify the landscape into high, moderate, low (or high and moderate) strata where beaver can be monitored, and identify non-beaver habitat watersheds. In addition, it is likely that beaver populations on many National Forest System lands occur at levels lower than they did historically (Emme and Jellison 2004). Analyses may be furthered by comparing current estimates of abundance and distribution to historical levels; however, formal analysis of monitoring data should be based on current populations surveyed as part of this protocol to provide detection of declines relative to current management activities.

V. METHODS – TRENDS IN ABUNDANCE

5.0. Background –Selecting a Stratified Random Sample

When initiating the monitoring protocol for the Bighorn and Black Hills National Forests, we used existing information to stratify sampling units. We used habitat suitability modeling to stratify sampling units (Appendix A). Modeling provided a way to stratify 6th level HUC watersheds into sampling units of low, moderate, or high quality for the Bighorn National Forest and moderate and high quality for the Black Hills National Forest. Each 6th level HUC sampling unit per Forest was categorized into a stratum based on the proportion of suitable beaver habitat, and a random sample of sampling units was selected within each habitat strata for monitoring (Tables 2 - 4).

A common question that has been raised is whether one changes the designation of a stratum as habitat conditions change over time (e.g., willow and aspen regenerate in watersheds or streams lacking substantial woody riparian cover at the beginning of the monitoring period). The unequivocal answer is no because changing stratum designation would not allow for the ability to evaluate changes in the beaver food cache density over time as related to management activities. Managers need to strictly adhere to strata designation throughout monitoring programs. If desired, future modeling and monitoring of sampling strata could be implemented to compliment habitat restoration efforts.

5.1. Definition of Sample Elements

The sampling elements to monitor trend in abundance of beaver are defined as food caches. The number of food caches counted per sampling unit is used to compute an index of food caches density as: number of food caches / total length (km) of perennial water in sampling unit. Food cache densities are then incorporated into a stratified random estimate of the mean density for the entire area of suitable beaver habitat on each Forest (Thompson et al. 1998), and to estimate trends in food cache density across areas of suitable habitat.

5.2. Sample Size

If no information from pilot studies are available for a National Forest or National Grassland, we suggest sampling 30 sampling units (6^{th} level HUCs) and dividing these samples equally among strata if stratification is used. For example, after delineating strata on the Bighorn

National Forest, we numbered each 6th level HUC and then randomly selected 15 HUCs within both the moderate and high quality habitat strata to yield a total of 30 sampling units (Table 4; Figure 1). Our choice of 30 samples is based on the Central Limit Theorem; however, the efficacy of this sample will be re-evaluated after 2 monitoring sessions (6 years) to determine its adequacy.

In 2007 the Black Hills National Forest initiated MIS monitoring for beaver. They sampled 23 high-quality watersheds and 17 moderate quality watersheds to monitor beaver abundance and an additional 12 watersheds to monitor distribution (Table 5; Figure 2). Although data from one monitoring year cannot be used to assess the sample sizes to estimate trends in food cache density over time, it can be used to determine the sample size to estimate mean food cache density across the Black Hills National Forest for a single monitoring year. We used data from 2007 monitoring and determined that sampling 40 watersheds allowed mean food cache densities to be estimated to within 20% of the true [unknown] mean across the Forest (Appendix C). This level of 20% precision is considered sufficient for monitoring the effects of management (Norris et al. 1996).

5.3. Field Methods

Helicopters and fixed wing aircraft have been used successfully to monitor beaver food caches (Payne 1981, Swenson et al. 1983, Smith 1999). However, Payne (1981) indicated that while Super Cub fixed wing aircraft were less expensive, they were only about half as efficient in censusing beaver food caches as were helicopters in Newfoundland. In Yellowstone National Park, Smith (1999) counted beaver food caches at an altitude of 152.4 m (500 ft) above ground at a speed of 88.5 to 104.6 km per hour (55 to 65 mph) from a Super Cub fixed wing aircraft.

Beaver food caches were counted from helicopter and fixed wing aircraft at 90–150 m (295.3–492.1 ft) altitude above ground in Newfoundland. Beaver food caches were surveyed on prairie rivers in Montana in Super Cub (150 hp) fixed wing aircraft at an altitude of 100–200 m (328.1–656.2 ft) above ground level at a speed of 100 km per hour (62.1 mph; Swenson et al. 1983).

It should be noted that Forest Service aircraft safety requirements preclude use of Super Cub aircraft by Forest Service personnel, but helicopters are considered to be sufficiently safe for use (J. Warder, personal communication, Bighorn National Forest, Sheridan, Wyoming, USA, 2005). A Helicopter Operation Safety Plan and Risk Analysis will be prepared and signed prior to surveys. Daily pre-flight safety briefings will occur to reiterate and/or identify known or anticipated hazards, such as overhead powerlines, for the predicted survey flight path.

When monitoring declines in abundance of beaver, a helicopter survey searches the entire length of perennial streams, rivers, and water bodies in each randomly selected HUC. If helicopters are used for aerial surveys, it may be feasible to use different flight patterns (flying, hovering, circling) to ensure that all perennial waterbodies are observed for caches. Each time a new beaver food cache is located in each sampling unit, it is recorded. Additional data that should be recorded include the date of the survey flight, time each food cache was located, and Universal Transverse Mercator (UTM) coordinates for each food cache (Appendix D).

Since food cache abundance is standardized by length of perennial stream, there needs to be consistency in how the perennial streams within watersheds are measured for length and surveyed for food caches. The fraction of the stream network that is perennial can increase or decrease with wet and dry climate cycles. Thus, the length of perennial stream within a watershed may change over time. We recommend that all perennial water during a monitoring year be surveyed for beaver food caches, and the extent of perennial water within a watershed should also be determined for each monitoring year (possibly as a separate activity). It is also important to maintain consistency in whether short segments of perennial tributary streams are surveyed or if surveys are only conducted on the main streams within each watershed. These details are to be recorded by individual Forests so that consistency in field sampling protocols can consistently be used during each monitoring year. The total length of perennial stream should be reported in addition to food cache densities and trends in densities so it is known whether the length of stream is changing when densities are not changing.

It is important to control the probability of detecting beaver food caches during helicopter surveys. Detection probabilities during helicopter surveys can vary due to observer bias and training, riparian vegetation, watershed valley shape, time of day, and weather. Observers will often differ in their ability to detect food caches during helicopter surveys. For example, if one experienced observer can detect 90% of caches and a less experience observer only detects 70% of caches, this could result in a difference of 20% in cache density between observers. Observer training can improve detection probability and reduce differences among observers. Detection probability also has the potential to increase over time due to observer experience, but an initial training period can reduce the change in detectability due to experience over time. Thus, we recommend that new observers be trained. We recommend that distribution watersheds be used for training and, thus, at least one or two distribution watersheds be monitored prior to any abundance watersheds being monitored. Food caches located during previous monitoring years in distribution watersheds, such as those initially surveyed in 2007 on the Black Hills National Forest, can be used to train new observers and retrain previous observers on determining the presence or absence of food caches at a known location. It is only appropriate to conduct training in distribution watersheds since the influence of training on detection probabilities and monitoring trends will be negligible. Several distribution watersheds should be sampled prior to abundance watersheds so that observers can reliably detect beaver food caches during abundance surveys. Photos of previously located food caches can also be used for the training period prior to sampling watersheds for abundance monitoring.

Beaver food caches located in abundance watersheds during prior monitoring surveys should not be used for training new observers or retraining previous observers. Doing so could result in different detection probabilities between previously located caches and new or previously undetected caches. As more caches are detected during subsequent monitoring surveys, the probability of detection throughout the watershed could increase over time. This could result in trends in beaver food cache abundance that are not due to real changes in abundance but trends that are caused by increasing probability of detection over time. Ensuring that each cache has detection probability that is unaffected by prior observer experience will result in trends in beaver food cache abundance that are real and not a result of increased detection probability.

The probability of detecting beaver food caches during helicopter surveys can also be affected by riparian vegetation, watershed valley shape, time of day, and weather. For example, caches can be difficult to see in coniferous vegetation when compared to alder and willow. Cache detectability is typically high (70-100%) in broad valleys, but it can decrease in narrow valleys and canyons (Steve Hirtzel, Black Hills National Forest, personal communication; Swenson et al. 1983). Since this protocol requires that every watershed be resampled during subsequent monitoring periods and probabilities of detection are generally high, indices of food cache densities should be useful in detecting trends in beaver abundance as long as detection probabilities do not change over time within individual watersheds.

Time of day and weather can influence light conditions and visibility of beaver food caches from the air. Overcast conditions produced a "flat light" that improved cache detectability because the harsh shadows and glare present on sunny days was lacking (Steve Hirtzel, Black Hills National Forest, personal communication). Changes in detectability due to these factors can add variation to food cache densities when monitored over time. Thus, aerial surveys should attempt to maximize survey time when weather and light conditions provide the best detectability, if practicable. This will ensure that detection probabilities are maximized and variation in detectability over time is minimized. Controlling these sources of variation in detection probability due to sampling will allow trends in food cache abundance to be more reliably detected over time. Periodically assessing detection probabilities over time will elucidate whether they are changing over time.

5.4. Assessing Population Trends over Time

Beaver food cache monitoring will not require annual surveys because of the persistence of beaver family groups. Consequently, food cache surveys could be conducted every 2 or 3 years. In this protocol we suggest 3 year intervals; however, Forests may decide to survey more frequently due to management goals or suspected changes in beaver abundance.

As previously mentioned, the number of food caches counted per sampling unit is used to compute an index of food caches per stream km. These index values are then incorporated into a stratified random estimator to estimate the mean food cache density for the entire area of suitable beaver habitat on each Forest (Thompson et al. 1998). The estimator for the population mean (μ) for a stratified random sample is:

$$\bar{y}_{st} = \frac{1}{N}\sum_{l=1}^{L} N_l \bar{y}_l$$

where $\bar{y}_{st}$ = the mean food cache density for a stratified random sample, N = the number of sampling units in the sample frame, N_l = the number of sampling units in the sample frame of stratum l, L = the total number of strata, and, $\bar{y}_l$ = the mean food cache density for stratum l computed as for a simple random sample:

$$\bar{y}_l = \frac{\sum_{i=1}^{n_l} y_{il}}{n_l}$$

where y_{il} = the food cache density for watershed i in stratum l, and n_l = the number of watersheds sampled in stratum l.

The variance for $\bar{y}_{st}$ is:

$$s^2_{\bar{y}_{st}} = \frac{1}{N^2}\sum_{l=1}^{L} N^2\left(\frac{N_l - n_l}{N_l}\right)\left(\frac{s_l^2}{n_l}\right)$$

Where N, N_l, and n_l are as defined above, the expression $\left(\frac{N_l - n_l}{N_l}\right)$ is a finite population correction factor that is used to reduce variance according to the proportion of the sampling frame included in the sample. For example, 23 out of 37 watersheds in the high-quality stratum on the Black Hills National Forest were selected for beaver monitoring. s_l^2 = the sample variance for stratum l computed as for a simple random sample:

$$s_l^2 = \frac{\sum_{i=1}^{n} y_{il}^2 - n_l \bar{y}_l^2}{n_l - 1}$$

with all terms as defined above.

Forest-wide trends in the food cache density will be estimated. First, trends in food cache density are estimated for each watershed. Trends for each watershed are estimated by regressing the $\log_e$(food cache density [1] +0.001) on year and determining the slope coefficient for this relationship. Using the $\log_e$-cache density is appropriate because animal abundances often change by some percentage each time unit (e.g., 2% decline per year) rather than changing by a constant amount (e.g., a decline of 1 cache/km per year) (Thompson et al. 1998). By using $\log_e$-cache density, the slope of the regression line $\hat{b}_{1i}$ is an estimate of the percent annual change in the cache density when slope estimates are small (<0.20), and exp($\hat{b}_{1i}$) × 100 estimates the percentage of individuals in the population at time *i* that remain at time *i* + 1 (Thompson et al. 1998). After trends are estimated for each watershed, then trends (i.e., slope coefficients) are averaged across watersheds to estimate average trend among the watersheds in the sampling frame.

For each watershed, the $\log_e$(cache density + 0.001) is calculated for each year and then regressed versus year to estimate trend:

$$\log_e\left(\hat{D}_{ilj} + 1\right) = b_{0il} + b_{1il}\left(Year_j\right)$$

where $\hat{D}_{ilj}$ = the estimated cache density (cache / km) for watershed *i* of stratum *l* in year *j*. The regression slope estimates ($\hat{b}_{1il}$; approximate estimate of percent annual change in cache density) for each watershed *i* in stratum *l* will then be used to compute a mean slope $\hat{\bar{b}}_{1l}$ among watersheds in stratum *l*:

---

1 It is recommended that the index value be expressed as the number of caches per perennial stream kilometer (N/km) so that adding a constant of 0.001 will result in little change in the index value. However, if it is desired that the index value be expressed on a different scale, such as N/mile, then a smaller constant should be added (e.g., 0.00001). A constant needs to be added to the index value because there is no natural logarithm for an index value of zero if no caches are observed in a watershed.

$$\hat{\bar{b}}_{1l} = \frac{\sum_{i=1}^{n} \hat{b}_{1il}}{n_l}$$

where $\hat{b}_{1il}$ is as above, and n_l is the number of sampling units (watersheds) sampled in stratum l.

The variance estimate for mean slope $s^2_{\hat{\bar{b}}_{l1}}$ is calculated as:

$$s^2_{\hat{\bar{b}}_{1l}} = \frac{\sum_{i=1}^{n_l} \left(\hat{b}_{1il} - \hat{\bar{b}}_{1l} \right)^2}{n_l - 1}$$

where $\hat{b}_{1il}$, $\hat{\bar{b}}_{1l}$, and n_l are defined as above. Then the stratified estimate of the mean slope is:

$$\hat{\bar{b}}_{1st} = \frac{1}{N} \sum_{l=1}^{L} N_l \hat{\bar{b}}_{1l}$$

and the variance of $\hat{\bar{b}}_{1st}$ is:

$$s^2_{\hat{\bar{b}}_{1st}} = \frac{1}{N^2} \sum_{l=1}^{L} N_l^2 \left(\frac{N_l - n_l}{N_l} \right) \left(\frac{s^2_{\hat{\bar{b}}_{1l}}}{n_l} \right)$$

where all terms are defined as above.

The variance estimate can be used to compute a confidence interval for the mean slope $\hat{\bar{b}}_{1st}$ estimate. The lower confidence limit is:

$$\hat{\bar{b}}_{1st} - t_{n-1,\alpha} \times \sqrt{s^2_{\hat{\bar{b}}_{1st}} / n}$$

and the upper limit is:

$$\hat{\bar{b}}_{1st} + t_{n-1,\alpha} \times \sqrt{s^2_{\hat{\bar{b}}_{1st}} / n}$$

where $\hat{\bar{b}}_{1st}$, $s^2_{\hat{\bar{b}}_{1st}}$, and *n* are as before, and $t_{n-1,\alpha}$ = the t-value from a t distribution table with n-1 degrees of freedom and specified α. One can be 100(1 – α)% sure that the true but unknown mean percent annual change in the food cache density is within this interval.

A one-tailed t-test can be used to determine if the mean slope $\hat{\bar{b}}_{1st}$ is significantly less than zero. A Type I error rate of α = 0.20 is recommended over a more conservative rate (e.g., α = 0.05) to reduce the chance of missing a decline that is real (Type II error). However, a higher Type I error rate will, by definition, increase the risk of detecting false changes. Type I and II error rates are inversely related, but not proportional. Management context will determine what are acceptable levels of each risk (Mulder et al. 1999). If the slope is significantly less than 0 and the estimated mean slope ($\hat{\bar{b}}_{1st}$) is -0.05 (equivalent to a 5% annual decline) or less after 9 years, then management action should be taken to examine causes for declines in beaver abundance.

The monitoring objective is designed specifically to examine potential declines in abundance of beaver at the scale of the National Forest. Therefore, one-tailed statistical tests should be used to evaluate change in abundance. Although the design for abundance focuses in detecting declines, graphical analyses and descriptive statistics can be employed to examine the monitoring results to identify increases in abundance.

VI. METHODS – TRENDS IN DISTRIBUTION

6.0. Background

Forest-wide distribution of beaver is not expected to change rapidly unless the species experiences a significant stressor (e.g., epidemic disease). Therefore we recommend monitoring distribution through surveys every 3 (concurrent with abundance monitoring) to 6 years.

6.1. Definition of Sample Elements

Distribution of beaver is defined as the proportion of sampling units (watersheds) with beaver food caches present. Trends will be determined by documenting changes in the proportion of sampling units with food caches present over time.

6.2. Sample Size

To select a sample of sampling units to monitor every 3 to 6 years we recommend the following steps. First, all sampling units that potentially can be occupied by beaver (i.e., low, moderate, and high quality strata) are identified and numbered but excluding sampling units that were originally selected for abundance monitoring. Next, a random sample will be drawn from sampling units in all strata. The estimator for this sample will not be stratified because we are interested in obtaining an estimate for beaver distribution irrespective of habitat quality. Because monitoring sampling units for presence of beaver (rather than counting all beaver in the watershed) could be relatively fast, it may be feasible to monitor a larger number of sampling units than when monitoring abundance. Pilot data from individual Forests or Grasslands can be used to determine the number of sampling units required to estimate the proportion with caches present and the trend in proportion with caches present over time.

Monitoring data from 2007 on the Black Hills National Forest were evaluated to determine the number of sampling units required to estimate the proportion of watersheds with caches present for an individual monitoring year. The Black Hills National Forest is currently monitoring 52 watersheds for presence of beaver food caches, and that sample size allows them to determine the proportion of watersheds with caches present within 0.05 to 0.10 of the true [unknown] proportion (Appendix C). This precise estimate of the proportion of watersheds with beaver food caches will allow trends in proportion of watersheds with caches to be detected with a high level of precision. However, there currently is no information from multiple years to determine the sample size needed to precisely detect trends in the proportion of watersheds with caches present using logistic regression.

6.3. Field Methods

To evaluate changes in distribution of beaver, a helicopter survey would begin searching the length of perennial streams, rivers, and water bodies in each randomly selected sampling unit (watershed). In this case, monitoring would cease as soon as 1 food cache was located in each sampling unit because active beaver presence in that sampling unit has been confirmed. On the other hand, if no cache is observed, the search would not cease until the entire sampling unit had been surveyed. Data from the presence/absence survey to document distribution would be combined with data from the abundance survey to determine distribution every 3 to 6 years.

6.4. Assessing Trends in Distribution over Time

Changes in distribution of beaver populations through presence or absence of food caches need not be an annual activity due to the persistence of beaver colonies. Monitoring to evaluate expansion and contraction of populations could be conducted as infrequently as every 6 years. The parameter of interest in this case is the proportion of sampling units monitored that have food caches present (Thompson et al. 1998). Following notation in Schaeffer et al. (1996), the estimator for the population proportion (p) for a simple random sample is:

$$\hat{p} = \frac{\sum_{i=1}^{n} y_i}{n}$$

Where, $\hat{p}$ = the estimator for a population proportion (p) for a simple random sample, y_i = presence of beaver in sampling unit i (presence = 1; absence = 0), and n = the total number of sampling units monitored. The variance estimator for $\hat{p}$ is (Scheaffer et al. 1996):

$$s_{\hat{p}}^2 = \frac{\hat{p}\hat{q}}{n-1}\left(\frac{N-n}{N}\right)$$

where, $\hat{q} = 1 - \hat{p}$, $\hat{p}$ and n are as defined above, and N = number of sampling units in the sampling frame. Again, $\left(\frac{N-n}{N}\right)$ is a finite population correction that reduces variance according to the fraction of sampling frame sampled. Although the variance is not used directly in detecting changes in beaver distribution, it is recommended that a measure of precision be reported for the proportion of sampling units that have beaver food caches present for each monitoring time period.

To place a bound on the error of estimation, the lower $(1 - \alpha)\times 100\%$ confidence limit is:

$$\hat{p} - t_{n-1,\alpha} \times \sqrt{s_{\hat{p}}^2}$$

and the upper limit:

$$\hat{p} + t_{n-1,\alpha} \times \sqrt{s_{\hat{p}}^2}$$

After Forest-wide estimates are computed for a monitoring time period, time trends in the proportion of sampling units having beaver food caches present can be determined. Logistic regression can be used whereby the proportion of sampling units occupied $\hat{p}_k$ for years 1 through k will be logit-transformed and regressed on year:

$$\log_e\left(\hat{p}_k / 1 - \hat{p}_k\right) = b_0 + b_1\left(year_k\right)$$

A statistical software package that performs logistic regression should be used to account for the correct variation behavior of $\hat{p}_k$, as opposed to regressing logit of $\hat{p}_k$ versus time using least squares methods (Hosmer and Lemeshow 2000). The statistical test of interest is whether the estimated rate of change b_1 is significantly different from zero, using a two-tailed t-test. The parameter b_1 is the change in the log-odds of a watershed having beaver food caches present, and $\exp(b_1)$ is the annual increase or decrease in the odds a watershed having caches present per year. A two-tailed test is used to determine whether b_1 is less than or greater than zero because both increasing and decreasing trends in beaver distribution are of interest. Again, a Type I error rate of $\alpha = 0.20$ is recommended, but it should be set according to the risks associated with making a Type I versus a Type II statistical error. When conducting the statistical test, a finite population correction should be applied to the variance estimate (and subsequently the standard error estimate) for b_1. For example, 52 of the 73 watersheds in the sample frame for the Black Hills National Forest are included in the sample. As above, the finite population correction is multiplied by the variance estimate: $s^2_{\hat{b}_1}$ $(N - n / N)$. If the slope is significantly less than or greater than 0 and the estimated mean slope is ±0.058 (equivalent to 2-fold increase or decrease in the odds of a watershed having caches present after 12 years; exp[0.058 × 12] = 2, exp[-0.058 × 12] = 0.5), then management action should be taken to examine causes for the change in distribution.

VII. DATA RECORDING AND ARCHIVING

A data form containing the basic information to record when conducting beaver food cache surveys to monitor abundance is provided in Appendix D and a data form to use when monitoring beaver distribution is provided in Appendix E. Recording UTM coordinates at food cache locations will facilitate entry of these data points in NRIS Fauna or other wildlife databases. Further details for counting beaver food caches are found on pages 19–21 of the British Columbia Ministry of Environment, Lands and Parks publication "Inventory Methods for Beaver and Muskrat" (http://srmwww.gov.bc.ca/risc/pubs/tebiodiv/bemu/index.htm).

VIII. LITERATURE CITED

Allen, A. W. 1983. Habitat suitability index models: beaver. FWS/OBS-82/10.30 Revised, U.S. Fish and Wildlife Service,

Anderson, D. R. 2003. Response to Engeman: index values rarely constitute reliable information. Wildlife Society Bulletin 31:288-291.

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29:1284-1297.

Beier, P., and R. G. Barrett. 1987. Beaver habitat use and impact in Truckee River basin, California. Journal of Wildlife Management 51:794-799.

Boyce, M. S., P. R. Vernier, S. E. Nielsen, and F. K. A. Schmiegelow. 2002. Evaluating resource selection functions. Ecological Modelling 157:281-300.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information - theoretic approach, Second edition. Springer-Verlag, New York.

Caughley, G., and A. R. E. Sinclair. 1994. Wildlife ecology and management. Blackwell Science, Cambridge, Massachusetts, USA.

Collins, T. C. 1976. Population characteristics and habitat relationships of beavers, *Castor canadensis*, in northwest Wyoming. Ph.D. Ph.D. University of Wyoming.

Elzinga, C. L., D. W. Salzar, and J. W. Willoughby. 1998. Measuring and monitoring plant populations. BLM Technical Reference 1730-1, Denver, Colorado.

Emme, T. J., and B. A. Jellison. 2004. Managing for beaver on the Bighorn National Forest. Wyoming Game and Fish Department, Cheyenne, Wyoming.

Engeman, R. M. 2003. More on the need to get the basics right: population indices. Wildlife Society Bulletin 31:286-287.

Fryxell, J. M. 2001. Habitat-suitability and source-sink dynamics of beavers. Journal of Animal Ecology 70:310-316.

Gibbs, J. P. 1995. Monitor version 6.2 users manual. Exeter Software, Setauket, New York.

Grasse, J. E., and E. F. Putnam. 1955. Beaver management and ecology in Wyoming. Federal Aid to Wildlife Restoration Project, Bulletin Number 6, Wyoming Game and Fish Department, Cheyenne, Wyoming.

Hay, K. G. 1958. Beaver census methods in the Rocky Mountain region. Journal of Wildlife Management 22:395-402.

Hayward, G. D., D. G. Miquelle, E. N. Smirnov, and C. Nations. 2002. Monitoring Amur tiger populations: characteristics of track surveys in snow. Wildlife Society Bulletin 30:1150-1159.

Hosmer, D. W., and S. Lemeshow. 2000. Applied logistic regression, second edition. John Wiley & Sons, Inc., New York.

Howard, R. J., and J. S. Larson. 1985. A stream classification system for beaver. Journal of Wildlife Management 49:19-25.

Jenkins, S. H., and P. E. Busher. 1979. *Castor canadensis*. Mammalian Species No. 120, American Society of Mammologists, Shippensburg, Pennsylvania.

McComb, W. C., J. R. Sedell, and T. D. Buchholz. 1990. Dam-site selection by beavers in an eastern Oregon basin. Great Basin Naturalist 50:273-281.

Mulder, B. S., B. R. Noon, T. A. Spies, M. G. Raphael, C. J. Palmer, A. R. Olsen, G. H. Reeves, and H. H. Welsh. 1999. The strategy and design of the effectiveness monitoring program for the Northwest Forest Plan. General Technical Report PNW-GTR-437, U.S. Department of Agriculture, Forest Service, Northwest Research Station, Portland, Oregon.

Norris, R. H., E. P. McElravy, and V. H. Resh. 1996. The sampling problem. Pages 184-208 in G. Petts and P. Calow, editors. River biota: dynamics and diversity. Blackwell Science, Cambridge, Massachusetts.

Novak, M. 1987. Beaver. Pages 283-312 in M. Novak, J. A. Baker, M. E. Obbard, and B. Malloch, editors. Wild furbearer management and conservation in North America. Ministry Natural Resources, Toronto, Ontario.

Olson, R., and W. A. Hubert. 1994. Beaver: water resources and riparian habitat manager. University of Wyoming, Laramie, Wyoming.

Payne, N. F. 1981. Accuracy of aerial censusing for beaver colonies in Newfoundland. Journal of Wildlife Management 45:1014-1016.

Retzer, J. L., H. M. Swope, J. D. Remington, and W. H. Rutherford. 1956. Suitability of physical factors for beaver management in the Rocky Mountains of Colorado. Technical Bulletin 2, Colorado Department of Fish and Game,

Robel, R. J., L. B. Fox, and K. E. Kemp. 1993. Relationship between habitat suitability index values and ground counts of beaver colonies in Kansas. Wildlife Society Bulletin 21:415-421.

Satterthwaite, F. E. 1946. An approximate distribution of estimates of variance components. Biometrics Bulletin 2:110-114.

Scheaffer, R. L., W. Mendenhall III, and R. L. Ott. 1996. Elementary survey sampling, 5th edition. Duxbury Press, Boston, Massachusetts.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters, 2nd edition. Charles Griffith, London.

Slough, B. G. 1978. Beaver food cache structure and utilization. Journal of Wildlife Management 42:644-646.

Slough, B. G., and R. M. F. S. Sadleir. 1977. A land capability classification system for beaver (*Castor canadensis* Kuhl). Canadian Journal of Zoology 55:1324-1335.

Smith, D. W. 1999. Beaver survey: Yellowstone National Park 1998. YCR-NR-99-3, Yellowstone Center for Natural Resources, Yellowstone National Park, Wyoming.

Suzuki, N., and W. C. McComb. 1998. Habitat classification models for beaver (*Castor canadensis*) in streams of the central Oregon Coast Range. Northwest Science 72:102-110.

Swenson, J. E., S. J. Knapp, P. R. Martin, and T. C. Hinz. 1983. Reliability of aerial cache surveys to monitor beaver population trends on prairie rivers in Montana. Journal of Wildlife Management 47:697-703.

Thompson, S. K. 1992. Sampling. John Wiley & Sons, Inc., New York.

Thompson, S. K., and G. A. F. Seber. 1994. Detectability in conventional and adaptive sampling. Biometrics 50:712-724.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, San Diego, California.

Table 1. Literature review of important predictors of beaver habitat suitability.

Reference	Location	Important predictors of beaver habitat suitability
Retzer et al. (1956)	Colorado	Excellent habitat was characterized by valley grades of ≤6%; valley width of >46 m; and, rock types of glacial till, schist, or granite. Unsuitable habitat had a valley grade of >15%, and a valley width that was not much wider than the stream itself.
Slough and Sadleir (1977)	British Columbia	Beaver occupancy along lakes and streams was most related to food availability (aspen along lakes and cottonwood along streams).
Beier and Barrett (1987)	eastern California and western Nevada	Increasing stream width and depth and decreasing stream gradient were most influential in beaver habitat use. Food availability added little explanatory power.
Howard and Larson (1985)	Massachusetts	Watershed size (ha) above the colony site, stream width (m) below the final dam, stream gradient, soil drainage class, % hardwood vegetation within 100m of the site center, percentage hardwood vegetation within 200 m of the site center, and percentage abandoned fields within 100 m of the site center all affected beaver colony site longevity
McComb et al. (1990)	Eastern Oregon	Stream reaches with beaver dams were shallower and had a lower gradient than unoccupied reaches. Beaver did not build dams at sites with a rocky substrate. Bank slopes at occupied reaches were not as steep as those at unoccupied reaches. Occupied streams had greater tree canopy cover, especially of thinleaf alder
Suzuki and McComb (1998)	Western Oregon	Beaver built dams in areas with wide valley-floors, low gradient streams, high graminoid cover, low red alder cover, and low shrub cover.
Fryxell (2001)	Ontario, Canada	Beaver abundance was related to food availability.

Table 2. Sampling units identified for monitoring beaver through a food cache index on the Bighorn National Forest, Wyoming. Sampling units are 6th level hydrologic unit code (HUC) watersheds with suitable beaver habitat and are stratified according to probabilities of beaver habitat suitability. A random sample of 15 moderate and 15 high quality sampling units is identified in the strata column. Hydrologic unit code 12 codes and names for are provided because they represent 6th level HUC information.

HUC 12 Code	HUC 12 Name	Stratum
100800100402	Bighorn River-Willow Creek	Non habitat
100800080502	Brockenback Creek	Non habitat
100800160108	East Pass Creek	Non habitat
100901010209	Goose Creek	Non habitat
100800100107	Horse Creek-Shell Creek	Non habitat
100901010205	Lower Big Goose Creek	Non habitat
100800080406	Lower Canyon Creek	Non habitat
100901010110	Lower Quartz Creek	Non habitat
100800100602	Middle Porcupine Creek	Non habitat
100902060303	North Piney Creek	Non habitat
100800100307	Salt Creek	Non habitat
100800160104	West Fork Little Bighorn River	Non habitat
100800100103	Cedar Creek	Low sample
100800160102	Dry Fork Little Bighorn River	Low sample
100902060302	Kearny Creek	Low sample
100800080403	Lower Tensleep Creek	Low sample
100901010207	Middle Goose Creek	Low sample
100902060305	North Prong Shell Creek	Low sample
100800100104	Shell Creek-Cottonwood Creek	Low sample
100901010106	Tongue River-Columbus Creek	Low sample
100800100106	Trapper Creek-Shell Creek	Low sample
100901010206	Upper Little Goose Creek	Low sample
100901010203	West Fork Big Goose Creek	Low sample
100800100105	White Creek	Low sample
100800080402	East Tensleep Creek	Moderate sample
100800100401	Five Springs Creek	Moderate sample
100800080404	Leigh Creek	Moderate sample
100800080602	Long Park Creek	Moderate sample
100800100204	Lower Beaver Creek-Shell Creek	Moderate sample
100901010101	North Tongue River	Moderate sample
100800100101	Shell Creek-Willett Creek	Moderate sample
100800100604	Trout Creek	Moderate sample
100800100305	Upper Bear Creek	Moderate sample
100800100203	Upper Beaver Creek	Moderate sample

Table 2. Continued.

HUC 12 Code	HUC 12 Name	Stratum
100901010204	Upper Big Goose Creek	Moderate sample
100901010201	Upper East Fork Big Goose Creek	Moderate sample
100902050106	Upper Middle Fork Crazy Women Creek	Moderate sample
100901010109	Upper Quartz Creek	Moderate sample
100800160107	West Pass Creek	Moderate sample
100800100603	Deer Creek	Moderate - no sample
100800160103	Little Bighorn River-Red Canyon Creek	Moderate - no sample
100800160101	Little Bighorn River-Wagon Box Creek	Moderate - no sample
100901010107	Little Tongue River	Moderate - no sample
100800160301	Lodge Grass Creek-Line Creek	Moderate - no sample
100901010202	Lower East Fork Big Goose Creek	Moderate - no sample
100800080601	Paint Rock Creek-Trout Creek	Moderate - no sample
100902060301	South Piney Creek	Moderate - no sample
100800160109	Twin Creek	Moderate - no sample
100902060304	Upper Piney Creek	Moderate - no sample
100902060104	Clear Creek-Grommund Creek	High sample
100800100309	Crystal Creek	High sample
100800080606	Lower Medicine Lodge Creek	High sample
100901010104	Lower South Tongue River	High sample
100902060102	Middle Clear Creek	High sample
100902050103	Muddy Creek	High sample
100800080604	Paint Rock Creek-Luman Draw	High sample
100902050107	Poison Creek	High sample
100902060202	Rock Creek-Clear Creek	High sample
100800080605	Upper Medicine Lodge Creek	High sample
100902050101	Upper North Fork Crazy Women Creek	High sample
100902010301	Upper North Fork Powder River	High sample
100800100601	Upper Porcupine Creek	High sample
100800080401	Upper Tensleep Creek	High sample
100901010103	Upper Tongue River	High sample
100901010102	Fool Creek	High - no sample
100902060106	French Creek	High - no sample
100902050102	Middle North Fork Crazy Women Creek	High - no sample
100902060201	North Rock Creek	High - no sample
100800080603	Paint Rock Creek-South Paint Rock Creek	High - no sample
100902060103	Seven Brothers Creek	High - no sample
100800100102	Shell Creek-Granite Creek	High - no sample
100902060101	South Clear Creek	High - no sample
100901010105	Tongue River-Sheep Creek	High - no sample
100800080405	Upper Canyon Creek-Tensleep Creek	High - no sample

Table 3. Sampling units identified for monitoring beaver through a food cache index on the Black Hills National Forest, South Dakota and Wyoming. Sampling units are 6th level hydrological unit code (HUC) watersheds with suitable beaver habitat and are stratified according to probabilities of beaver habitat suitability. A sample of 17 moderate and 23 high quality sampling units and 12 distribution sampling units is identified in the stratum column.

HUC 12 Code	HUC 12 Name	Stratum
101201060201	Cheyenne River-Moss Agate Creek	Non-habitat
101201060203	Cheyenne River-Driftwood Creek	Non-habitat
101201060204	Cheyenne River-Sheep Canyon	Non-habitat
101201060206	Chilson Canyon	Non-habitat
101201060207	Cheyenne River-Little Tepee Creek	Non-habitat
101201060208	Cheyenne River-Tepee Creek	Non-habitat
101201060209	Dry Creek-Cheyenne River	Non-habitat
101201060403	Pleasant Valley-Red Canyon Creek	Non-habitat
101201060404	Nitche Spring-Red Canyon Creek	Non-habitat
101201060405	Hawkwright Creek	Non-habitat
101201060406	Craven Canyon	Non-habitat
101201060407	White Draw-Red Canyon Creek	Non-habitat
101201060509	Angostura Reservoir-Horsehead Creek	Non-habitat
101201070308	Line Creek-Beaver Creek	Non-habitat
101201070402	Beaver Creek-Rats Valley Creek	Non-habitat
101201070405	Whoopup Creek	Non-habitat
101201070406	Roby Canyon	Non-habitat
101201070407	Beaver Creek-Rock Canyon	Non-habitat
101201070502	Middle Pass Creek	Non-habitat
101201070503	Teepe Canyon	Non-habitat
101201070504	Lower Pass Creek	Non-habitat
101201070505	Pass Creek-East Pass Creek	Non-habitat
101201090101	Upper Cold Brook	Non-habitat
101201090102	Lower Cold Brook	Non-habitat
101201090103	Cottonwood Springs Creek	Non-habitat
101201090104	Hot Brook	Non-habitat
101201090105	Fall River	Non-habitat
101201090301	Slate Spring Draw-Cheyenne River	Non-habitat
101201110604	Pleasant Valley Creek	Non-habitat
101201110605	Morris Creek	Non-habitat
101202010801	Upper Inyan Kara Creek	Non-habitat
101202020105	Horse Creek-Belle Fourche River	Non-habitat
101202020703	Upper Spring Creek	Non-habitat
101201090201	Upper Beaver Creek	High-quality abundance - sample
101201090202	Middle Beaver Creek	High-quality abundance - sample
101201090401	South Fork Lame Johnny Creek	High-quality abundance - sample
101201090601	Ruby Creek-French Creek	High-quality abundance - sample

Table 3. Continued.

HUC 12 Code	HUC 12 Name	Stratum
101201090602	Stockade Lakes-French Creek	High-quality abundance - sample
101201090902	Newton Fork	High-quality abundance - sample
101201090903	Newton Fork-Spring Creek	High-quality abundance - sample
101201090904	Sheridan Lake-Spring Creek	High-quality abundance - sample
101201090905	Johnson Gulch-Spring Creek	High-quality abundance - sample
101201100101	North Fork Rapid Creek	High-quality abundance - sample
101201100109	Slate Creek	High-quality abundance - sample
101201100201	Victoria Creek-Rapid Creek	High-quality abundance - sample
101201110304	Jim Creek-Boxelder Creek	High-quality abundance - sample
101201110601	Town of Roubaix-Elk Creek	High-quality abundance - sample
101201110602	Little Elk Creek-Elk Creek	High-quality abundance - sample
101202010903	Blacktail Creek-Belle Fourche River	High-quality abundance - sample
101202010906	Beaver Creek-Lame Jones Creek	High-quality abundance - sample
101202020701	Park Creek	High-quality abundance - sample
101202030101	Upper Redwater Creek	High-quality abundance - sample
101202030102	South Redwater Creek	High-quality abundance - sample
101202030105	Crow Creek-Redwater Creek	High-quality abundance - sample
101202030302	Middle Spearfish Creek	High-quality abundance - sample
101202030402	Upper False Bottome Creek	High-quality abundance - sample
101201070401	Beaver Creek-Bear Run	Mod-quality abundance - sample
101201090803	Deadman Gulch Creek-Battle Creek	Mod-quality abundance - sample
101201090901	Headwaters Spring Creek	Mod-quality abundance - sample
101201090906	Rockerville Gulch-Spring Creek	Mod-quality abundance - sample
101201100102	South Fork Rapid Creek	Mod-quality abundance - sample
101201100103	Silver Creek-Rapid Creek	Mod-quality abundance - sample
101201100104	Upper Castle Creek	Mod-quality abundance - sample
101201100105	South Fork Castle Creek	Mod-quality abundance - sample
101201100106	Deerfield Lake-Castle Creek	Mod-quality abundance - sample
101201100107	North Fork Castle Creek	Mod-quality abundance - sample
101201100108	Lower Castle Creek	Mod-quality abundance - sample
101201110303	Estes Creek-Boxelder Creek	Mod-quality abundance - sample
101202020108	Oak Creek	Mod-quality abundance - sample
101202020702	Boulder Creek	Mod-quality abundance - sample
101202030301	Upper Spearfish Creek	Mod-quality abundance - sample
101202030303	Little Spearfish Creek	Mod-quality abundance - sample
101202030405	North Fork Hay Creek	Mod-quality abundance - sample
101201090203	Highland Creek	High-quality distribution - sample
101201090603	Glen Erin Creek-French Creek	High-quality distribution - sample
101201090801	Grizzly Bear Creek-Battle Creek	High-quality distribution - sample
101201090802	Iron Creek	High-quality distribution - sample
101201100110	Pactola Reservoir-Rapid Creek	High-quality distribution - sample
101201110301	North Boxelder Creek-Boxelder Creek	High-quality distribution - sample
101202010707	Lytle Creek	High-quality distribution - sample
101202010907	Lame Jones Creek	High-quality distribution - sample

Table 3. Continued.

HUC 12 Code	HUC 12 Name	Stratum
101202020102	Belle Fourche River-Deep Creek	High-quality distribution - sample
101202030304	Lower Spearfish Creek	High-quality distribution - sample
101201110302	South Boxelder Creek-Boxelder Creek	Mod-quality distribution - sample
101202030203	Sand Creek	Mod-quality distribution - sample
101201090804	Upper Grace Coolidge Creek	High-quality no-sample [a]
101201090805	Lower Grace Coolidge Creek	High-quality no-sample [a]
101202010705	Miller Creek-Arch Creek	High-quality no-sample [a]
101202020107	Deep Creek-Pine Creek	High-quality no-sample
101201060401	Fourmile Creek	Mod-quality no-sample [a]
101201060402	Lightning Creek-Red Canyon Creek	Mod-quality no-sample [a]
101201070203	Upper Oil Creek	Mod-quality no-sample
101201070501	Upper Pass Creek	Mod-quality no-sample
101201100202	Canyon Lake-Rapid Creek	Mod-quality no-sample
101201110305	Blackhawk Creek-Boxelder Creek	Mod-quality no-sample
101201110603	Stagebarn Canyon Creek	Mod-quality no-sample
101202010806	Hudson Creek	Mod-quality no-sample
101202020207	Upper Whitewood Creek	Mod-quality no-sample
101202020901	Headwaters Alkali Creek	Mod-quality no-sample
101202030103	Sundance Creek	Mod-quality no-sample
101202030201	Cold Springs Creek	Mod-quality no-sample
101202030202	Grand Canyon	Mod-quality no-sample
101202030204	Red Canyon Creek	Mod-quality no-sample
101202030205	Bear Gulch	Mod-quality no-sample
101202030403	Polo Creek	Mod-quality no-sample
101202030406	South Fork Hay Creek	Mod-quality no-sample

[a] Sampled during Fall 2007 monitoring season as auxillary samples

Table 4. Number of beaver sampling units by stratum on the Bighorn National Forest, Wyoming and Black Hills National Forest, South Dakota and Wyoming. Sampling units are 6th level hydrologic unit code watersheds.

Stratum	Objective	Number of sampling units	
		Bighorn	Black Hills
Low – sample	Distribution	12	NA[a]
Moderate – sample	Abundance	15	17
	Distribution	0	2
High – sample	Abundance	15	23
	Distribution	0	10
Moderate – no sample		10	17
High – no sample		10	4
Non habitat		12	33
Total		74	106

[a]Stratum not used in the Black Hills National Forest beaver monitoring

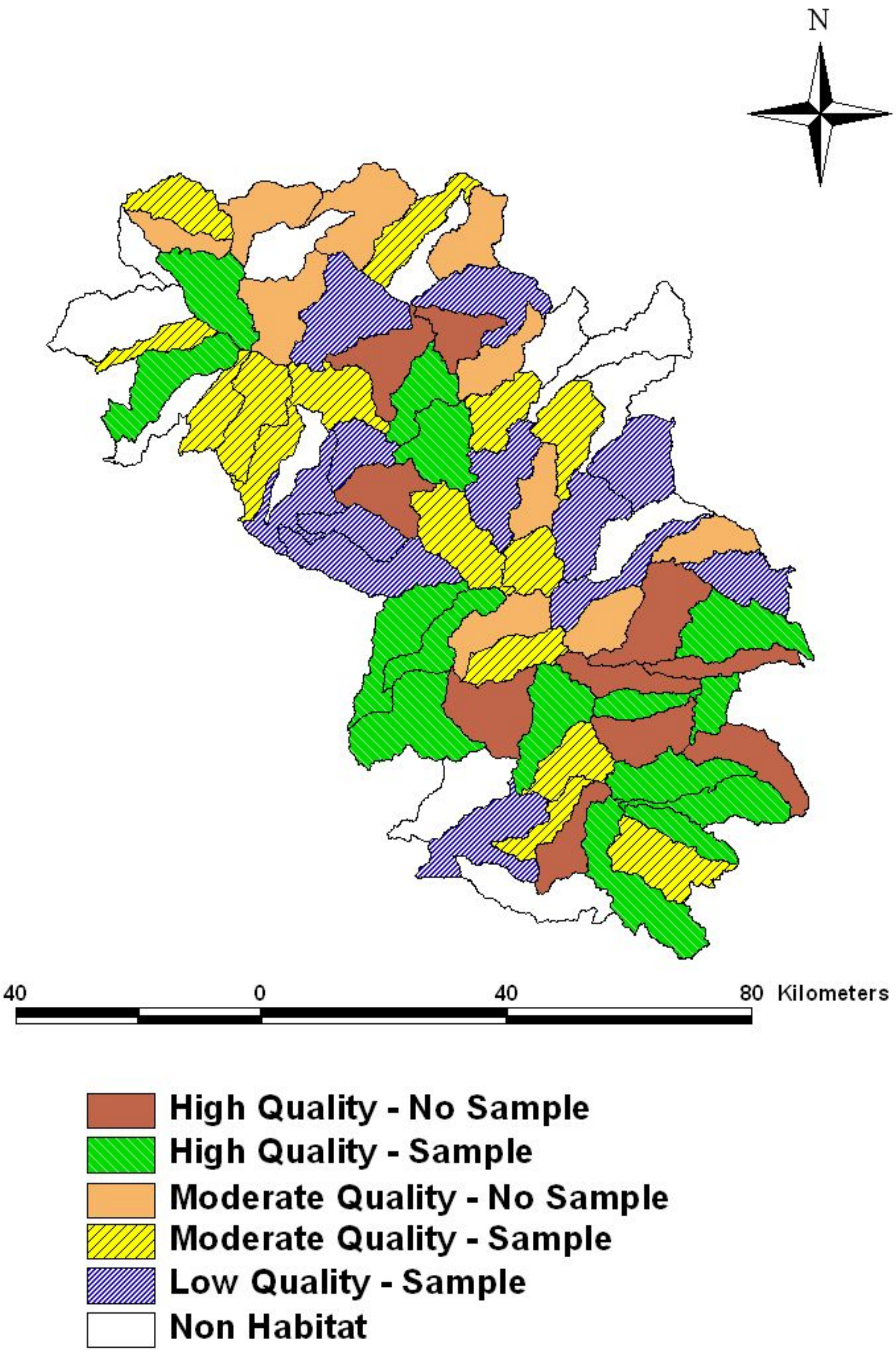

Figure 1. Proposed sampling units to monitor abundance of beaver using food cache density on the Bighorn National Forest, Wyoming. Sampling units are 6th level hydrologic unit codes that have been designated as strata.

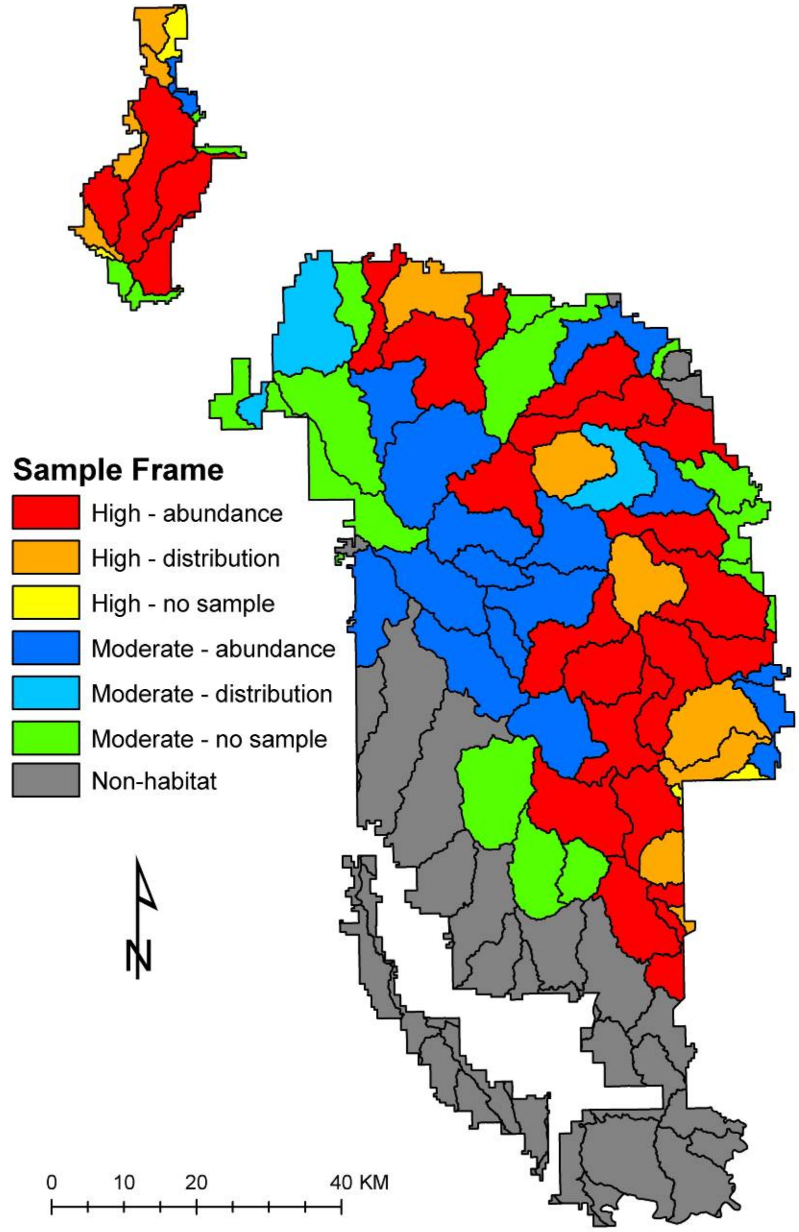

Figure 2. Proposed sampling units within high- and moderate-quality strata to monitor abundance of beaver using food cache density and distribution (cache presence-absence) on the Black Hills National Forest, South Dakota and Wyoming. Sampling units are 6th level hydrologic unit codes that have been designated as high- and moderate-quality strata.

APPENDIX A. HABITAT SUITABILTY MODELING TO DEFINE STRATA

A.0 Methods

We used data collected on the 448,259 ha (1,107,671 ac) Bighorn National Forest in fall 2003, and data collected on the 485,623 ha (1,200,000 ac) Black Hills National Forest in September 2004 to model beaver habitat suitability and ultimately to define sampling strata on each Forest. On the Bighorn National Forest, a helicopter and fixed wing aerial survey was conducted to locate beaver food caches from 26 August to 23 October, 2003. The Bighorn survey located 60 active food caches and 106 historical lodges (Emme and Jellison 2004). Among the 6th level HUCs on the Bighorn National Forest that contain perennial water in streams, rivers, or other water bodies, there are 66, 6th level HUCs with potential beaver habitat. An analysis provided by GIS personnel on the Bighorn National Forest identified 74, 6th level HUCs on the Forest, 65 (88%) of which contained perennial streams or rivers within the boundaries of the Bighorn National Forest. We modeled suitable beaver habitat on the Bighorn National Forest with 154 beaver locations identified in 2003 within the boundaries of the Bighorn National Forest and 500 random locations across the Forest.

A helicopter survey conducted from September 14 to 16, 2004 on the Black Hills National Forest resulted in locating 74 active beaver lodges on Forest Service lands (K. Burns. 2004. Beaver Survey Report. Black Hills National Forest, South Dakota and Wyoming, Black Hills National Forest, Custer, South Dakota, USA). We used the 74 beaver lodge locations identified in 2004 and 400 random locations to model beaver habitat suitability on the Black Hills National Forest.

For each model, we developed a logistic regression model to stratify 6th level HUCs into non habitat, low-quality habitat, moderate-quality habitat, and high-quality habitat strata for the

Bighorn National Forest, and non habitat, moderate-quality, and high-quality strata for the Black Hills National Forest, based on criteria indicative of beaver habitat suitability. All data used to evaluate habitat suitability were obtained from geographic coverages provided by GIS personnel with the Bighorn and Black Hills National Forests. Predictor variables included elevation (m), slope (gradient), elevation deviation (m), distance (m) to water source (perennial), and distance to aspen and/or willow. Elevation was recorded for each 30-m pixel (USGS National Elevation Dataset). Percentage slope was calculated using a 3 × 3 neighborhood window centered on each cell. We determined for the center cell, the down slope direction, or the greatest rate of change within a window. The rate of change in elevation values was then calculated and assigned to the center cell. We repeated this analysis for every grid cell within the analysis window with units expressed as percentages. We measured the distance to perennial streams and water bodies and distances to aspen and/or willow with the Euclidean distance (i.e., right angle distance) from each cell to the nearest water feature or vegetation type coded as aspen or willow. We calculated elevation deviation to estimate the topographic complexity in the area surrounding a grid cell with a 5 × 5 neighborhood window. For each cell, the standard deviation of elevation values within the 5 × 5 window was calculated and assigned to the center cell. Neighborhoods of high complexity have higher standard deviations relative to those areas of lower topographic complexity.

We used independent sample *t*-tests on raw data to test for differences in habitat variables at beaver and random locations (Proc TTEST; SAS Institute 2001). We evaluated equality of variances with the Folded F method and used the Satterthwaite (1946) method to calculate *t*-values in those instances where variances were unequal. The Satterthwaite statistic is an approximate *t* statistic and is used if the population variances of two groups are unequal. We computed degrees of freedom for this statistic with the Satterthwaite (1946) approximation.

We used binary logistic regression (Proc Logistic; SAS Institute 2001) to model suitable beaver habitat on each Forest, where we coded the beaver lodge or food cache locations as 1s and random points as 0s. We assessed the strength of evidence for several candidate models with Akaike's information criterion for small samples (AIC_c) (Burnham and Anderson 2002). We selected the model with the lowest AIC_c value as the best-fitting model, and we used Δ_i, the difference between AIC_c for the best model and AIC_c for the *i*th candidate model, to identify models competing with the best model. Akaike weights (w_i) allowed us to assess the weight of evidence in favor of each model (Burnham and Anderson 2002).

We performed a 5-fold cross validation to evaluate goodness-of-fit of our beaver habitat suitability models for both Forests (Boyce et al. 2002). For each Forest, independent model validation was done using k-fold cross-validation. The data set was partitioned into k = 5 sets, and the best model was fitted to 80% of the dataset (4 of 5 sets) and the remaining 20% was used for cross-validation. The cross-validated dataset was partitioned into five bins, and Spearman rank correlation was used to compare the association between the median (independently) predicted probability of occurrence and the percentage of observations with beaver present among bins. This process was repeated five times for each 20% of the original dataset, and correlations were averaged to test for model fit. Strong positive correlations would indicate our models had good predictive performance because more suitable beaver habitat would fall into higher ranked probability bins (Boyce et al. 2002).

We used the parameter estimates from the best model to compute probabilities of beaver habitat suitability across the Forest with the logistic function, $\exp(\beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4)$

/ $(1 + \exp[\beta_o + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4])$. We used 6th level HUCs on each Forest as analysis regions to clip and subsequently summarize beaver habitat suitability by dividing the probabilities (i.e., *P*-values) on Forest Service lands in each HUC into quartiles.

For each Forest we used similar criteria to define sampling strata. We designated the non habitat stratum where at least 95% of Forest Service area in each HUC was dominated by probabilities in quartile 1. For the Bighorn National Forest, we designated the low quality habitat stratum as all 6th level HUCs where 85.0 to 94.9% of Forest Service area was dominated by probabilities in quartile 1. We did not define a low quality stratum for the Black Hills National Forest. To define moderate and high quality strata we used an index where probabilities in each quartile were multiplied as: Index = $0 \cdot Q1 + 1 \cdot Q2 + 5 \cdot Q3 + 10 \cdot Q4$. We then ranked the index values for the HUCs not previously classified as non-habitat or low quality strata and designated the top 50% HUCs as high quality stratum and the bottom 50% as moderate quality stratum.

A.1 Results

Beaver locations in the Bighorn National Forest were closer to suitable vegetation and a perennial water source, situated at lower elevations and slopes, and in areas with lower topographic complexity than at random (Table A.1). The same patterns were observed for beaver locations on the Black Hills National Forest with the exception that topographic complexity and slope did not differ between beaver and random locations (Table A.2). We reason that these differences are related to the fact that topography is more rolling and less abrupt on the Black Hills than on the Bighorn National Forest.

The best model identifying suitable beaver habitat on the Bighorn and Black Hills National Forests was the model with predictor variables including percent slope, distance to water (perennial streams on the Black Hills National Forest), distance to aspen and/or willow, and elevation deviation (Tables A.3 and A.4). Parameter estimates were used to predict habitat suitability across the Forests (Table A.5). The northern portion of the Black Hills National Forest contained more suitable beaver habitat than in the southwest portion of the Forest (Figure A.1). Our cross validation analyses indicated the best model was a strong, positive predictor of beaver habitat suitability on the Bighorn ($r_s = 0.82$, $P = 0.004$, $n = 10$) and Black Hills National Forests ($r_s = 0.96$, $P < 0.001$, $n = 10$). After the distribution of predicted probabilities was summarized into quartiles (Tables A.6 and A.7), watersheds were classified into suitability strata based on index values (Figure A.2).

Table A.1. Mean (±1 SE) habitat characteristics at 154 beaver and 500 random location pixels, Bighorn National Forest, Wyoming, 2003. Independent sample *t*-tests evaluated differences between beaver and random locations.

Habitat variables	Beaver	Random	*t*	*Df*	*P*
Distance (m) to vegetation[a]	754 ± 93	1,940 ± 67	–10.38	327	<0.001
Distance (m) to water[b]	167 ± 9	319 ± 10	–11.15	496	<0.001
Elevation (m)	2,539 ± 13	2,607 ± 17	–3.23	592	0.001
Elevation deviation (m)	3.6 ± 0.2	6.6 ± 0.2	–9.34	525	<0.001
Slope (%)	13.8 ± 0.9	25.6 ± 1.0	–9.18	519	<0.001

[a] aspen and/or willow.

[b] stream or water body.

Table A.2. Mean (±1 SE) habitat characteristics at 74 active beaver dam and 400 random location pixels, Black Hills National Forest, South Dakota and Wyoming, 2004. Independent sample t-tests evaluated differences between beaver and random locations.

Habitat variables	Beaver	Random	*t*	*df*	*P*
Distance (m) to vegetation[a]	558 ± 68	3,296 ± 313	–3.76	472	<0.001
Distance (m) to stream[b]	452 ± 105	2,062 ± 107	–6.38	472	<0.001
Elevation (m)	1,536 ± 20	1,618 ± 12	–2.74	472	0.006
Elevation deviation (m)	13.7 ± 0.6	13.3 ± 0.4	0.45	472	0.655
Slope (%)	18.1 ± 1.4	19.9 ± 0.7	–1.10	107	0.272

[a] aspen and/or willow.

[b] perennial stream.

Table A.3. Model selection results for suitable beaver habitat (n = 654), on the Bighorn National Forest, Wyoming, 2003. Models are based on 154 beaver location pixels and 500 random pixels and are listed according to the model that best fits the data and ranked by ΔAIC_c, the difference between the model with the lowest Akaike's information criterion for small samples (AIC_c) and the AIC_c for the current model. The strength of evidence for each model is assessed with Akaike weights (w_i). Model fit is described with the value of the maximized log-likelihood function (log[L]) and the number of parameters (K).

Model[a]	Log(L)	K	AIC_c	ΔAIC_c	w_i
Slope + Veg + Stream + ED	–257.43	5	524.951	0.000	0.689
Slope + Veg + Stream	–259.24	4	526.539	1.588	0.311
Stream	–305.75	2	615.520	90.570	0.000
Veg	–320.01	2	644.040	119.090	0.000
ED	–323.24	2	650.495	125.545	0.000
Slope	–324.40	2	652.825	127.875	0.000
Elevation	–354.53	2	713.078	188.128	0.000
Null	–356.96	1	715.921	190.970	0.000

[a]Explanatory variables used are percentage slope (Slope), distance (m) to nearest aspen and/or willow (Veg), distance (m) to nearest perennial stream (Stream), elevation deviation (ED), and elevation (m) at each beaver or random pixel.

Table A.4. Model selection results for suitable beaver habitat (n = 474), on the Black Hills National Forest, South Dakota and Wyoming, 2004. Models are based on 74 active beaver dam location pixels and 400 random pixels and are listed according to the model that best fit the data and ranked by ΔAIC_c, the difference between the model with the lowest Akaike's information criterion for small samples (AIC_c) and the AIC_c for the current model. The strength of evidence for each model is assessed with Akaike weights (w_i). Model fit is described with the value of the maximized log-likelihood function (log[L]) and the number of parameters (K).

Model[a]	Log(L)	K	AIC_c	ΔAIC_c	w_i
Slope + Riparian + Water + Elev.	–137.42	5	284.961	0.000	1.000
Slope + Riparian + Water	–148.22	4	304.518	19.557	0.000
Water	–164.90	2	333.831	48.870	0.000
Riparian	–182.32	2	368.658	83.697	0.000
Elevation	–201.56	2	407.138	122.177	0.000
Null	–205.33	1	412.659	127.698	0.000
Slope	–204.75	2	413.527	128.566	0.000

[a]Explanatory variables used are percentage slope (Slope), distance (m) to nearest aspen and/or willow (Riparian), distance (m) to nearest perennial stream or water body (Water), and elevation (m) at each beaver or random pixel.

Table A.5. Parameter estimates (±1 SE) for best logistic regression model describing beaver habitat suitability on the Black Hills National Forest, South Dakota and Wyoming.

Variable	Parameter estimate	Standard error
Intercept	7.0647	1.5030
Slope (%)	-0.0277	0.0118
Distance to vegetation (m)	-0.0011	0.0002
Distance to stream (m)	-0.0014	0.0002
Elevation (m)	-0.0038	0.0009

Table A.6. Bighorn National Forest – probabilities of beaver habitat suitability

Number	Quartile	Range of *P*-values
1	0.0–24.9	0.000000000–0.013595855
2	25.0–49.9	0.013595855–0.074777203
3	50.0–74.9	0.074777203–0.305906738
4	75.0–100	0.305906738–0.866735756

Table A.7. Black Hills National Forest – probabilities of beaver habitat suitability

Number	Quartile	Range of *P*-values
1	0.0–24.9	0.000000000–0.025267522
2	25.0–49.9	0.025267522–0.093850798
3	50.0–74.9	0.093850798–0.245455593
4	75.0–100	0.245455593–0.920459747

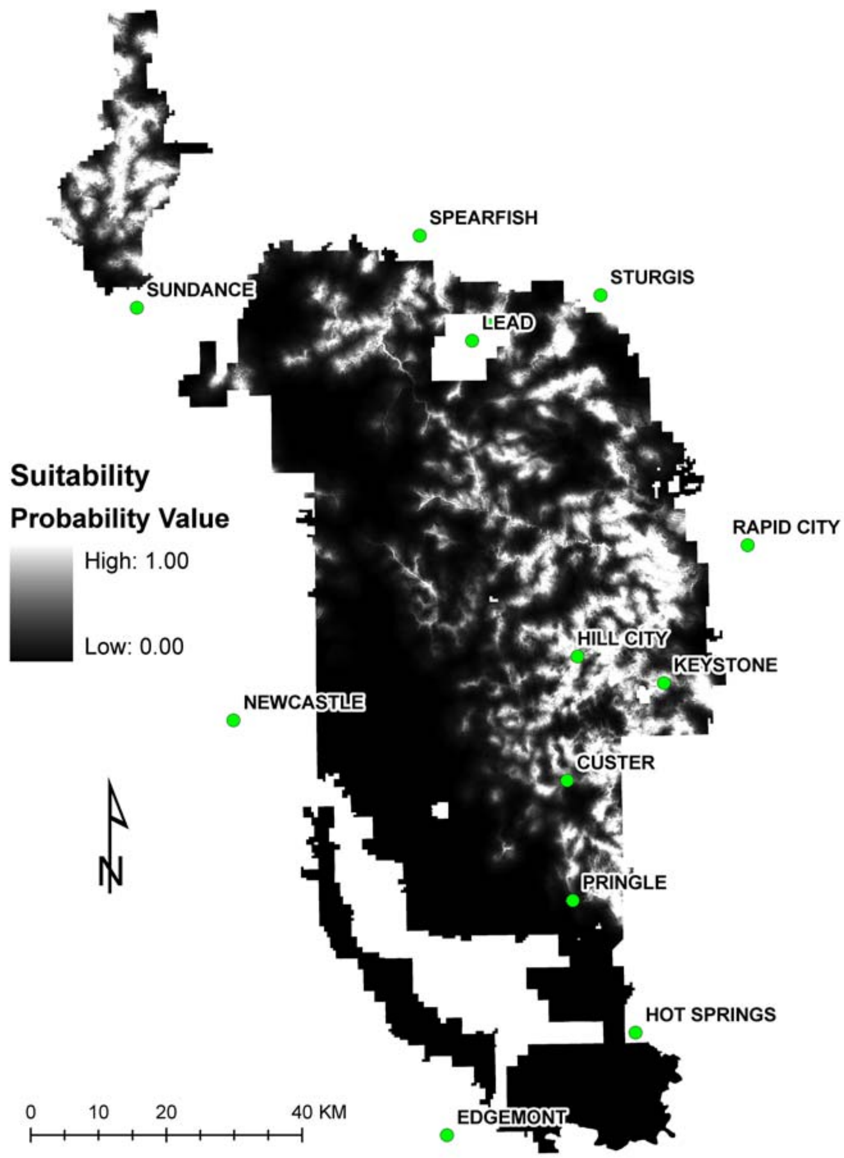

Figure A.1. Predicted beaver habitat suitability for the Black Hills National Forest

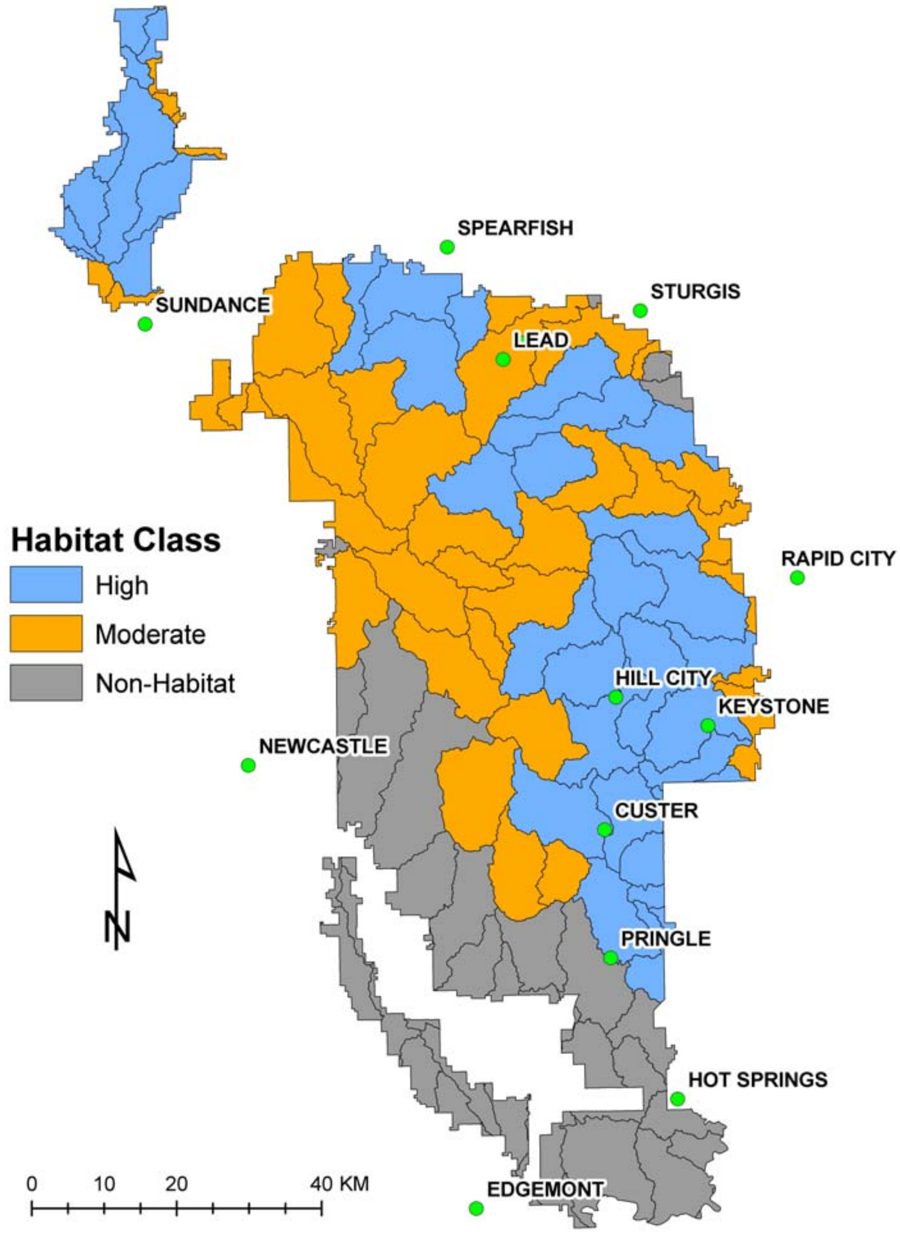

Figure A.2. Habitat suitability of watersheds for beaver on the Black Hills National Forest.

APPENDIX B. APPROACHES TO ESTIMATE DETECTION PROBABILITY OF BEAVER FOOD CACHES

Counts of beaver food caches made during aerial surveys likely under-represent the true number of food caches, i.e., detection probabilities of food caches are less than 1 (Swenson et al. 1983). Thus, counts of food caches during aerial surveys represent an index of abundance. Indices can be problematic because they assume equal detection probabilities across time, habitat types, and observers (Anderson 2001, Anderson 2003). Changes in detection probabilities over time can cause index values to change over time despite no change in abundance of food caches. For example, changes in detection probability over time might result from temporal changes in habitat or increased experience of observers conducting surveys. If it is not feasible to estimate detection probability during every survey, it can periodically be estimated to determine if it is changing over time.

There are several different methods for estimating detection probabilities of beaver food caches. Each method varies in its level of sampling intensity and degree of precision.

B.0 Ground Count versus Aerial Count

One method for estimating detection probability of beaver food caches is to compare ground counts of food caches to counts made during aerial surveys. Probability of detection is then estimated by dividing aerial counts by ground counts:

$$\hat{p} = \frac{count_{aerial}}{count_{ground}}$$

This is the best estimate of detection probability if it can be assumed that 100% of caches are observed during ground counts. Multiple estimates can be made to make comparisons between habitat, watershed characteristics, or observers; that is, factors that affect detection probability. There are two ways to estimate detection probability using this method.

One way is to conduct aerial surveys and ground counts at the level of an individual watershed. First, an aerial survey is performed were all food caches are counted within the watershed. Then, a ground count is conducted by walking the entire length of stream in the watershed. If the same observer is used for both counts, it is important to conduct the aerial survey first so that the observer does not know the location of food caches from the ground survey. This would increase detection probability and result in a biased estimate.

A second method is to determine if individual food caches can be detected during an aerial survey. Here, aerial surveys would be conducted over known food caches to determine if individual caches can be detected. Then, the fraction of known food caches detected during aerial surveys can be used as an estimate of detection probability. However, it is important that the observer conducting the aerial survey is naïve to the location of the food cache because knowledge of the food cache location will result in higher, and thus biased, estimates of detection probability.

B.1 Capture – Recapture with Two Observers

Capture – recapture estimators of abundance can also be used to estimate the number of beaver food caches within a watershed. Using this method, two observers would independently count beaver food caches during an aerial survey, or possibly during two separate surveys, and note their location. Food caches counted by one observer would be considered the initial marking period during a capture-recapture event, and counts made by the second observer would represent the recapture period. These data would be used in a capture-recapture estimator to determine the number of food caches present. There are several capture-recapture abundance estimators. A simple estimator is the Petersen index (Seber 1982):

$$\hat{N} = \frac{MC}{R}$$

where $\hat{N}$ = the estimated number of beaver food caches present; M = number of food caches observed by observer 1; C = number of food caches observed by observer 2; and R = number of food caches observed by both observer 1 and 2. The estimated number of food caches can be used to estimate detection probabilities for each observer:

$$\hat{p}_{observer1} = \frac{M}{\hat{N}}$$

$$\hat{p}_{observer2} = \frac{C}{\hat{N}}$$

or an average detection probability between observers can be estimated as:

$$\hat{\bar{p}} = \frac{M + C}{2\left(\hat{N}\right)}$$

B.2 Adjusting Counts with Detection Probability

One way to alleviate the problem of unequal detection probabilities among habitats over time is by adjusting food cache counts by detection probability. Dividing counts by detection probability gives an unbiased estimate of beaver food cache abundance:

$$\hat{N} = \frac{C}{\hat{p}}$$

where $\hat{N}$ = unbiased estimate of food cache abundance; C = number of caches counted during aerial survey; and $\hat{p}$ = estimated detection probability. Variances associated with $\hat{N}$ determined this way are readily available from several sources (Thompson 1992, Thompson and Seber

1994). Estimates of detection probability could also be used to retrospectively adjust counts of food caches from past surveys given that the estimates apply to previous monitoring years.

B.3 Factors Affecting Detection Probability

The probability of detecting beaver food caches during aerial surveys can be influenced by several factors. Observers conducting surveys may differ in their ability to detect caches. Detection probability may also differ among types of riparian vegetation, valley shape, or time of day. If counts of food caches are to be adjusted for detection probability, then estimates of detection probability specific to observers or habitat types may be required. Alternatively, detection probability could be modeled as a function of these important factors using logistic regression. Then, the model-predicted probabilities of detection could be used to adjust each count within watersheds based on the characteristics of that watershed. This modeling approach is useful (Thompson and Seber 1994), but we recommend that a statistician be consulted if it is to be used.

APPENDIX C. BEAVER FOOD CACHE DENSITIES (CACHE / KM) AMONG STRATA AND SAMPLES SIZE ESTIMATES – ANALYSIS OF 2007 MONITORING DATA FROM THE BLACK HILLS NATIONAL FOREST.

In fall 2007 the Black Hills National Forest sampled 23 watersheds in the high-quality stratum and 17 in the moderate-quality stratum. Twelve additional watersheds were monitored for beaver cache distribution. These data were used to: 1) compare mean cache densities and proportion of watersheds with beaver food caches between habitat strata, and 2) determine the sample sizes needed to estimate beaver food cache density and the proportion of watersheds with food caches present at various levels of precision.

C.0 Cache Densities and Distribution Between Strata

Beaver monitoring is based on 6^{th} level hydrologic unit code watersheds, and these watersheds on the Black Hills National Forest were classified as: non-habitat, moderate-quality habitat, and high-quality habitat. Watersheds considered to be non-habitat were excluded from the sampling frame, and only watersheds in the moderate- and high-quality strata were selected for monitoring beaver food cache abundance and distribution. Stratification is intended to increase the precision of the Forest-wide estimate of food cache density and facilitate comparisons between watersheds representing different habitat quality for beaver.

The densities of beaver food caches were similar between high-quality and moderate-quality watersheds. The average density (±1 SD) in 23 high-quality watersheds was 0.030 (± 0.053) caches / km. The average density in 19 moderate-quality watersheds was 0.021 (± 0.043) caches / km. Variances did not differ between strata (Variance ratio test; $F = 1.47$; df = 22, 26; P

= 0.431), and there was no significant difference in food cache densities between strata (one-tailed t-test; t = 1.197; df = 38; P = 0.119). Twelve additional watersheds were sampled for beaver food cache distribution. Including watersheds sampled for abundance, beaver food caches were present in 20 of 52 watersheds (proportion = 0.38; 1 SE = 0.02). In all, food caches were observed in 15 of 33 high-quality watersheds and 5 of 19 low-quality watersheds. There was no difference in the proportion of watersheds with beaver food caches between strata (Chi-square test; X^2 = 1.87; df = 1; P = 0.172).

Although no differences in food cache densities were observed between moderate- and high-quality watersheds, retaining the stratified monitoring framework is still useful. Monitoring trends in food cache between different strata allows comparison of trends between strata. For example, food cache abundance may increase over time in the high-quality stratum but remain constant in the moderate-quality stratum. Stratification may also improve trend detection capability across the Forest. Forest-wide trends will be estimated by estimating trends in cache density within a watershed and then averaging trends to make inference regarding Forest-wide trend (limited to sampling frame). Stratification may reduce the variance in trends across the Forest and, consequently, improve trend detection capability. However, there is no current information on the variation in trend estimates within a stratum. Thus, it is recommended that the stratified sampling be retained.

C.1 Sample Size Estimates

Detecting changes in the abundance of animal populations requires that their abundance for any monitoring time period be precisely estimated. The protocol states that existing data should be used to determine the number of watersheds (i.e., sample size) needed to estimate food cache densities within a certain level of precision. This information can help to determine if the sample sizes currently used are adequate to precisely estimate food cache densities and food cache distribution within a monitoring year.

Sample sizes were computed according to the sample size formulas in Section 3.2 of this protocol for several levels of precision. A sample size of 68 watersheds is needed to be within 5% of the mean food cache density across the Forest, and a sample size of 28 is needed to be within 30% of the mean (Table C.1). The sample size needed to be within 0.05 of the estimated proportion is 61, whereas the sample size needed to be within 0.30 of the estimated proportion is 9 (Table C.1).

Table C.1. Number of watersheds (sample size) that require sampling to estimate beaver food cache density within a certain percentage of the true mean (bound of error), and the number of watersheds (sample size) that require sampling to estimate the proportion of watersheds with beaver food caches present within a certain error of the true proportion.

Food cache density		Proportion with caches present	
Bound of error	Sample size	Bound of error	Sample size
5%	68	0.05	61
10%	60	0.10	41
15%	51	0.15	26
20%	42	0.20	18
30%	28	0.30	9

The current sampling scheme allows the mean food cache density among watersheds to be estimated within ~20% of the true [unknown] mean. It also allows the proportion of watersheds with food caches present to be estimated within 0.05-0.10 of the true [unknown] proportion. Therefore, the current sampling scheme allows good estimates of beaver food cache density among watersheds and good estimates of the proportion of watersheds that have food caches present. This should allow trends in beaver abundance and distribution to be detected with a reasonable amount of certainty.

APPENDIX D. SUGGESTED FORM WHEN RECORDING FOOD CACHES TO MONITOR TRENDS IN BEAVER ABUNDANCE.

Food Cache	Sampling unit name/number	Date of flight	Time of Location	UTM East Coordinate	UTM North Coordinate
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					

APPENDIX E. SUGGESTED FORM WHEN RECORDING FOOD CACHES TO MONITOR TRENDS IN BEAVER DISTRIBUTION.

Sampling unit name/number	Present (1)[a] Absent (0)	Date of flight	Time of Location	UTM East Coordinate[b]	UTM North Coordinate
Total	[c]				

[a]To indicate presence, record a 1 when the first beaver food cache is located in each sampling unit. Record a 0 to indicate absence when no beaver food caches are observed after entirely searching each respective sampling unit.

[b]UTM coordinates are only recorded when food caches are present.

[c]Total is $\sum y_i$ used to estimate proportion of sampling units occupied by beaver.