Government of Northwest Territories Gouvernement des Territoires du Nord-Ouest

Optimal Survey Design, Survey Intervals and Analysis Strategies for Caribou Calving Ground Surveys, Reconnaissance Surveys and Composition Surveys

John Boulanger
Integrated Ecological Research, Nelson, BC

2021

Manuscript Report No. 289

The contents of this paper are the sole responsibility of the author.

ABSTRACT

This report provides statistical commentary on survey methods used to monitor barren-ground caribou herds in the Northwest Territories (NWT) with an emphasis on intervals between surveys and survey precision. I provide some comments on the statistical design of surveys and some ideas to improve precision.

Calving ground photo surveys and post-calving photo surveys are the most important surveys for barren-ground caribou as they provide benchmarks for herd status and management. Results of power analyses suggest that the sampling interval for these surveys should never be less than three years unless a very large change in abundance is expected. For the most likely rates of change in population size (+/- 10% per year) then a survey interval of five to six years is adequate.

Composition surveys in June, fall (rut, usually late October), and late winter (March/April) are used to assess initial calf productivity, calf survival to four to five months, calf survival to nine to ten months, and sex ratio (in the fall). Representative sampling across a herd's range is key to obtaining reliable results. Late-winter surveys are best carried out annually to capture frequently high year-to-year variation. Fall surveys to assess sex ratio are usually carried out in years of calving ground photo surveys (every three years in most NWT herds 2006-2018) and may be conducted more often if a substantial male-dominated harvest is in place.

Reconnaissance surveys on the calving grounds of some herds have been used to assess trend in caribou abundance on calving grounds in years between full photographic surveys. They are much simpler and far less costly. However, variance on survey results is usually high and assessment of composition (breeding cows, non-breeding cows, yearlings and bulls) on or near the calving grounds may not be reliable. I provide recommendations to improve precision of these surveys.

I note that the primary analyses in this report occurred in 2011 and since then some of the methodologies have evolved. I provide updated citations to this current work.

PREFACE

The analyses described in this report were carried out in 2011 to assist biologists with the Government of the Northwest Territories (GNWT) Department of Environment and Natural Resources (ENR) in planning population surveys and other monitoring surveys of barren-ground caribou herds. The work was carried out under contract with ENR ungulate biologist Jan Adamczewski. The main focus was on how often these surveys should be carried out. A recommended survey interval depends on the risk status of the herd and in part on the likelihood of management decisions based on herd status (e.g. harvest limits). The ability to detect population change in a caribou herd depends on the rate of change but also on the precision of survey results, thus the design of surveys to achieve good precision must also be considered. Since 2011, a number of NWT caribou herds, most notably the Bathurst herd, have declined to very low numbers, and there have been refinements of survey methods. The analyses in this report are reported as they were described in 2011, recognizing that status of several herds and survey methodologies have changed since that time. We suggest readers review recent calving ground and post-calving survey reports for an update on current methodologies employed and herd status.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This report provides statistical commentary on methods used to monitor barren-ground caribou herds with an emphasis on intervals between surveys. Given that the design and efficiency of each metric will affect the precision of each survey, I also provide some comments on the statistical design of surveys and provide potential ideas to improve each survey type. This report is not meant to be an exhaustive treatment of any method and I attempt to reference applicable reports for more details when needed. Some of the topics in this report were first discussed by (Heard 1985) and I attempt to refer to this report when applicable. A demographic monitoring manual ((Gunn and Russell 2008) produced by the Circum Arctic Rangifer Monitoring and Assessment Network (CARMA) also provides details on many of the methods discussed in this paper.

There are many factors that can potentially affect the optimal sampling design and sampling intervals. These are:

1. Herd status and how much risk is acceptable in terms of management. The intensity of sampling and associated sampling intervals will affect how quickly changes in population size can be detected. Therefore, reduced sampling effort can result in increased risk of a large decline being not detected and therefore any study design should be considered in the context of herd status and potential levels of decline that are acceptable to management (Figure 1). A study by Hauser et al. (2006) of survey intervals suggested that annual survey intervals are only needed if the results of the survey will directly affect management actions. They recommended the use of population models as a secondary means to evaluate status, and suggest that annual surveys are not needed for populations that are far from critical status thresholds.

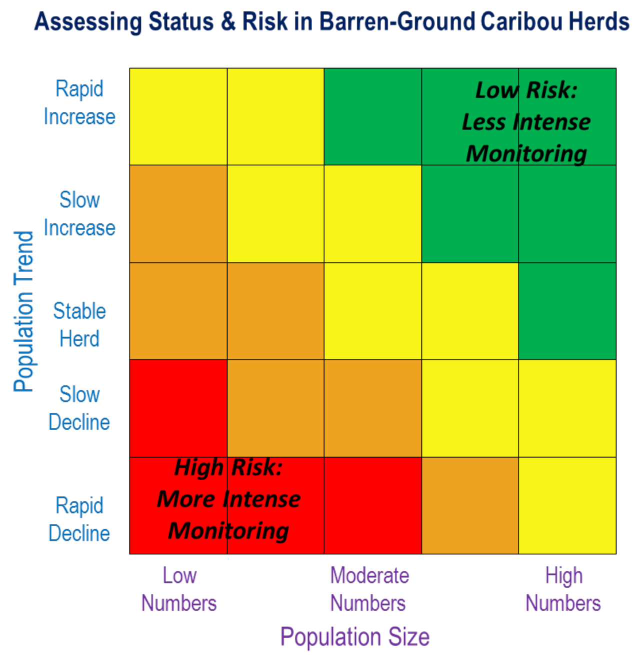

Figure 1. The relationship between monitoring intensity, population size, and population trend. Figure is courtesy of Jan Adamczewski (ENR).

2. The type of data collected and attributes of the data set. Data can be categorized into direct (i.e. trend or population estimates) and indirect indicators (calf-cow ratios and survival estimates) of population status. Each of these data types will have different levels of sampling variation and biological (process) variation. The degree of variation will determine how much sampling effort is needed to detect trends.
3. The cost of collecting the data. In general, estimates of population size are the most expensive measurements to collect compared to indices of productivity and trend.
4. The ultimate method of analysis. The power to detect change will also depend on how the data are analyzed. For example, analysis methods that consider multiple data sources simultaneously rather than single data sources should have higher power to detect trends (Boulanger et al. 2011). In addition, analyses that consider multiple years of data will be more powerful than methods that consider only sequential surveys.

Fundamental to determining survey intensity and associated sampling intervals is the amount of risk that managers are willing to accept in terms of detecting change in population size of caribou herds (Figure 2).

In terms of population demography we often quantify change in terms of annual change in population size. The actual ability of power to detect change in population size often takes years of time and with annual change being compounded yearly to produce a larger net change. For example, a population declining at 10% per year will be at 60% of its size in five years. In this context risk and associated sampling intensity to detect a decline would be based on current status of the population and the target level of decline that managers would like to detect.

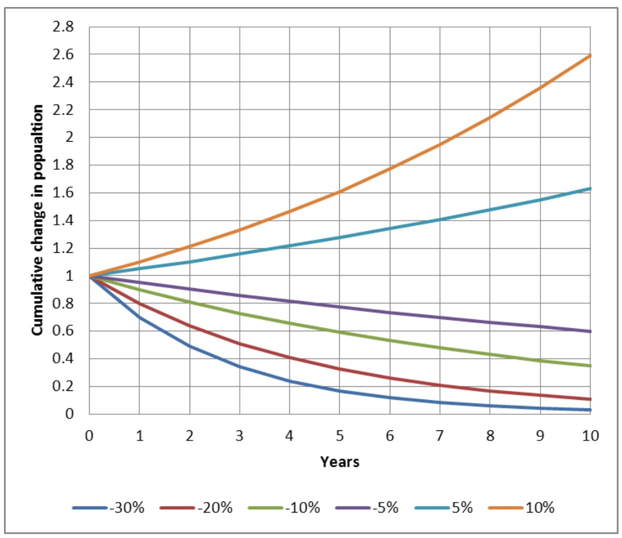

Figure 2. The relationship between annual change and cumulative change in population size as a function of the number of years surveys between surveys. Each line represents a different level of annual change.

RECONNAISSANCE SURVEYS

Summary of Current Methods

The general method for calving ground reconnaissance surveys has been the survey of core calving areas with transects at 10 km spacing (Nishi 2010). From this, either change in segment counts or replicate transect counts are used to estimate trend (Boulanger 2011). The power to detect change in population size depends on the amount of variation in counts during surveys and the amount of natural or process variation in population trend. Given this, one of the first relationships to consider for power analyses is how the precision of counts varies with observed counts or density of caribou on the calving ground. Plots of coefficient of variation of counts for the Bathurst and Beverly/Ahiak (Queen Maud Gulf) herds suggest that variation in counts is directly proportional to abundance (Figure 3). The mean coefficient of variation for the Bathurst was 44.6% (min=25.1, max=81.3%) whereas the mean coefficient of variation for the Beverly/Ahiak was 12.7% (min=8.1, max=19.9). The large difference in observed coefficients of variation between the Bathurst and Beverly/Ahiak herds is due to the fact that the Bathurst calving ground is composed of clustered groups of caribou in a relatively small area whereas the Beverly/Ahiak is composed of less clustered groups across a very large area. Given that transect spacing was similar for each of these herds (10 km), the Beverly/Ahiak had a much larger sample size of transects (mean=8.5 transects for Bathurst, mean=34.4 transects for Beverly/Ahiak) which also reduced the coefficient of variation of counts.

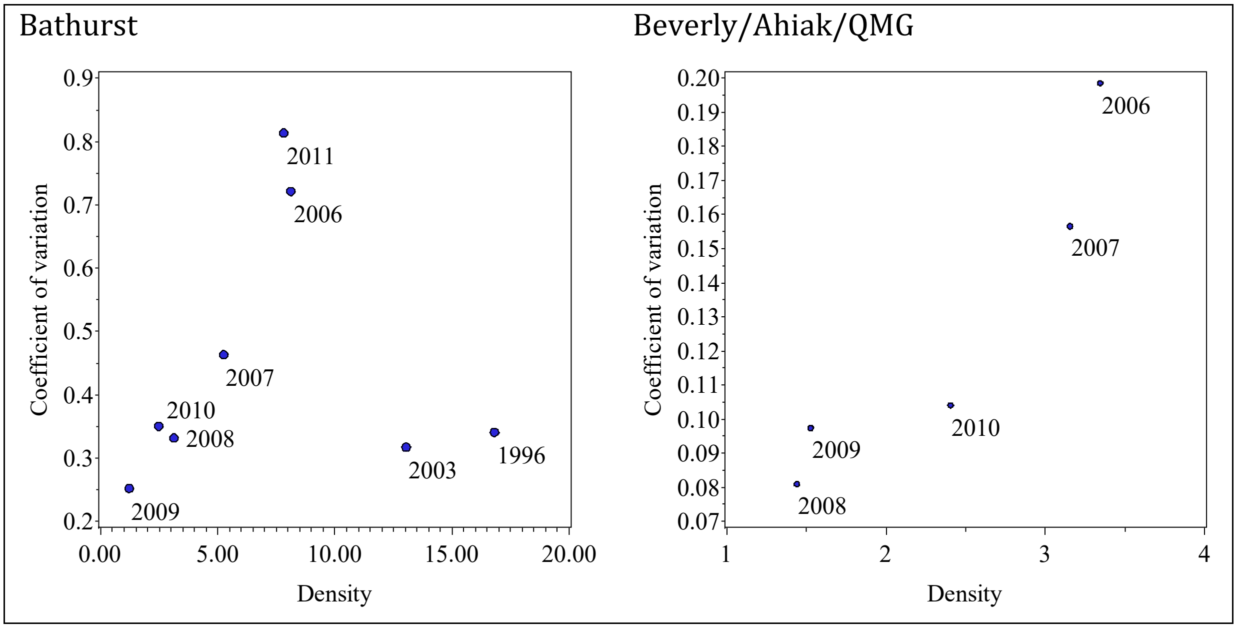

Figure 3. Relationship between CV of transect counts and average density from reconnaissance surveys of the Bathurst and Beverly/Ahiak (Queen Maud Gulf) caribou herds. Year of survey is given next to the data points. Note the different scales of the y and x axes of graphs.

The coefficients of variation from transect counts are only one component of variance in trend. The other component is process variance, or the natural biological variance in trend (Thompson et al. 1998). At this point, there are not enough successive years of data for herds to estimate process variance from reconnaissance surveys. However, I suspect it would add at least 5 to 10% to the observed sample variation CV's.

Power Analyses

The general result of coefficient of variation increasing with abundance runs counter to the general relationship with line transect sampling where precision increases with increasing abundance (Gerrodette 1987). Further inspection of the distribution of segment counts, and caribou biology, revealed that as density increases it is more likely that caribou will be clumped into higher density areas therefore creating a larger degree of variance than if the distribution of caribou was random (Figure 4).

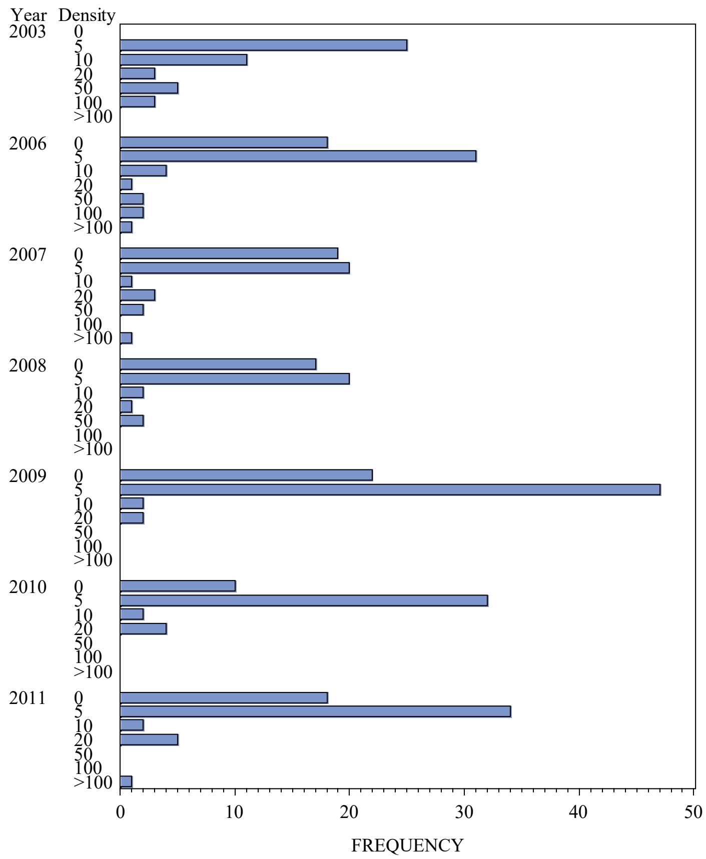

Figure 4. The dispersion of caribou density on calving grounds as indicated by histograms of segment densities from the Bathurst reconnaissance surveys. Note how in years of higher abundance (2003, 2006, and 2011) the distribution of caribou densities becomes roughly bimodal (due to clustering of caribou).

The preferred method to analyze calving reconnaissance survey data to assess trend is regression analysis. Comparison of annual counts is problematic with reconnaissance data given the low precision of any single year of data. I used negative binomial regression in which transects were the sample unit to analyze the Beverly/Ahiak and Bathurst data sets (Boulanger 2010b;2011). Both of these analyses suggested non-linear population trends which made application of power analyses difficult.

The main objective of power analyses was to assess the relative power of regression type methods to detect change with the observed sample variation and likely process variance, and if applicable determine how power is influenced by sampling intervals. For this objective, I decided to use a more general power analysis that allows direct variation of sampling parameters without consideration of specific project attributes such as population size. In this context the power analysis is meant to allow relative interpretation of various sampling strategies in terms of relative risk. Risk in this context is the number of years needed to detect a change in population size and the associated cumulative change in population size when it was detected. This approach is similar to that first used by Heard (1985). As discussed later, it is possible to run more exact analyses that consider historic data available for each herd as well as current population size and trends. However, this approach was beyond the scope of this current effort.

For the power analysis I used program TRENDS (Gerrodette 1993) which allows estimation of the number of surveys needed to detect given changes in population size as a function of survey precision. For program TRENDS, I assumed that the CV of counts was proportional to abundance (as suggested in Figure 1). I used an alpha level of 0.2, an exponential model of population change, and a two-tailed test to consider changes. A power level of 0.8 was considered adequate to detect a decline. I initially considered annual sampling intervals but also considered biannual sampling intervals. I simulated changes in population size from proportion changes of -0.3 to +0.3.

Power analysis results suggest that at the highest level of precision observed (CV=0.25) for the Bathurst it would take six annual surveys to detect an annual 10% change and four years to detect a 20% change. If CV's were higher the number of years to detect change increased. If the precision of the 2011 survey (CV=0.83) is used then it would take many years to detect any sort of trend in the population, at which point population size could be substantially reduced. For the Beverly/Ahiak (average CV approximately 15%) it would take at least five years to detect a 10% annual change in abundance. As mentioned earlier, process variance would add further variation into the observed sample coefficients of variation.

Table 1. Results from power analyses of reconnaissance surveys. The number of annual surveys needed to detect annual changes in population size are given. The corresponding change in population size given the number of years needed for detection is given.

CV counts	**number of annual surveys to detect given annual change**				**Proportion change in population given years of survey to detect change and annual change**			
	0.05	**0.1**	**0.2**	**0.3**	**0.05**	**0.1**	**0.2**	**0.3**
Years to detect decrease								
0.15	*7*	*5*	*3*	*3*	0.74	0.66	0.64	0.49
0.25	9	6	4	3	0.66	0.59	0.51	0.49
0.3	**10**	**7**	**4**	**4**	0.63	0.53	0.51	0.34
0.35	11	7	5	4	0.60	0.53	0.41	0.34
0.45	13	8	5	4	0.54	0.48	0.41	0.34
0.55	14	9	6	4	0.51	0.43	0.33	0.34
0.75	16	10	6	5	0.46	0.39	0.33	0.24
Years to detect increase								
0.15	8	5	4	4	1.41	1.46	1.73	2.20
0.25	**11**	**7**	**5**	**4**	1.63	1.77	2.07	2.20
0.3	12	8	6	5	1.71	1.95	2.49	2.86
0.35	14	9	7	5	1.89	2.14	2.99	2.86
0.45	16	11	8	6	2.08	2.59	3.58	3.71
0.55	18	12	9	7	2.29	2.85	4.30	4.83
0.75	22	15	10	8	2.79	3.80	5.16	6.27

Note that the results of these power analyses should be interpreted in the context of the amount of survey effort that has already occurred for the given calving grounds. The estimate of years to detect trend basically assumes that no prior data have been collected which often will not be the case. For example, since 2006, there has been seven years of calving reconnaissance surveys of the Bathurst herd (with a potential increasing trend in the last three years). Therefore, in the context of these results, we are not "starting at year 1" in terms of trend monitoring. However, the power analyses results mainly consider a single trend rather than multiple trends (i.e. a decline then a recovery) so therefore the actual power to detect changes will be somewhat dependent on the number of "recovery" years that are used to estimate trend. It would be possible to conduct specific power analyses that better consider the underlying trends, and prior data for each herd, however this work is beyond the scope of this analysis.

One conclusion from this power analysis is that reconnaissance survey methodology needs to be enhanced for the Bathurst herd if this method is to provide estimates of trend with adequate precision to detect population change. Until enhancements occur and the subsequent data evaluated, it is difficult to recommend changing the survey intervals to bi-annual or longer time increments. More succinctly, the amount of sampling variance introduced by yearly sampling efforts needs to be reduced by better study design.

For herds with reasonable survey precision, can we sample bi-annually?

I ran analyses with annual and bi-annual sampling for CV levels of 0.15 and 0.25 which is the approximate range of CV levels for the Beverly/Ahiak herd (given that process variance will increase the actual amount of observed sample variation). It can be seen that bi-annual sampling adds four more years for adequate power to be achieved to detect declines in population size. This amounts to an added 10-25% cumulative change in population size when adequate power to detect the decline is achieved.

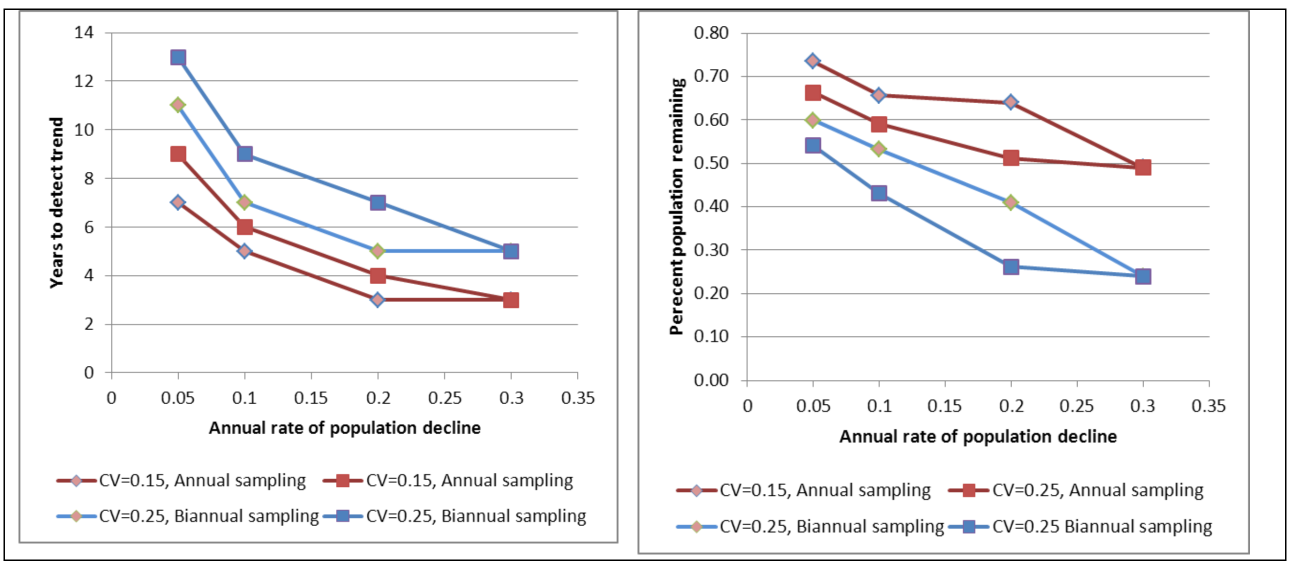

Figure 5. Comparison of annual and bi-annual sampling of reconnaissance surveys as a function of CV of counts on calving grounds.

How can we improve the current calving ground reconnaissance method to improve power and reduce the need for annual survey intervals?

The two main improvements to calving ground reconnaissance surveys are (1) increased coverage in areas of high density and (2) better classification of caribou groups to allow better definition of core calving areas that mainly contain breeding caribou. Increasing coverage will reduce annual sampling variation. Better classification of caribou will decrease sampling variation and also decrease across year variation in counts by ensuring that the same "target population" of primarily breeding caribou is surveyed each year.

Increasing sample coverage for the Bathurst herd

The logical question for the Bathurst herd becomes whether the degree of count precision could be increased by increasing coverage in the higher density areas encountered on the calving grounds. To explore this, I randomly resampled the 2006 and 2009 photo survey calving ground data at the recon (approximately 8%) coverage and a 5 km spacing coverage of 16% (Figure 5, Table 2).

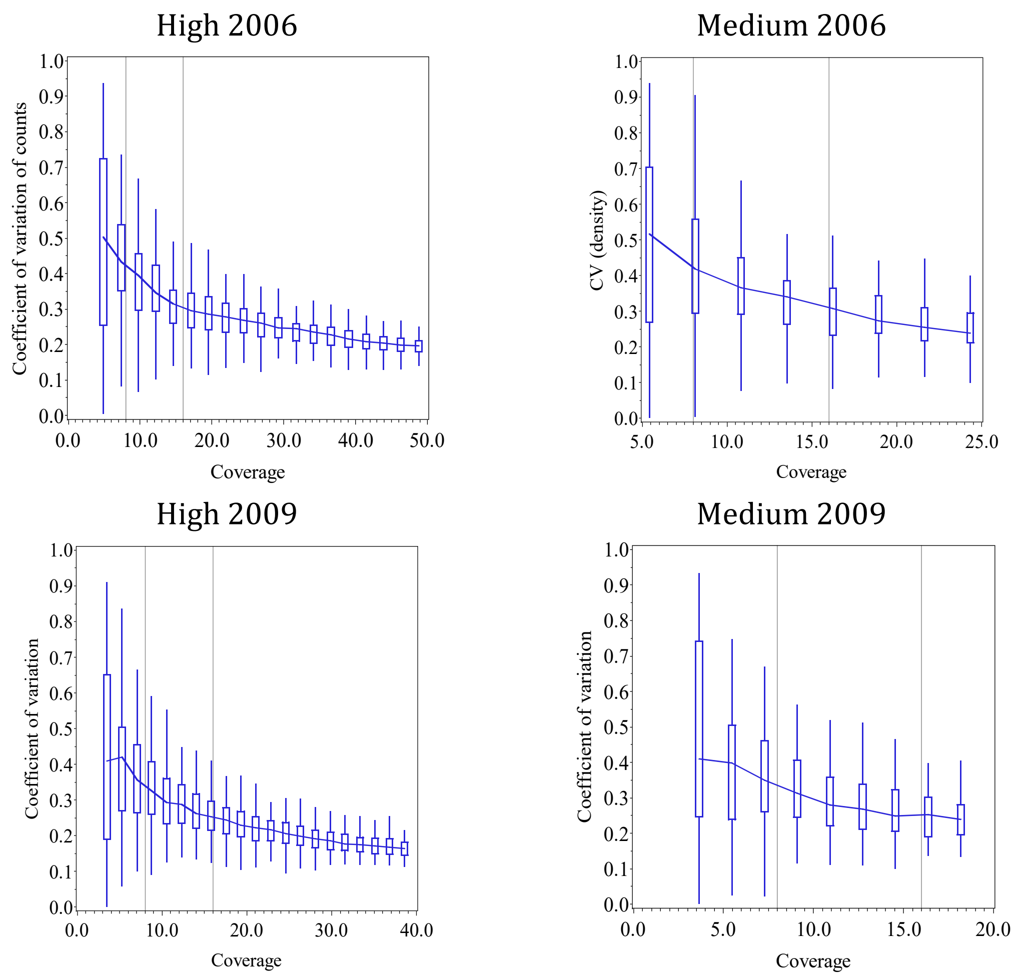

Figure 6. Estimated relationship between transect coverage and coefficient of variation of transect counts for the 2006 and 2009 photo-survey data for the Bathurst herd. Boxplots show the spread of points resulting from the resampling exercise.

The results of this analysis suggest that across the densities observed in these surveys, the CV of counts could be reduced to 25-30% even at the higher densities observed during the 2006 survey by reducing line spacing to 5 km (Table 2).

Table 2. Results of resampling analysis of 2006 and 2009 photo survey data from Bathurst calving ground surveys.

Year	Strata	Mean density	Surveyed coverage	CV at full coverage	Resampling estimates of CV at given levels of coverage	
					8%	16%
2009	high	6.76	40.6%	13.7%	0.34	0.25
2006	high	49.3	53.3%	16.0%	0.44	0.32
2006	medium	2.57	23.8%	24.0%	0.43	0.31
2009	medium	2.49	19.1%	26.4%	0.34	0.25

Better classification of caribou groups

The second method to increase precision is through better classification of caribou during the reconnaissance survey. This topic has previously been discussed by Gunn and Russell (2008), and Poole et al. (2010) and it is highly recommended that these reports be reviewed when considering strategies for classification of caribou groups. Better classification would allow better definition of the core calving area, and perhaps even the eventual analysis of breeding caribou (females that were or are pregnant) without the added variation of non-breeding caribou. Non-breeding caribou, which are often yearling caribou and bulls, are likely to introduce variation for a variety of reasons. First, the proportion of yearlings in the population is likely to vary due to yearly variation in productivity as shown by spring calf-cow ratios from the Bathurst herd (Figure 6). Therefore, the estimated trend becomes that of the core adult breeding female population and the varying non-breeder yearling caribou. I suspect that surges of productivity along with sampling variation are one reason why year to year comparisons of counts on calving ground surveys often suggest biologically infeasible increases. A rough approximation of the degree of variation in the context of calf-cow ratios is that the number of yearlings (assuming most survive from the spring survey until the calving survey) varies between 0.1 up to 0.5 yearlings per adult female caribou. If non-breeding groups can be better classified, or the core calving ground can be defined to exclude or minimize these groups, then the actual trend from counts will be a better indicator of trend in the core adult females in the population.

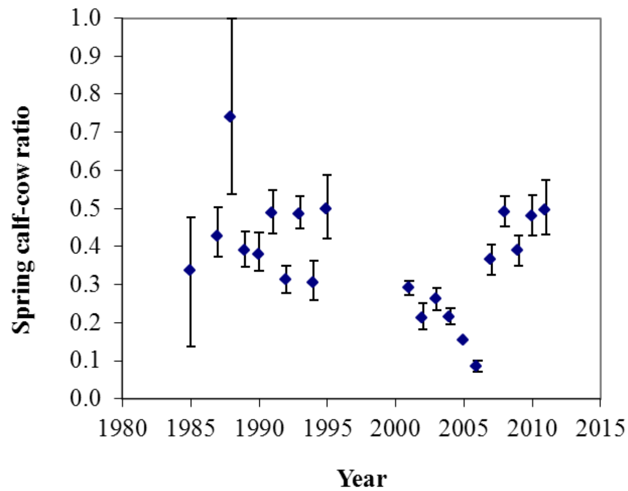

Figure 7. Calf cow ratios from the Bathurst herd from 1985 to2011. Calf cow ratios can be interpreted as the number of calves per adult female in the herd. In this context, calf-cow ratios are proportional to the amount of potential variability introduced into calving ground estimates due to the partial presence of non-breeding yearlings (which are calves for the spring calf cow ratios) on the calving ground.

The issue of sightability

I note that the issue of sightability, and in particular yearly variation in sightability, cannot be completely ignored for calving ground reconnaissance surveys, especially if there is a lot of year to year variation in ground conditions or other factors that influence sightability. As Heard and Williams (1990) stated: "We have spent ten years and over half a million dollars demonstrating the inaccuracies of visual transect strip surveys. However, the technique is still being used and results are interpreted as if they are accurate and consistent". For example, recent research has shown that the use of double rather than single observers influences the proportion of caribou sighted by up to 30% (Boulanger et al. 2010). In addition, group sizes, ground conditions, plane type, and observer experience will influence sightability. As it stands now, it is assumed that sightability is constant each year so that counts of caribou indicate trend. I suggest that there are now double observer methods available to estimate sightability and potentially counting bias (that were not available during the review of Heard and Williams 1990) which will result in a more robust trend estimate.

Other analytical strategies to optimize inference

I note that there are methods to better separate the biological variation in the population (process variance) compared to the current method that pools sampling and process variance (Dennis et al. 2006, Humbert et al. 2009). This type of analysis would allow an estimate of actual biological variation with sampling variation removed, therefore providing a better estimate of trend. These methods work best with annual surveys and do require longer time series of data sets to provide valid process variance estimates. In addition, the directional non-linear trends observed in many of the caribou herds (i.e. decline and recovery) also need to be considered when estimating process variance, which complicate analyses. Further development of these methods is beyond the scope of this current work but should be pursued as more annual data become available.

There are also other analytical strategies such as assessment of dispersion of caribou on the calving ground (Poole et al. 2010) and use of plots of segment density classes (Nishi 2010) that can be used to assess trends in caribou distribution and relative herd status.

Table 3 provides a summary of modifications to the calving ground reconnaissance survey that should be considered to allow better estimates of trend. If these changes are successful then more precise yearly counts will be produced which may allow a less frequent survey interval. I note that the recommendations about increased line spacing mainly pertain to the Bathurst herd rather than the Beverly/Ahiak herd.

Table 3. Summary of recommendations for enhancement of reconnaissance surveys to increase the overall power to detect changes in population size.

Issue	Statistical implications of issue	Recommendations
Clustering of caribou groups results in only one or two transects or transect segments sampling the high density areas of the calving ground *for the Bathurst caribou herd*	Substantially reduced trend estimate precision and less power to detect trends (as exemplified with the 2011 data). This results in higher overall cost of surveys given that more annual surveys are needed to detect trends when data is imprecise	• Adaptive sampling design (for the Bathurst herd) where transect spacing is changed to 5 km or less when high densities (>10caribou/km^2) are encountered.
Problems with classifying caribou results in segments with unknown caribou composition and minimal inference about the proportion of non-breeders on the calving ground each year *for the Bathurst, Beverly/Ahiak, and other caribou herds.*	Difficulties in identifying the core calving ground where the majority of breeding caribou are found. In addition, grouping breeders and non-breeders for trend estimates creates higher variance of trend estimates given that variable productivity can cause the number of non-breeders to vary greatly from year to year (compared to breeders).	• Ensure that both sides of the plane have at least one experienced observer with prior experience classifying caribou. • Re-fly transects or portions of transects at lower speeds and/or lower altitudes to reclassify groups when needed.
Groups of caribou are added together and summarized by segment or by further spaced waypoint data points.	Inability to interpret if counts in database represent larger groups or the summation of smaller groups which makes it difficult to determine the true degree of clustering of caribou as well as assessment of how groups were classified (previous point).	• Use tablet computers, or other methods that allows group-specific data entry so that the database better reflects the actual observations from the aerial survey.
Ground conditions, group sizes of caribou, and observer experience influence the ability of observers to observe caribou, and count caribou group sizes	If ground conditions or other factors influence sightability then added variance to trend estimates is introduced. If directional trends (i.e. sightability better one year than the other), then biased estimates can result.	• Use double observer methods to allow assessment of sightability and replicate counts between observers. In addition, ground conditions and weather conditions should be documented.

Given the limited power of the calving ground reconnaissance survey method to assess trend I suggest that annual surveys are conducted especially for caribou herds at risk. I suggest that further discussion of improvement of the field component of this methodology is essential to ensure statistically rigorous trend estimates. Finally, methods that integrate multiple data sources such as the OLS model should be considered for herds at risk given that they utilize data from survival analysis, reconnaissance surveys, photo surveys, and composition surveys into one analysis, therefore improving the power to detect trends in caribou demography (compared to single data source analyses).

BREEDING FEMALE ESTIMATES FROM CALVING GROUND SURVEYS

The estimates of breeding females from calving ground photo-surveys and extrapolated herd size are often used for management purposes. For example, having a firm herd estimate is often used to set harvest levels and other management actions through co-management processes. Estimates can also be used to give a better indication of trends and population change between surveys.

The statistical criterion for determining optimal survey intervals is the situation when it is possible to detect a difference between two estimates. There are two methods to detect trends from photographic surveys of calving grounds. First there is comparison of sequential estimates using a t-test to determine if there are significant differences. Second, there is weighted regression analysis of multiple years of data to assess longer term trends. Of these, the most powerful is regression analysis since it uses multiple years of data. Using a segmented or polynomial regression approach, which is similar to the reconnaissance survey analysis, it is possible to estimate if data from the current survey indicate a change in trend from previous surveys (Boulanger 2010a).

One of the key distinctions between the stratified photographic surveys and the reconnaissance level surveys *is that survey effort is adjusted to optimize estimate precision during the photo survey rather than using a fixed amount of survey effort (with reconnaissance surveys). Therefore the estimates from these data sets are likely to be more precise and vary less with abundance compared to the reconnaissance level surveys.* For the Bathurst herd, the coefficient of variation for breeding female estimates ranged between 6% and 23% for surveys conducted between 1986 and 2009 with an average CV of 15% (Table 4) (Gunn et al. 2005b, Nishi et al. 2010). After 2003 both sequential t-tests and regression analyses were applied to estimates of breeding females. In general, the regression analysis was more powerful at detecting trends compared to sequential t-tests. This was due to the simple fact that regression analyses utilize multiple years of data whereas sequential t-tests only utilize two adjacent surveys' worth of data.

Table 4. Summary of breeding female estimate precision and results of statistical tests. The regression analyses results were based on estimate of trend from 1986 to the given year compared to the t-tests that tested differences between sequential estimates.

Year	$\widehat{N}$	SE	CV	Population change		Decline detected?	
				Between surveys	Annual	Sequential t-tests	Regression analysis
1986	203,800	12,696	0.06				
1990	151,927	25,805	0.17	-25.5%	-7.1%	yes	
1996	151,393	35,144	0.23	-0.4%	-0.1%	no	
2003	80,658	13,149	0.16	-46.7%	-8.6%	no	yes
2006	55,593	8,813	0.16	-31.1%	-11.7%	no	yes
2009	16,604	2,176	0.13	-70.1%	-33.2%	yes	yes

However, for some herds (i.e Beverly/Ahiak) only one estimate of abundance is available and therefore the use of sequential t-tests is still needed. Also, it is inevitable that sequential estimates will be compared rather than overall trend estimates from regression analyses (Gunn et al. 2005b). For this reason, power of sequential t-tests to detect differences between estimates is still useful for management purposes despite the fact that it is less statistically efficient than regression analyses.

To explore the power of sequential t-tests I wrote a simulation program that simulated sequential estimates with varying levels of estimate precision, yearly change in population size, and survey intervals. For this analysis, the amount of survey effort (i.e. number of transects employed) was assumed to be similar to the 2009 Bathurst survey.

Results suggested that annual surveys were not optimal, and that only large annual changes (+/- 30%) could be detected if surveys were conducted on a bi-annual basis when the CV of surveys was 15% (Figure 8). Annual changes of 20% could be detected if survey interval was three years (with CV=15%). Table 5 summarizes the power analysis results as well as the cumulative change in the population size at the given survey interval. These results suggest that the populations would be reduced to 60-70% of original size at target CV levels of 0.15.

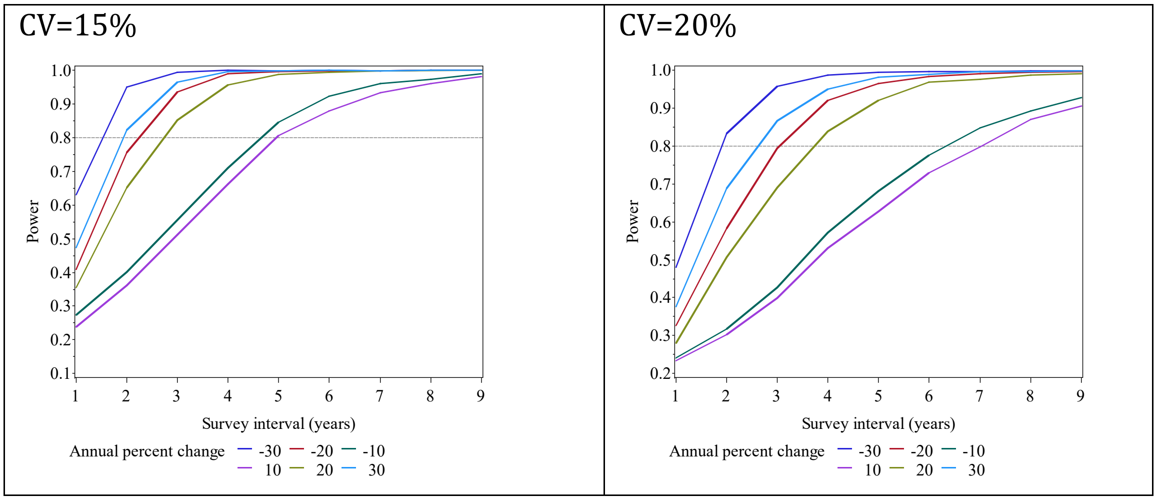

Figure 8. Power to detect declines at various survey intervals (years) when a t-test is used to compare sequential counts.

The results of power analyses that use both regression and t-tests suggest that the sampling interval for calving ground surveys should never be less than three years unless a very large change in abundance is expected. For the most likely rates of change in population size (+/- 10% per year) then a survey interval of five to six years is adequate. These general results are similar to those of Heard and Williams (1990) who also recommended survey intervals of six years. As discussed later, multiple data source modeling can be used in interim periods to better determine population trends and likely demography in the absence of annual population estimates (Hauser et al. 2006).

Table 5. Survey interval needed to detect annual population change at various levels of survey precision (CV) for sequential t-tests.

CV N estimate	Annual change	Year when power>0.8	Power at survey interval	Proportion population remaining at given year
0.15	-30	2	0.95	70.0%
0.15	-20	3	0.94	64.0%
0.15	-10	5	0.83	65.6%
0.15	10	6	0.89	161.1%
0.15	20	3	0.87	144.0%
0.15	30	2	0.84	130.0%
0.2	-30	2	0.83	70.0%
0.2	-20	4	0.91	51.2%
0.2	-10	7	0.84	53.1%
0.2	10	7	0.80	177.2%
0.2	20	4	0.83	172.8%
0.2	30	3	0.88	169.0%
0.25	-30	3	0.88	49.0%
0.25	-20	5	0.88	41.0%
0.25	-10	8	0.80	47.8%
0.25	10	10	0.84	235.8%
0.25	20	5	0.81	207.4%
0.25	30	4	0.86	219.7%

POST-CALVING ESTIMATES

Post calving methods provide an estimate of the entire herd (caribou at least one year old) rather than an estimate of breeding females. As with calving ground photo surveys, one criterion for post-calving surveys is the need for a reliable entire herd estimate with acceptable variance for management purposes.

There are some important differences between post-calving and calving ground estimates that should be considered when evaluating trends from each data source. First, post-calving estimates include the entire herd of caribou, including yearlings, and exclude only calves of the year. In contrast, breeding female estimates and extrapolated herd estimates do not include yearlings given that they are estimated by the number of breeding females divided by pregnancy rate and sex ratio. Pregnancy rate, which is usually set at 0.7 to 0.72 is based upon caribou that have some probability of being pregnant (in this case caribou that were yearlings or older in the previous fall) (Dauphine' 1976). Given this, yearlings (that were calves of the year the previous fall) are not included. Therefore caution should be exercised when comparing trend estimates from post calving and extrapolated calving ground herd estimates.

The same general power analysis of sequential t-tests used for calving ground estimates also applies to post-calving estimates. However, caution should be applied to any trend estimates from post-calving surveys given the likelihood of underestimated variances due to non-uniformity of collared caribou relative to group size if the Lincoln Petersen method is used for population estimate (Bechet et al. 2004). Recent research has applied the Rivest estimator to the 2010 Bluenose East post calving data set as well as other data sets(Adamczewski et al. 2017, Boulanger et al. 2018). The Rivest estimator includes more realistic variance estimation.

COMPOSITION SURVEYS

I note that obtaining an unbiased adult sex ratio is challenging and often requires adequate sample sizes of collars to delineate areas where various groups of caribou are distributed (Otto et al. 2003), as well as reconnaissance surveys to ensure that all bull and cow groups are adequately sampled (Gunn et al. 2005a, Gunn and Russell 2008).

Spring calf-cow ratios

Calf-cow ratios that occur in the spring provide valuable information about the yearly productivity of herds and calf survival (Gunn et al. 2005a). Overall, the precision of spring calf-cow ratios is high for the Bathurst herd (mean CV=0.078, min=0.03, max=0.16). However, sole interpretation of calf-cow ratios as an indicator of herd status is problematic due to the fact that they can be misleading when there are changing trends in adult cow survival (Harris et al. 2007, Boulanger et al. 2011). A good example of problems with interpreting calf-cow ratios occurred with the Bathurst herd in 2007-2009 where the ratios indicated reasonable productivity when in fact productivity was low, and ratios were biased high due to low adult cow survival (Boulanger et al. 2011). In addition, it is known that ungulate populations can tolerate large short-term fluctuations in productivity without direct impact on overall population trend (Gaillard et al. 1998). Therefore, year-specific trends in calf-cow ratios do not provide an overall indication of population status. However, they can provide a general estimate of yearly herd productivity which is useful information especially for recovering or threatened populations.

The first determinant of the overall utility of collecting composition surveys depends partly on herd status and the way that the data will be analyzed. For a declining or recovering herd (such as the Bathurst) composition surveys can provide a valuable indicator of population recovery especially if it is analyzed jointly with other data sources (Boulanger et al. 2011). If interpreted in unison with trend and survival data, calf-cow ratios can help determine if productivity is offsetting observed mortality rates. For example, OLS models basically can help answer the question if observed trends in population size can be explained just by trends in productivity or also by trends in adult survival.

The optimal survey interval for composition surveys is either annual or tri-annual depending on the status of the herd. Annual collection is optimal to capture the annual variation in calf cow ratios (Figure 4) (Heard and Williams 1990). Often, caribou can exhibit a "saw-blade" pattern in productivity where high productivity one year is offset by low productivity the following year (Gunn et al. 2005a). Thus bi-annual sampling may be problematic since it could capture either the low or the high years. Given this, for herds that are not threatened, a tri-annual survey can capture variation and still be potentially useful for multi-data source models.

Fall composition surveys

Fall composition surveys provide estimates of adult sex ratio and calf-cow ratios to index productivity. Of these metrics, the adult sex ratio is most useful especially for deriving population estimates from breeding females. Calf-cow ratios in fall are less indicative of productivity given that they do not account for over-winter survival of calves. Therefore they overestimate yearly productivity, and are less useful than spring calf cow ratios (Boulanger et al. 2011). The adult sex ratio is useful to track the male segment of the caribou population, and can provide indirect inference of male population size and adult male survival rates (Boulanger et al. 2011) when used in multi-data source demographic models.

The survey interval for sampling of adult sex ratios depends partially on whether there is a male-dominated harvest (Boulanger and Adamczewski 2015). In this case, more frequent survey intervals are required. In other cases, the adult sex ratio is most useful if conducted at the same frequency as calving ground photo surveys.

METHODS TO INCREASE INFERENCE AND POWER FROM DEMOGRAPHIC INDICATORS

One take-home message of power analyses is that it is difficult to quickly detect changes in population size with the current levels of precision possible with caribou surveys and financial constraints on the survey frequency. Given this, I suggest that an adaptive management approach that utilizes population models, observed harvest levels, and all data sources available be used to adaptively define survey intervals as a function of likely herd status (Boulanger et al. 2011). This approach also involves inclusion of other indictors of herd status such as harvest levels, environmental covariates, and indices of predator abundance. This approach follows the work of Hauser et al. (2006) who used a population model along with environmental covariates to assess likely trend in population size, and with these data, optimized population survey intervals.

Use of population models that combine data sources.

Use of population models can also help determine the actual cause of population change. For example, demographic analysis of the Bathurst herd using a multi-data source population model (Boulanger et al. 2011) determined that negative trends in survival rates in association with low productivity were associated with observed declines in population size. The negative trends in adult survival were not detectable using the collar data alone, and it was difficult to determine if observed trends in calf-cow ratios could have caused the observed declines. In this case, the demographic model allowed detection of trends and rigorous exploration of causes of the decline in the population.

I note that the OLS model used for the demographic analysis is in essence a series of regressions on each data source where the predictions for the regression are generated by a population model. In fact, this method has also been called "Seemingly unrelated regressions" (White and Lubow 2002) given that it is really a set of regression analyses based on predictions of a population model. This objective and use of model is distinctly different than PVA (Population Viability Analysis) or simulation models used for population demography research.

Recently the OLS model has been updated to a Bayesian integrated population model (IPM: Kery and Schaub 2012) as illustrated in recent calving ground survey reports (Adamczewski et al. 2019, Boulanger et al. 2019). The IPM approach is similar to the OLS model; however, it provides more robust estimates of demographic parameters.

The main constraint with the OLS and IPM approach is the need for adequate collection of various data sources for the model. One key component is the estimation of adult survival rates from collar data. Most monitoring programs utilize radio collared animals for use in delineation of seasonal ranges. However, the tracking of collared animal fates is essential for survival analysis, and often fate of collared caribou is not adequately determined so that it is not possible to determine fate, which makes survival analysis problematic. I suggest that further development of methods to track collar fates is essential to allow full utilization of the data from collared caribou. If this is done then survival data can be combined with survey data to maximize inference about herd status using all data sources.

The use of simulation models for adaptive management to further refine study designs

The OLS method can also be extended to better appreciate future management scenarios. The OLS model was modified to generate stochastic variation and used to explore the effect of varying harvest levels on the Bathurst herd (Boulanger and Adamczewski 2010). This simulation model simulated population trajectories as a function of various levels of productivity and harvest level. The model generated predictions in terms of herd status, but these predictions were illustrated in terms of the power of statistical tests to discern the given model outcomes. For example, in Figure 8, simulations were run with 0 harvest but with varying levels of productivity (based upon calf-cow ratios). The bars represented different levels of herd status resulting from each simulation scenario. The red and green bars represented outcomes in which a change would be detected from sequential t-tests. The other shades represented changes that could not be detected. These results suggested that if productivity was low (0.18) then it would still take at least nine years to detect a change in population size using sequential t-tests, however, the majority of simulations still suggested a declining population (the brown bars). This information could then be used to prioritize further monitoring of productivity, and would help assess optimal intervals for subsequent surveys. If productivity was high (0.57) then the risk of decline was lessened suggesting that monitoring actions could be relaxed. This model also generated predicted calf-cow ratios for fall and spring sampling, and predicted sex ratios for fall composition surveys.

One important component of the use of models for adaptive management would be the incremental refinement of model predictions as new data became available. For example, productivity in Figure 9 was simulated across a large range of values and variation. These model runs could be further refined based upon recent composition survey results and other recent data sources. This would result in a more refined and focused set of model predictions.

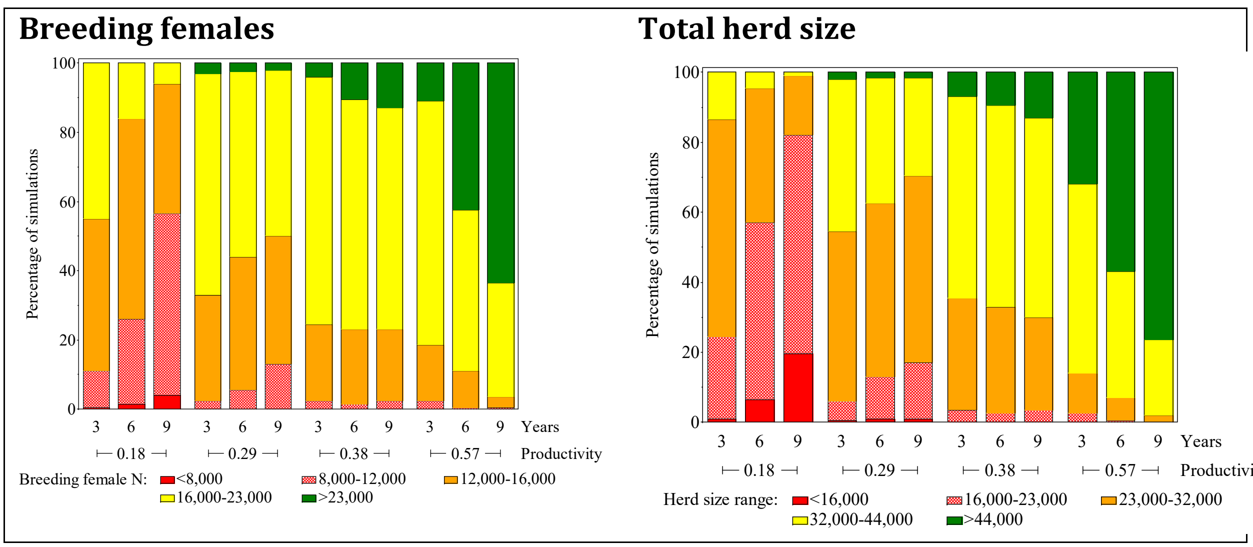

Figure 9. Results of simulations with no harvest (male or female) as a function of mean productivity and years since 2009. Each colour on the bar denotes the relative proportion of simulations that resulted in a given range of herd sizes/management targets with the estimates of 16,000 cows and 32,000 caribou as a baseline. Declines that are coloured red and increases that are coloured green are statistically detectable. For these simulations adult female survival was 0.88 since no harvest was simulated. Productivity estimates of 0.18, 0.29, 0.38 and 0.57 correspond to mean annual rates of change of 0.96, 0.98, 1.00 and 1.04 respectively.

SUMMARY OF RESULTS

Table 6 provides a summary of results of the analyses conducted in this report. I note that some analyses, such as reconnaissance trend analyses for the Beverly/Ahiak herd (with the 2011 data included), and the Bluenose-East herd are yet to be conducted. Therefore, more specific recommendations for these herds would depend on more detailed demographic analyses.

Table 6. Summary of optimal survey intervals and analysis strategies for the Bathurst herd and other caribou herds (when noted).

Data type	Analysis method	Recommendations
Reconnaissance analysis of calving grounds	Regression analysis of mean counts in transect surveys OLS model to integrate findings with other data sources	• Sample coverage is constant (8%) so precision is reduced as caribou densities increase due to aggregation of caribou in high density clusters. • Bathurst (herd at risk): Annual surveys are needed to compensate for the high amount of sample variation with the present sampling design. • Enhancements (Table 3) to field survey methods are required for this method to provide better trend estimates, especially for the Bathurst herd. If enhancements are applied, it may be possible to reduce the survey interval to bi-annual sampling. • Some of the enhancements (tighter transect spacing in high density areas) may not apply to other herds (i.e. Beverly/Ahiak) that occupy a larger calving ground area, have a larger number of transects sampled, and have less high density clustering of caribou.
Calving ground surveys to estimate breeding females	Sequential t-test of N estimates Regression analyses OLS model	• Survey coverage is *adjusted* during each survey so that survey precision (CV) *is approximately constant* as densities increase. • Sampling intervals of three to six years are optimal to detect any changes in population size (if CV of estimates is 0.15). In some cases, surveys at greater than tri-annual intervals are optimal for lower risk herds. • Regression analyses are more powerful than sequential t-tests to detect trends in estimates for herds that have multiple years of data.
Post-calving surveys to estimate breeding females	Sequential t-test of N estimates Regression analyses	• Same general recommendations apply as calving surveys. • Post-calving Lincoln Petersen estimates are not as statistically robust as calving ground estimates and therefore trend estimation is less certain. Rivest estimates are more robust. • Post-calving estimates include yearlings and bulls and are therefore more variable than calving ground estimates which results in reduced power to detect trends in the core breeding female population.
Calf-cow ratios conducted in spring	Regression analyses of trends OLS model with all data sources considered	• As a stand-alone indicator of population status, calf cow ratios are of limited utility. • In unison with population estimates they can provide valuable inference for populations. • For recovering or threatened populations annual sampling is optimal. • For less threatened populations sampling every three years is adequate • Calf-cow ratios are best interpreted using models that simultaneously consider other data sources.
Fall composition surveys	Used to estimate extrapolated population size Sex ratio used in OLS models to estimate male population size	• Adult sex ratios are most useful for extrapolated population estimates from calving ground surveys • Calf-cow ratios sampled in the fall have less utility than spring calf cow ratios. • Adult sex ratios should be estimated more often if male-dominated harvest ocurs. • Otherwise, frequency should coincide with calving ground surveys

How to apply the results of this analysis.

The following steps should be used to apply the results of this analysis;

1. Define how much risk is acceptable with a herd given its current status and determine thresholds in terms of trend and population size where management actions will occur. Figure 2 demonstrates the cumulative change in population size as a function of annual change and the number of years in-between surveys. If an estimated population size is available then it is relatively easy to translate this figure into actual population sizes. Once this is done, a threshold population level should be considered where management action would be required. The various annual decline rates and associated number of years associated with the target management level should be noted and used when interpreting power analyses.
2. Assess the level of precision with the various survey metrics or use the levels associated with the Beverly/Ahiak and Bathurst herds (Figure 2) as probable levels of precision for reconnaissance, calving ground, or post-calving estimates.
3. Determine the relative power and number of years to detect change with reconnaissance or population estimate surveys. For reconnaissance surveys, use Table 1 and also assess if bi-annual sampling is acceptable (Figure 4). For calving ground or post-calving surveys, survey intervals can be assessed using Table 5 and Figure 7. If data have already been collected, then the estimates of power (i.e. years until trend detected) will be conservative since the power analyses assume that no prior data has been collected when assessing years to detect trend. Power analyses that consider past data can be run for each herd which would give a better indication of optimal sampling strategies. This type of analysis was beyond the general scope of this report.
4. Dependent on herd risk, and likely analyses to be used, assess the utility of calf-cow ratios. The utility of calf-cow ratios will depend on herd status, and how the calf-cow ratio data will be analyzed. Annual surveys to every three years are the most recommended survey intervals.
5. Formulate the best analysis strategy that considers all of the data types. Various metrics will be available to assess population status and better assessment of overall herd status. Consider the use of population models to allow a better forecast of likely population status and optimal survey

intervals and intensities. It would make sense to repeat this exercise as new data become available.

RECOMMENDATIONS FOR FUTURE ANALYSES

The analyses conducted in this paper demonstrate the complexities of determining optimal management and monitoring strategies for caribou herds. The main objective of this work was to provide a general framework for determination of sampling design. There are analyses that were beyond the scope of this work that could be considered to further. These analyses are outlined in point form:

- It is likely that optimal sampling designs, data collections, and sampling intervals will vary for each herd dependent on population status, management requirements, and limitation on data collection. I suggest that this work could be best delivered in the format of a workshop where each herd's status, management concerns, and data attributes are considered. This would ensure that sampling designs are best tailored for the management and research needs for each herd.
- In some cases, more exact power analyses that consider prior data collected for each herd would be useful to better determine optimal sampling designs. A Monte Carlo simulation approach could be used to generate more refined sampling design recommendations.
- A subsampling analysis of data sets from other calving ground surveys than the Bathurst (Figure 5) would give a better indication of how coefficient of counts varies with coverage for herds that show less-clustered distributions on the calving ground.
- A more exact set of recommendations for calf-cow ratio sampling could be formulated through the analysis of longer-term complete calf-cow data sets from herds such as the Western Arctic Herd. A sub-sampling/simulation analysis could be used to determine optimal sampling intervals to detect various trends in productivity.
- The use of collared caribou to determine herd location and distribution is an important component of the sampling design for composition surveys as well as calving ground surveys. Analyses to assess optimal sample sizes of collars for delineation of seasonal ranges are currently in progress.

LITERATURE CITED

Adamczewski, J., J. Boulanger, B. Croft, T. Davison, H. Sayine-Crawford and B. Tracz. 2017. A comparison of calving and post-calving photo-surveys for the Bluenose-East herd of barren-ground caribou in northern Canada in 2010. Canadian Wildlife Biology and Management 6:4-30.

Adamczewski, J., J. Boulanger, H. Sayine-Crawford, J. Nishi, D. Cluff, J. Williams and L.M. LeClerc. 2019. Estimates of breeding females and adult herd size and analyses of demographics for the Bathurst herd of barren-ground caribou: 2018 calving ground photographic survey. Environment and Natural Resources, Government of the Northwest Territories. Manuscript Report No. 279.

Bechet, A., A. Reed, N. Plante, J.F. Giroux and G. Gauthier. 2004. Estimating the size of the greater snow goose population. Journal of Wildlife Management 68: 639-649.

Boulanger, J. 2010. Estimates of trend for caribou utilizing the Beverly/Ahiak calving ground 2006-2010. Environment and Natural Resources, Government of Northwest Territories. Draft report.

Boulanger, J. 2011. Estimates of trend for caribou on the Bathurst calving ground 2003-2011 Environment and Natural Resources, Government of Northwest Territories. Draft report.

Boulanger, J. 2014. Appendix K in Nishi, J.S., B. Croft, J. Boulanger and J. Adamczewski. 2014. An estimate of breeding females in the Bathurst herd of barren-ground caribou, June 2009. Environment and Natural Resources, Government of the Northwest Territories. File report 144.

Boulanger, J. and J. Adamczewski. 2015. Simulations of harvest and recovery for the Bathurst caribou herd, with annual variation. Environment and Natural Resources, Government of the Northwest Territories. File Report 145.

Boulanger, J., K.G. Poole, J. Williams, J. Nishi and B. Croft. 2010. Estimation of sighting probabilities from caribou calving ground surveys using double observer methods. Governments of Northwest Territories and Nunavut. Draft report.

Boulanger, J., A. Gunn, J. Adamczewski and B. Croft. 2011. A data-driven demographic model to explore the decline of the Bathurst caribou herd. Journal of Wildlife Management 75: 883-896.

Boulanger, J., J. Adamczewski, and T. Davison. 2018. Estimates of caribou herd size using post-calving surveys in the Northwest Territories and Nunavut, Canada: A meta-analysis. Rangifer 38:39-78.

Boulanger, J., J. Adamczewski, J. Nishi, D. Cluff, J. Williams, H. Sayine-Crawford, and L. LeClerc. 2019. Estimates of breeding females and adult herd size and analyses of demographics for the Bluenose-East herd of barren-ground caribou: 2018 calving ground photographic survey. Environment and Natural Resources, Government of the Northwest Territories. Manuscript Report No. 278.

Dauphin'e, T.C. 1976. Biology of the Kaminuriak population of barren ground caribou, Part 4: Growth, reproduction and energy reserves. Canadian Wildlife Service Report No. 38. Canadian Wildlife Service.

Dennis, B., J.M. Ponciano, S.R. Lele, M.L. Taper and D.F. Staples. 2006. Estimating density dependence, process noise, and observation error. Ecological Monographs 76: 323-341.

Gaillard, J.M., M. Festa-Bianchet and N.G. Yoccoz. 1998. Population dynamics of large herbivores: variable recruitment with constant adult survival. Trends in evolution and ecology 13: 58-63.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology 68: 1,364-1,372.

Gerrodette, T. 1993. TRENDS: software for a power analysis of linear regression. Wildlife Society Bulletin 21: 515-516.

Gunn, A., J. Boulanger and J. Williams. 2005a. Calf survival and adult sex ratio in the the Bathurst Herd of barren ground caribou 2001-2004. Department of Resources, Wildlife and Economic Development, Government of the Northwest Territories. Manuscript report No. 163.

Gunn, A., J. Nishi, J. Boulanger and J. Williams. 2005b. An estimate of breeding females in the Bathurst Herd of the barren-ground caribou, June 2003. Dept. of Environment and Natural Resources, Government of Northwest Territories. Manuscript Report No. 164.

Gunn, A. and D.E. Russell, editors. 2008. Monitoring Rangifer herds (population dynamics): Manual. Circumarctic Rangifer Monitoring and Assessment Network (CARMA), www.carmanetwork.com.

Harris, N.C., M.J. Kauffman and L.S. Mills. 2007. Inferences about ungulate population dynamics derived from age ratios. Journal of Wildlife Management 72: 1,143-1,151.

Hauser, C.E., A.R. Pople and H.P. Possingham. 2006. Should managed populations be monitored every year? Ecological Applications 16: 807-819.

Heard, D.C. 1985. Caribou census methods used in the Northwest Territories. McGill Subarctic Research Papers 40: 229-238.

Heard, D.C. and M. Williams. 1990. Caribou project summary and review. Department of Resources, Wildlife, and Economic Development, Government of Northwest Territories. Unpublished report.

Humbert, J.Y., L.S. Mills, J.S. Horne and B. Dennis. 2009. A better way to estimate population trends. Oikos 118: 1,940-1,946.

Kery, M. and M. Schaub. 2012. Bayesian population analyses using WinBugs: A hierarchical perspective. Volume 1. Academic Press, Watham, MA.

Nishi, J. 2014. Appendix M in in Nishi, J.S., B. Croft, J. Boulanger and J. Adamczewski. 2014. An estimate of breeding females in the Bathurst herd of barren-ground caribou, June 2009. Environment and Natural Resources, Government of the Northwest Territories. File report 144.

Nishi, J. S., B. Croft, J. Boulanger and J. Adamczewski. 2014. An estimate of breeding females in the Bathurst herd of barren-ground caribou, June 2009. Environment and Natural Resources, Government of the Northwest Territories. File report 144.

Otto, R.D., N.D. Simon, S. Couturier and I. Schmelzer. 2003. Evaluation of satellite collar sample size requirements for mitigation of low-level military jet disturbance of the George River caribou herd. Rangifer 14: 297-302.

Poole, K.G., A. Gunn and J. Wierzchowski. 2014. An operations guide to barren-ground caribou calving ground density, dispersion and distribution surveys, based on an assessment of the June 2007 and 2008 surveys, Northwest Territories and Nunavut. Environment and Natural Resources, Government of Northwest Territories. File Report 141.

Rivest, L.P., S. Couturier and H. Crepeau. 1998. Statistical methods for estimating caribou abundance using postcalving aggregations detected by radio telemetry. Biometrics 54: 865-876.

Thompson, W.L., G.C. White and C. Gowan. 1998. Monitoring Vertebrate Populations. Academic Press, San Diego, CA.

White, G.C. and B. Lubow. 2002. Fitting population models to multiple sources of observed data. Journal of Wildlife Management 66: 30