USE OF DISTANCE AND DOUBLE OBSERVER SAMPLING APPLIED TO BARREN-GROUND CARIBOU CALVING GROUND VISUAL SURVEY ESTIMATES

JOHN BOULANGER[1], ALLICIA KELLY[2], ROBIN ABERNETHY[2], JAMES HODSON[2] AND MITCH CAMPBELL[3]

[1]INTEGRATED ECOLOGICAL RESEARCH, [2]ENVIRONMENT AND CLIMATE CHANGE, GNWT, [3]DEPARTMENT OF ENVIRONMENT, GOVERNMENT OF NUNAVUT

2026

For review—contents may change in future drafts, including updates on collar data and enhanced demographic analyses.

MANUSCRIPT NUMBER XXX

Government of
Northwest Territories

TABLE OF CONTENTS, LIST OF FIGURES, LIST OF TABLES (TO BE ADDED)

INTRODUCTION

A challenge of ungulate aerial surveys is understanding how many animals that are in the survey area are likely to be detected (or missed) by observers, and how this affects the abundance estimates calculated from the observation data. Transect-based barren-ground caribou calving ground surveys have been used to estimate the abundance of breeding females and adult females on calving grounds for several barren-ground caribou herds in the Northwest Territories (NWT) and Nunavut since the 1980s. Visual surveys (i.e. caribou surveyed from aircraft by observers) are used for transects in areas with lower densities of caribou, where the ability to accurately estimate sightability of caribou on-transect will improve the accuracy of the resulting estimates.

There are three principal forms of bias in abundance estimates from wildlife surveys: counting bias, availability bias, and sightability bias. Counting bias occurs when the number of animals counted is consistently different than the number of animals actually present, and is more likely when group sizes are too large to count, leading to over or under counting of caribou groups (Elphick 2008, Campbell et al. 2025). To account for counting bias on barren-ground caribou calving ground surveys, an aerial photo survey is used in areas of higher density where caribou are counted from high-resolution images. Availability bias occurs when animals are not available to be sighted in surveyed areas such as in whale surveys where animals may under water when ships pass by. Availability bias is likely not an issue with caribou aerial surveys over open calving ground habitat. Sightability bias occurs when animals that are present in the survey area are missed (not detected) by the observer. Sightability of caribou during calving ground surveys is challenged due to wide variation in densities ranging from low abundance to large group sizes, varied ground conditions, and varying observer abilities which can make it difficult to effectively detect and accurately count caribou groups. When animals are systematically missed, abundance estimates will be negatively biased, i.e., lower than the true estimate. The main topic this report considers is sightability bias during visual abundance surveys of caribou on calving grounds, under what conditions observers miss caribou, and how best to estimate and correct for this bias, to more accurately estimate abundance.

Various methodologies have been proposed to estimate sightability on wildlife surveys including double observer methods and distance sampling (Laake et al. 2008a, Laake et al. 2008b, Buckland et al. 2010). The main purpose of each of these methodologies is to estimate the sightability of animals on the sampled transect area and account for differences in sightability related to factors that influence whether an animal in the area is detected. For example, one observer may simply be more skilled at detecting animals than another observer, but both observers may be more likely to detect large groups of animals than single animals. The likelihood of an animal being missed by an observer changes based on survey conditions or animal distribution; this is referred to as heterogeneity of detection probabilities. In addition, if most of the observation data results from animals that are relatively easy to detect (and thus influences our understanding of the factors that influence detection), and there is little data to inform animals that are hard to detect (because observers miss them), this can positively bias sightability estimates and lead to a negatively biased abundance estimates if modelling does not account for harder to detect animals (Laake and Collier 2024).

From the 1980s until 2011, calving ground visual surveys in NWT and Nunavut flew a 800 meter transect (400 meter strip on each side of the aircraft)and assumed that all caribou were detected within the 400m strip (Heard 1985). In 2011, the method was modified to a double dependent observer method with a front and rear observer on each side of the aircraft (Campbell et al. 2012). The two observers in this case represent an initial detection of caribou (the mark) and a potential re-detection of the caribou group (the recapture) and the double observer methods are analyzed using a mark-recapture analysis approach (Cook and Jacobsen 1979). Observers could communicate about what caribou groups were observed and therefore the method is termed dependent observer in contrast to independent observer methods where the two observers cannot communicate about groups they observed. This approach was an improvement over strip transect surveys (which assumed sightability=1), but it was potentially vulnerable to situations when sightability was lower and there was variation in caribou sightability as well as variation in observers' ability to sight caribou (Heard 1985, Boulanger et al. 2019, Campbell et al. 2019, Boulanger et al. 2022).

An alternative methodology used in aerial surveys is distance sampling where observers record the distance from the plane of each group of caribou to allow estimation of the decay in sightability as a function of distance from the survey plane. Distance sampling methods provide an additional way to correct for sightability by recognizing that observers are more likely to miss seeing caribou that are farther from the aircraft than caribou that are close to the aircraft. One challenge to this approach is that it requires collection of distance data, which is logistically challenging in areas with many caribou groups. However, distance sampling has the potential advantage of being more robust to sightability variation (Laake and Collier 2024).

In Nunavut, distance sampling methods have been used on spring and fall caribou abundance surveys, where caribou observations are assigned to one of five distance bins across a 1500m survey strip (Campbell et al. 2012, Campbell et al. 2015, Campbell et al. 2020, Campbell. et al. 2021). These surveys occurred in seasons with lower densities of caribou compared to calving ground surveys, making it possible to apply a binned-distance version of the distance sampling method. Results from all of these surveys estimate that detection probability declines with distance within the equivalent distance of the 400m survey strip (Boulanger and Campbell 2025) and this type of variation is not fully accounted for by double observer methods.

The benefits of distance sampling and the benefits of the double observer method have been leveraged in a combined analytical approach (Laake et al. 2008a). Distance sampling is relatively robust to heterogeneity variation of detection probabilities as long as sightability on the survey line is 1 or is estimated using double observer methods (Buckland et al. 1993, Rexstad et al. 2023, Laake and Collier 2024). Heterogeneity describes unequal detection probabilities between groups of caribou due to ground conditions, observer abilities, weather factors, or any other factor that might make 1 group more sightable than another. For this reason, the general approach used in distance sampling surveys that use mark-recapture methods is to focus mark-recapture (double observer) methods on estimating detection on the survey line with distance sampling used to model and estimate the decline in detection probability with distance (Laake et al. 2008a, Laake et al. 2008b). The distance sampling approach has been applied in surveys of ungulates, whales and other species worldwide (Thomas et al. 2009).

In the 2025 Bluenose-East and Bathurst calving ground surveys, we introduced a distance sampling component in addition to the double observer (mark-recapture) visual survey methods. The distance methods required observers to indicate which distance bin the observed caribou were in instead of simply counting caribou within the survey strip. As a second modification in 2025, data recorders were seated behind observers in the aircraft and separately recorded any caribou groups they observed that were missed by the primary and secondary observers (i.e. acted as a third observer on each side of the aircraft). This provided a higher count of caribou that was closer to the true number of caribou present in the survey area, and provided data to evaluate the detection probabilities estimated from the double observer (mark-recapture) analysis.

There were three main objectives for the research into visual survey methods during the 2025 surveys. First, we tested whether distance sampling methods, which model the decay in sighting probabilities with distance, could be effectively applied to a barren-ground caribou calving ground survey with data binned into 3 distance bins. Second, we compared the efficiency (ability to estimate missed caribou) of the double observer method of estimating detection probabilities versus a simpler 3 observer method (data recorder approach). Finally, we compared the overall efficiency of the various methodologies used to estimate caribou populations on calving ground surveys.

METHODS

Field methods

The 2025 Bluenose-East and Bathurst calving ground surveys were completed June 3-15, 2025 out of Kugluktuk and are reported in detail in Kelly et al. 2026a in prep (Bluenose-East herd) and Kelly et al. 2026b in prep (Bathurst herd). In brief, calving ground surveys are stratified strip-transect surveys, where a reconnaissance survey is flown to delineate and stratify the survey area into areas (strata) of relatively homogenous caribou densities and composition (primarily, breeding caribou present or absent). Following stratification, high caribou density areas are typically surveyed with photo aircraft that take high-resolution photos of the transects (later processed into 3D images with caribou counted on the images). Lower caribou density areas are typically surveyed with a visual aerial survey where observers count caribou on the transect.

Field methods specific to the 2025 surveys, including the distance sampling component and data recorder observations, are described below.

Both the reconnaissance survey and the visual survey component of the abundance survey used two Cessna 208 Caravan aircraft with 2 data recorders and 4 observers in each aircraft (Figure 1). The visual and reconnaissance surveys were both flown at 160 km/h and 120 m above ground. The aircraft had radar altimeter to ensure that altitude above ground remained constant.

The sample unit for these surveys was the observed group of caribou, regardless of size. Caribou were considered a group if they were noticed at the same time or if seeing an individual caribou was influenced by seeing other individual caribou nearby; e.g., a 'scattering' of 5 caribou observed at the same time would be a considered a group of 5 caribou (not 5 groups of 1 caribou). Observations of 1 caribou occurred when there was a single caribou in an area. Survey observations were collected using a Microsoft Surface Pro 10 with 5G tablet and Observation Collector (OC) software initially created for the 2011 Beverly survey (Campbell et al. 2012) and maintained by Caslys Consulting Ltd. The OC software was modified for the 2026 surveys to incorporate distance bins and data recorder observations. The touch-screen tablets recorded the names and positions of survey team members and aircraft for each survey session and recorded georeferenced survey observations and a lower-resolution flight track. Each caribou and incidental wildlife observation was linked to a GPS waypoint and survey transect. The data recorder also recorded percent snow cover and percent cloud cover. Percent snow and cloud cover could be entered with any caribou group observation (selecting from 0, 5, 10, 25, 50, 75 or 100%, or entering a value from 1-100%) or as a georeferenced observation independent of a wildlife observation. Percent snow and cloud cover values were applied to each subsequent caribou observation until a new value was entered by the data recorder (i.e. if there was 0% snow across an entire flight track it did not need to be recorded with each observation).

As in previous surveys since 2011, a dependent double observer pair was used to assess sightability of caribou. For the double-observer method, six basic steps were followed: (1) the front observer called out all groups of caribou (number of individuals and which distance bin) seen before they passed the front observer's position using an approximate dividing line between the front and rear observer perpendicular to the aircraft. Using a clock face as reference, this included caribou groups between 9 and 12 o'clock for the left side front observers and between 12 and 3 o'clock for the right side observers. (2) After the caribou passed behind this viewing area, the rear observer noted whether they had seen those same groups and if they had observed any additional groups. The rear observer waited to call out caribou until the group observed passed the dividing line (9 'clock or 3 o'clock). (3) When needed, the observers and data recorder discussed the group count to confirm they were calling the same or different group. (4) The data recorder noted whether an observation was seen by the "Front [only]", "Rear [only]" or "Both" observers. (5) New for this survey: If the data recorder saw any caribou groups not noted by either observer, they were recorded as seen by "Recorder". (6) Front and rear observers switched places between flights (Figure 1).

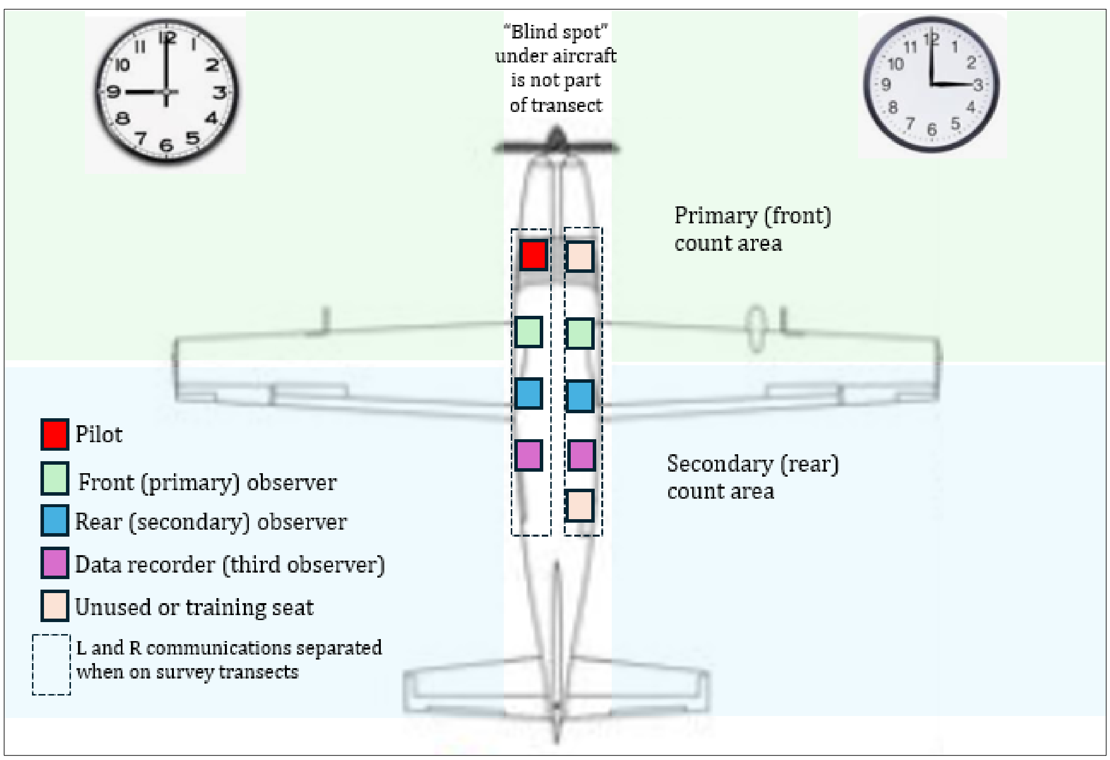

Figure 1: Observer and recorder positions for double observer method used in the Cessna 208 Caravan aircraft for the 2025 Bluenose-East calving ground survey. Distance sampling bins are not shown here. Data recorders acting as a third observer when time permitted was new for the 2025 surveys.

For the distance sampling methods, we established 3 distance bins of 0-200m, 200-400m, and 400-600m (instead of 5 bins used for Nunavut surveys), because of expected higher densities of caribou that would increase demands on observers to call and reconcile groups (compared to Nunavut surveys), and results from an empirical and simulation study of the Nunavut surveys that suggested that using 3 200-m distance bins would allow a decline in detection probability to be detected (Boulanger and Campbell 2025). The nearest two bins of 0-200m and 200-400m allowed us to analyze data that was consistent with previous calving ground surveys that used a 400m strip transect width.

To allow observers to identify the distance bin in which observations occurred, survey transects and bin widths (0-200m, 201-400m, and 401-600m) were marked on the wing struts using Norton-Griffith's 1978 formula as used in other calving ground surveys (Figure 2).

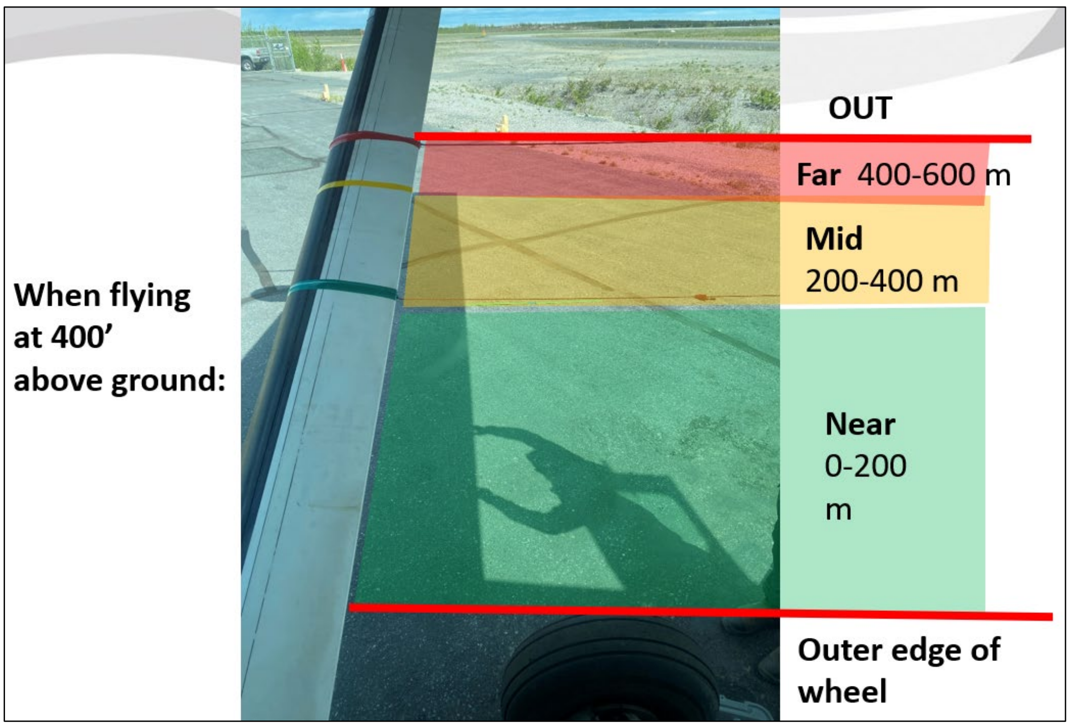

Figure 2: An annotated photo shows how strip widths are marked on the wing strut with coloured tape and streamers, and which ground area corresponds to each survey bin when the aircraft is 400' (120 m) above the ground. Survey bins that are farther from the aircraft appear thinner to the observer than survey bins closer to the aircraft.

Survey data

Caribou observations from both the reconnaissance and visual surveys of each calving ground (Bluenose-East and Bathurst herds) were used for detection modelling to estimate sightability. The reconnaissance data was filtered to maximize the applicability of the data set for visual survey estimates. Including reconnaissance group sizes larger than those observed during the visual abundance survey would likely create outlier observations that would affect relationships between distance sampling and double observer detection probabilities and group sizes. The largest group of caribou that occurred in the Bluenose-East and Bathurst visual survey strata was 35 so data from the reconnaissance survey was restricted to groups of 35 caribou or less.

Analysis methods

We detail each of the methods used to analyze the data. We note that each method takes a sample of caribou using line transects with a defined sampling area. The density from the defined sampling area is then used to estimate abundance for the entire survey strata (Jolly 1969). Where a sightability correction factor is estimated, it is used to adjust the density of caribou within the defined sampling area prior to estimating abundance for the entire survey strata.

Strip transect estimates (no correction for sightability)

We compared strip transect estimates (assuming perfect detection; Jolly 1969, Krebs 1998) based on 2 observers on each side of the aircraft. The strip transect approach basically treats each transect as a quadrat sample with the assumption that every caribou within the 0-400m survey strip was counted. Mean density from transects is estimated and multiplied by strata area to obtain an estimate of abundance. This approach was used for calving ground surveys prior to 2011 with one primary observer on each side of the aircraft. We also compared strip transect estimates using 3 observers per side of aircraft (including the data recorder). We did not calculate strip transect estimates using only the front observer on each side of the aircraft because on surveys prior to 2011, although there was typically one primary observer per side of aircraft, the pilot and navigator/data recorder contributed observations when possible, and other factors have changed (e.g. aircraft type and speed, expected density of caribou) that could affect detection of caribou.

Estimating a sightability correction factor

We estimated sightability using the 3 analytical methods described below: 1) Double observer (mark recapture), 2) Distance sampling (DS), 3) MRDS – mark-recapture (double observer) distance sampling.

For each method, covariates were used to describe and model factors influencing the sightability of caribou (Table 1). These included observer pair (the sample unit for dependent double observer methods is the pair of observers as opposed to single observers), caribou group size, log of group size, herd (calving ground surveyed, i.e., Bluenose-East or Bathurst), survey phase (reconnaissance or visual), strata surveyed, percent snow cover, percent cloud cover, and distance bin (0-200m, 201-400m, 401-600m). The basic approach used was to first test single covariate models followed by combining the most supported covariate models.

Table 1. Covariates used to model variation in sightability for distance and double observer pair analysis.

Covariate	Acronym	Description
Observer pair	obs	each unique observer pair
Group size	size	size of caribou group observed
	logsize	Natural log of group size
Herd	herd	Calving ground being surveyed
Phase	Phase	Reconnaissance or visual
Strata	strata	Visual survey strata (reconnaissance pooled)
Snow cover	snow	Percent cover (continuous)
distance	distance	Distance bin where group observed
Cloud cover	cloud	Percent cover (continuous)

The fit of models within each analysis (double observer (MR), DS, and MRDS) was evaluated using the Akaike Information Criterion corrected for small sample size (AIC_c). The model with the lowest AIC_c score was considered the most parsimonious, thus minimizing estimate bias and optimizing precision (Burnham and Anderson 1998). The difference in AIC_c values between the most supported model and other models (ΔAIC_c) was also used to evaluate the fit of models when their AIC_c scores were close. In general, any models with a ΔAIC_c score of less than 2 between them were considered to have equivalent statistical support. Overall model fit was also assessed using goodness of fit tests (Buckland et al. 1993, Buckland et al. 2004) as well as graphical comparison of detection functions with histograms of frequencies of observations from the survey.

Abundance estimates for strata were derived based on transect lengths and strata areas for the best fitting detection model.

Analyses were conducted in program *R* (R Development Core Team 2009) with plots being produced using the *ggplot* (Wickham 2009) R package with GIS analyses conducted in QGIS (QGIS Foundation 2020) and the *sf* (simple features) (Pebesma 2018) R package. The *MRDS* R package (Laake et al. 2012) was used to build mark-recapture, distance sampling and joint MRDS models. In addition, the *AICmodavg* (Mazerolle 2016), *lubridate* (Grolemund and Wickham 2011), and *ddply* (Wickham 2011) R packages were used in the analysis.

Double observer (mark-recapture) methods

The double observer sampling uses 2 observers on each side of the aircraft and can be considered a two-sample mark-recapture trial in which some caribou are seen ("marked") by the ("session 1") primary (front) observer, and some of these are also seen by the second (rear) observer ("session 2"). The second observer may also see caribou that the first observer did not see. This process is analogous to mark-recapture except that caribou are sighted and re-sighted rather than marked and recaptured. The double observer data are then analyzed using a mark-recapture analysis. Because the sighting probability of the second observer was not independent of the primary observer (they hear each other call observations), an independence removal estimator was used to model the mark-recapture estimates (Laake et al. 2008a, Laake et al. 2008b). The removal model assumed that the initial sighting probability of the primary and secondary observers was equal (Cook and Jacobsen 1979, Buckland et al. 1993, Buckland et al. 2004, Buckland et al. 2010).

Observers were switched midway in each survey day (on most days there were two flights with a re-fueling stop between them), and covariates were used to account for any differences that were caused by unequal sighting probabilities of primary and secondary observers.

Data recorder observations (any caribou groups that both observers missed but that the data recorder observed) were not incorporated directly into mark-recapture estimates because the sequential approach of the double dependent observer method makes direct incorporation of a 3rd observer problematic given that MRDS and other R packages only allow 2 observers. However, data recorder observations were used to test if double observer observations were adequately estimating groups missed by observers.

Front and rear observers working on the same side of an aircraft were grouped into pairs to model the effect of observer on sightability. Detection probabilities for each observer pair were estimated to allow a general comparison of observer pair efficiency. For the dependent double observer method used here, the probability of detection (p_{1x}) is based upon the conditional detection probability of observer 2 observing a caribou that observer 1 misses. This is estimated as ratio of total number of caribou observed by both observers divided by the total number observed by both observers and the caribou only observed by observer 2. Given the assumption of equal observation probabilities between observers, the double observer probability is then $1-(1-p_{1x})^2$.

We then estimated the total number of groups that the observer pair encountered (including those missed) as the total count of caribou groups observed (P+S+B) divided by the 2x observed probability ($E(P+S+B)= (P+S+B)/ p_{2x}$).

To explore the effects of using 3 observers on each side of the aircraft, we compared observations by the front and rear observer combined and the estimated number of caribou groups encountered with total observations including those of the 3rd observer (data recorder). For each double observer pair plus data recorder grouping, we summed the number of caribou groups observed by the primary observer (P), secondary observer (S), both observers (B), and the data recorder (R), and summed all observations without data recorders (P+S+B) and with data recorders (P+S+B+R). We calculated the proportion of observations where observer 1 was primary, single (p_{1x}) and combined detection probabilities (p_{2x}) for each observer pair. We calculated the expected total number of groups observed (P+S+B)/p2x. Finally, we calculated the ratio between groups observed with data recorder observations included as (%DR= E(P+S+B+R))/(P+S+B+R).

Factors influencing the sightability of caribou were included as covariates in the mark-recapture models to estimate sightability for the dependent double observer analysis as described in Table 1. The MR modelling analysis was conducted using both the 400m and 600m strip data sets. The 600m data set was the primary focus of the analysis given that it represented the main data set collected. The 400m data set was similar to previous surveys and therefore the same set of models were re-run for the 400m data set for comparison purposes.

Distance sampling (DS) methods

Distance sampling estimates were derived by fitting a detection function to the 3 survey bins (0-200m, 200-400m, 400-600m) under the assumption that sightability will decrease with distance from the survey plane (Buckland et al. 1993, Thomas et al. 2009). The main detection functions considered were half-normal, hazard rate, and uniform. Factors influencing the sightability of caribou were included as covariates in the DS models to estimate sightability as described in Table 1.

Data recorder observations (caribou that both observers missed but that the data recorder observed) were included in the distance sampling analysis, as inclusion of data recorder observations in distance methods as a third "platform" is acceptable and recommended as a means of increasing detection probabilities (Buckland et al. 2004, Thomas et al. 2009). Distance sampling estimates were generated using just the front and rear observer observations (pooled observations from 2 observers) and using the front, rear data recorder observations (pooled observations from 3 observers).

Mark-recapture/distance sampling (MRDS) methods

A combined distance sampling and mark-recapture used distance sampling to estimate change in detection with distance (the shape of the detection function), and mark-recapture to estimate sightability on the survey line (anchor the detection function). This approach potentially ensured a more robust estimate than using distance sampling methods alone which assume that the probability of detection of caribou groups at 0 distance from the plane is 1 (Borchers et al. 1998, Buckland et al. 2004, Laake et al. 2008a, Laake et al. 2008b, Buckland et al. 2010, Laake et al. 2012). The distance sampling component was based on 2 observers (i.e. did not include the data recorder observations). As in other models, covariates were used to improve estimates as described in Table 1. For MRDS it was possible to dichotomize sources of variance in the analysis into variation caused by estimation of detection and variation caused by estimation of encounter rate. Encounter rate variation is a measure of variation in density across transects (Buckland et al 2010).

Evaluation of methods

Table 2 summarizes methods employed for the analysis of the 2025 survey. The main differences between methods are 1) whether sightability was estimated (strip transects do not estimate sightability), 2) the method used to estimate sightability (double observer vs distance vs double observer/distance) and 3) whether 3rd observer (data recorder) observations were included.

Table 2: Summary of methods used to analyze the 2025 survey data

Method	Data included	Assumptions
Strip transect (ST)	All observations within the 400m survey strip including data recorder observations	All caribou are observed within the 400 m survey strip
Double observer Mark recapture (MR)	All observations within the 400m or 600m survey strip not including data recorder observations	Probability of caribou being observed is estimated as the proportion of caribou groups missed by the primary observer (see details on double observer)
Distance sampling (DS)	All within the 600m survey strip including data recorder observations. Double observer data pooled into single observation.	Probability of caribou being observed decreases with distance from plane. Probability of observing caribou at 0 distance from the plane is 1.
Mark-recapture/distance sampling (MRDS)	All observations within 600m of the survey strip not including data recorder observations	Assumptions of mark-recapture and distance sampling with the exception that the probability of observing a caribou in at the survey strip (at 0 distance from the plane) is estimated using MR

A challenge with empirical comparison of estimators is that the true number of caribou on the strata is unknown. Therefore, inference on bias can mainly be derived through assessment of the assumptions made with each method, and comparison of relative estimates between estimators. There were exact analysis questions that were used to evaluate each methodology. The main questions arose based on the assumption of whether sightability =1 (strip transects) versus estimation of sightability (distance and mark-recapture methods) versus the inclusion of a third observer as an alternative to estimating sightability. The questions being asked are shown in Table 3, which will be repeated in the results sections with further comments on each row item.

Table 3. Hypotheses from comparison of estimators with implications.

Hypothesis	Implications/biases
Distance sampling/2x observer methods estimates of sightability are less than 1	Strip transects are negatively biased
Distance sampling detects decline in detection probability that is not detected when distance is used as a covariate in double observer models	Increasing heterogeneity with distance from plane prevents the detection of decline in detection probabilities using double observer methods suggesting double observer methods are negatively biased
Estimates with data recorder included (3 observers) from strip transects are higher than estimates from strip transects using 2 observers	3 observers detect more caribou than 2 observers
Estimates with data recorder included from strip transects are higher than estimates from double observer models	Double observer models are negatively biased (do not fully account for missed caribou)
Distance sampling estimates with data recorder observations are higher than strip transect and/or double observer models using 2 observers	Distance sampling estimates using 3 observers provide a less biased estimate due to robustness to unequal sighting probabilities. Inclusion of data recorder observations (a 3rd observer) helps meet the assumption of sightability=1 on survey line
Mark-recapture distance sampling (MRDS) estimates are lower than distance sampling estimates with data recorder	Including data recorder observations provides a more efficient means to mitigate imperfect observer detection than double observer (mark-recapture) methods for distance sampling estimates

RESULTS

Survey data

Caribou observations from the 2025 Bathurst and Bluenose-East visual surveys are shown in Figure 3.

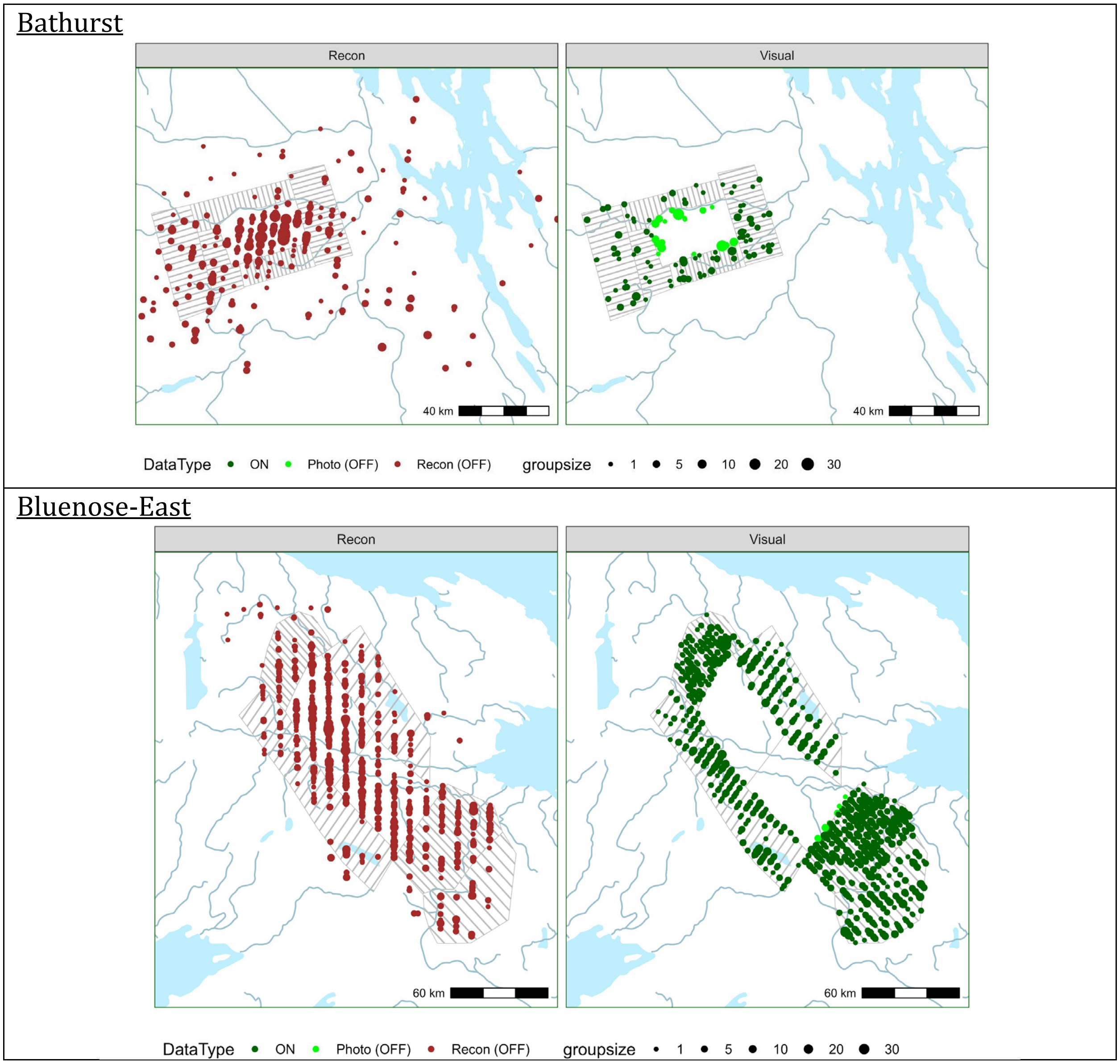

Figure 3: Data used for visual survey analysis. All on transect observations from the visual abundance survey were used for abundance estimates. For modelling of sighting probabilities, all caribou observations from the visual abundance survey and from the reconnaissance survey, including observations that were excluded from revised visual survey boundaries, were used. Recon (OFF) and Photo (OFF) refer to observations excluded from abundance estimates when survey strata boundaries were revised but still included in modelling sightability probabilities (Kelly et al. 2026a in prep).

The data used in subsequent analysis to compare methods of estimating sightability and to estimate abundance in the visual strata is shown in Table 4. Coverage (at 600m strip) of the visual strata was high for the Bathurst (46-58%) and moderate for the Bluenose-East (18-35%). Differences in coverage were mainly due to the large difference in overall survey areas (km^2) for the two herds, and to reduce variance around Bathurst estimates, because densities observed during the Bathurst reconnaissance survey were low and patchy. More detailed information on strata design and detailed data summaries are provided in survey reports (Kelly et al. 2006a in prep (Bluenose-East herd), Kelly et al. 2006b in prep (Bathurst herd)).

Table 4: Summary of Bathurst and Bluenose-East 2025 visual strata summarizing the main strata attributes and caribou detected in each strata. Coverage is based on the sum of strip width areas divided by the area of each stratum.

Strata	Area	transects	Ave transect (km)	Coverage (strip)		Caribou observed on transect	
				400m	600m	400m	600m
Bathurst							
VisEast	723.50	19	17.0	0.36	0.54	39	54
VisFarWest	844.57	19	17.2	0.31	0.46	24	47
VisNorth	443.54	20	9.7	0.35	0.53	10	16
VisSouth	349.76	16	10.6	0.39	0.58	49	51
VisWest	355.30	13	12.0	0.35	0.53	9	13
Total						131	181
Bluenose-East							
VisEast	2515.19	14	26.6	0.12	0.18	253	410
VisFarSouth	2366.84	14	30.5	0.14	0.22	441	604
VisNorth	1408.44	15	27.0	0.23	0.35	298	487
VisSouth	2670.11	18	38.5	0.21	0.31	1006	1434
VisWest	3312.14	25	25.5	0.15	0.23	499	706
Total						2497	3641

Using Jolly's (1969) strip-transect estimator (and 400 m strip width) with no correction for sightability, visual strata abundance estimates were 367 adult caribou (95% CI 281-480) for the Bathurst herd and 14,525 adult caribou (95% CI 12,631-16,702) for the Bluenose-East herd (including data recorder observations; see Table 9 for comparable estimates without data recorder observations, and for estimates derived from the other methods).

Double observer (mark-recapture) results

Observer pairs

There were 8 unique observer pairs in the dataset. Summary statistics were calculated for these 8 pairs (8 groups with the data recorder for each pairing) (Table 5). A single observer would see 70 – 96% of groups of caribou available on-transect. With the double observers, 95-99% of the estimated true number of groups would be counted.

Including data recorder observations

Comparing the estimated number of groups using the double observer (mark-recapture, MR) method to the total number of groups counted when using 3 observers (including data recorder observations) showed that the number of caribou groups is underestimated by the double observer (MR) method. The estimated counts using the double-observer method varied by observer pair and were 75-97% of the counts from 3 observers (both observers plus the data recorder). Figure 4 shows frequencies of detection by observer pairs as a function of detection type (seen by front observer only, rear observer only, both front and rear observer, or data recorder only).

Table 1: Double observer pairings with associated summary statistics. Columns heading indicated numbers of groups observed by primary (P), secondary (S), both observers (B), data recorder (R), and summations of observations without data recorders (P+S+B) and with data recorders (P+S+B+R). Also given are proportion observations where obs 1 was primary (%P), single (p_{1x}) and combined detection probabilities (p_{2x}) for each observer pair. The expected total number of groups observed (P+S+B)/p2x is given. The ratio between groups observed with data recorder observations included (%DR= E(P+S+B+R))/(P+S+B+R) is also given.

Pair	P	S	B	R	P+S+B	P+S+B+R	%P	p_{1x}	p_{2x}	E(P+S+B)	%DR
1	33	56	372	21	461	482	0.68	0.87	0.98	469	97.3%
2	45	45	451	30	541	571	0.92	0.91	0.99	545	95.5%
3	11	55	131	20	197	217	0.18	0.70	0.91	216	99.5%
4	8	6	40	6	54	60	0.58	0.87	0.98	55	91.6%
5	10	24	88	49	122	171	0.85	0.79	0.95	128	74.8%
6	19	30	206	11	255	266	0.47	0.87	0.98	259	97.4%
7	6	10	265	8	281	289	0.69	0.96	1.00	281	97.4%
8	23	38	373	20	434	454	0.65	0.91	0.99	438	96.4%

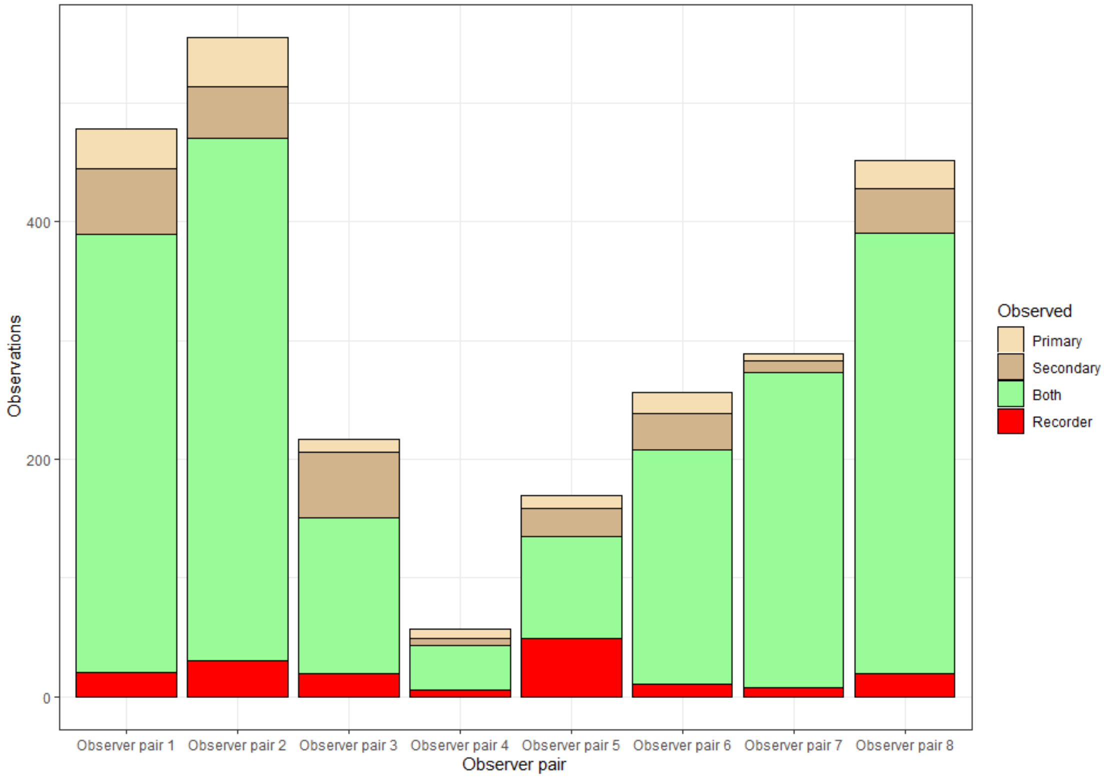

Figure 4. Frequencies of detection by observer pairs as a function of detection type (seen by front observer only, rear observer only, both front and rear observer, or data recorder only).

Double observer (MR) model comparison

The double observer (MR) modelling analysis was conducted using both the 400m and 600m strip data sets. The 600m data set was the primary focus of the analysis given that it represented the main data set collected. The 400m data set was similar to previous surveys and therefore the same set of models were re-run for the 400m data set for comparison purposes (Appendix A).

Model selection results suggested that double observer probabilities varied by the log of group size, observer pair, survey phase and herd (Table 6: Model 1). Distance from plane as a continuous (Models 11 and 12) or as a binned factor (i.e. 0-200,200-400, 400-600) was not supported as a covariate suggesting that the double observer data was unable to detect the decrease in detection probabilities with distance. Overall abundance estimates were very similar between models. The likely reason for this is that the overall double observer detection probability from Model 1 was 0.97 (SE=0.005) which infers that the majority of caribou within the strip were detected therefore reducing the effect of covariates on overall model estimates. Model selection results from the 400m strip was similar to the 600m strip analysis with the same models being most supported (Appendix A).

Table 2: Model selection results for double observer (mark recapture) analysis using 600m strip width. Model selection criteria are AIC_c, the difference in AIC_c values between the *i*th and most supported model 1 (ΔAIC_c), Akaike weights (w_i), and number of parameters (K), and log-likelihood. Total visual estimates are also given for each herd.

No	MR model	Model selection					BNE N		Bathurst N	
		AIC_c	ΔAIC_c	wi	K	LL	N	CV	N	CV
1	logsize + Obs + Phase + Herd	1500.81	0.00	0.70	11	-739.3	13,927	0.056	330	0.204
2	logsize + Obs + Phase * Herd	1502.86	2.05	0.25	12	-739.3	13,927	0.056	331	0.204
3	logsize + Obs + Phase	1506.01	5.20	0.05	10	-742.9	13,934	0.056	330	0.205
4	logsize + Obs	1528.39	27.59	0.00	9	-755.1	13,989	0.056	324	0.205
5	logsize + Phase * Herd	1565.69	64.88	0.00	5	-777.8	13,913	0.056	327	0.204
6	logsize + Phase	1572.17	71.36	0.00	3	-783.1	13,922	0.056	329	0.205
7	logsize + Herd	1572.76	71.95	0.00	3	-783.4	13,941	0.056	331	0.204
8	logsize + cloud	1576.40	75.59	0.00	3	-785.2	13,944	0.056	325	0.205
9	logsize + snow	1580.69	79.88	0.00	3	-787.3	13,958	0.056	322	0.206
10	logsize	1587.65	86.84	0.00	2	-791.8	13,964	0.056	323	0.206
11	logsize + distance	1589.47	88.66	0.00	3	-791.7	13,964	0.056	323	0.206
12	logsize * distance	1591.43	90.62	0.00	4	-791.7	13,964	0.056	323	0.206
13	size	1605.56	104.75	0.00	2	-800.8	13,973	0.056	323	0.206
14	distance* Phase	1628.58	127.77	0.00	4	-810.3	13,966	0.056	323	0.206
15	constant	1642.36	141.55	0.00	1	-820.2	14,040	0.056	322	0.207

A plot of predictions from Model 1 illustrates the effect of group size, observer pair, survey phase and herd (Figure 5). In most cases group sizes of greater than 10 have double observer detection probabilities that are close to 1 (regardless of distance from plane).

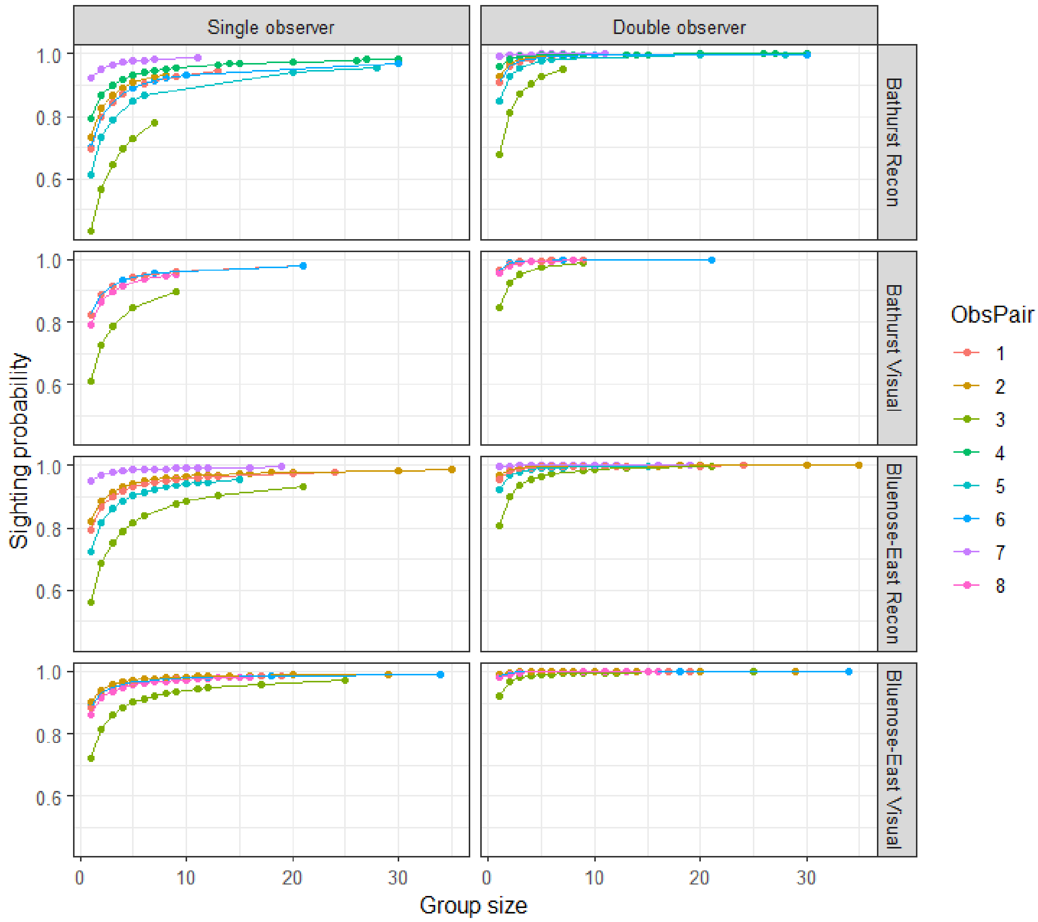

Figure 5: Predictions of detection probability from double observer analysis (600m) based on observer pairs, group size, survey phase, and herd.

Distance sampling (DS) results

Data summary

Overall plots of observations by distance demonstrate a decline in the frequencies of detection with successive distance bins. The proportion of observer detection types (front observer, rear observer, both front and rear, or data recorder only) were roughly proportional for each distance bin (Figure 6).

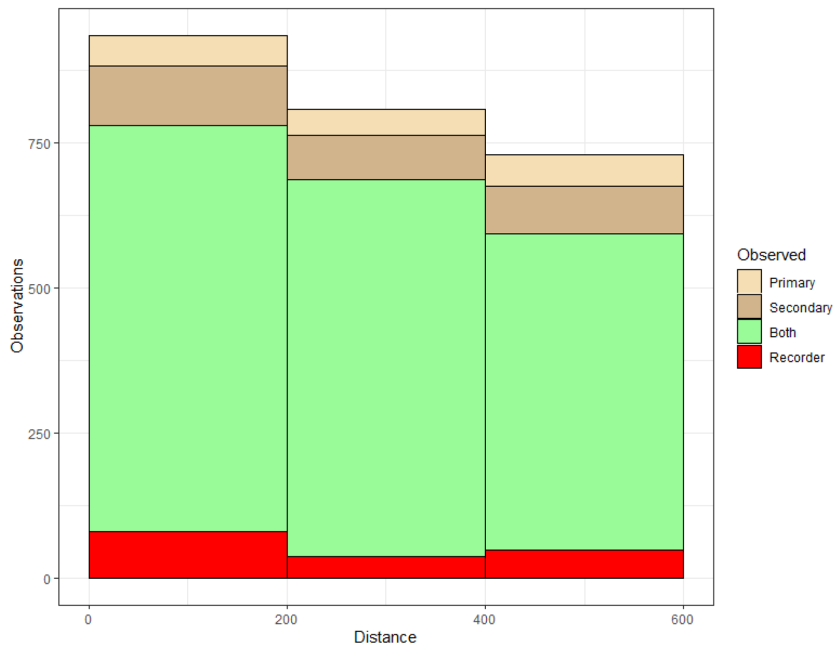

Figure 6: Pooled distance detection frequencies for visual and reconnaissance surveys. Sub-bars indicate whether the caribou group was observed by the primary, secondary or both observers, or only by the data recorder.

Detection frequencies were relatively similar for the two caribou herds with similar proportions of observer type for each distance bin (Figure 7).

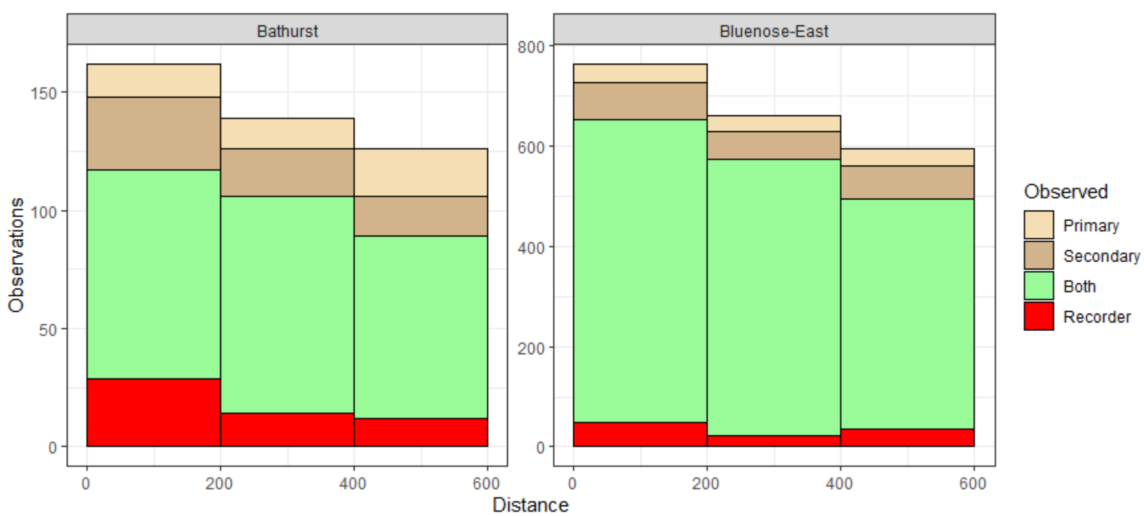

Figure 7: Distance detection frequencies for the Bluenose-East and Bathurst herds. Sub-bars indicate whether the caribou group was observed by the primary, secondary or both observers, or only by the data recorder.

If distance detection frequencies are viewed by herd and survey phase the decline in detections with distance was greatest for the Bluenose-East reconnaissance and Bathurst visual surveys with lesser declines for the Bathurst recon and Bluenose East visual surveys (Figure 8).

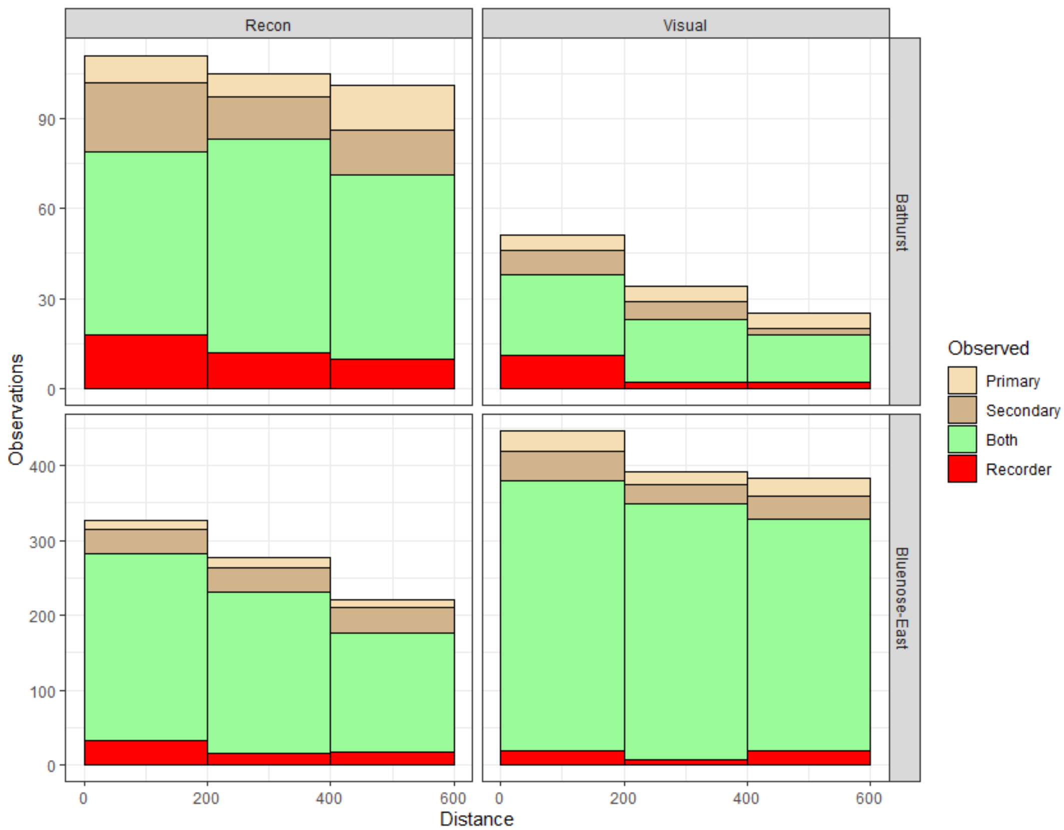

Figure 8: Detection frequencies for distance bins for visual and reconnaissance phases by herd. Sub-bars indicate whether the caribou group was observed by the primary, secondary or both observers, or only by the data recorder. Note different y-axis scales.

Distance detection frequencies by observer pair also suggested variation by observers with observer 3 and 6 showing minimal decline in detection frequency by distance bin. Observer pairs 1, 2, 7, and 8, which had higher sample sizes, did show declines with distance (Figure 9).

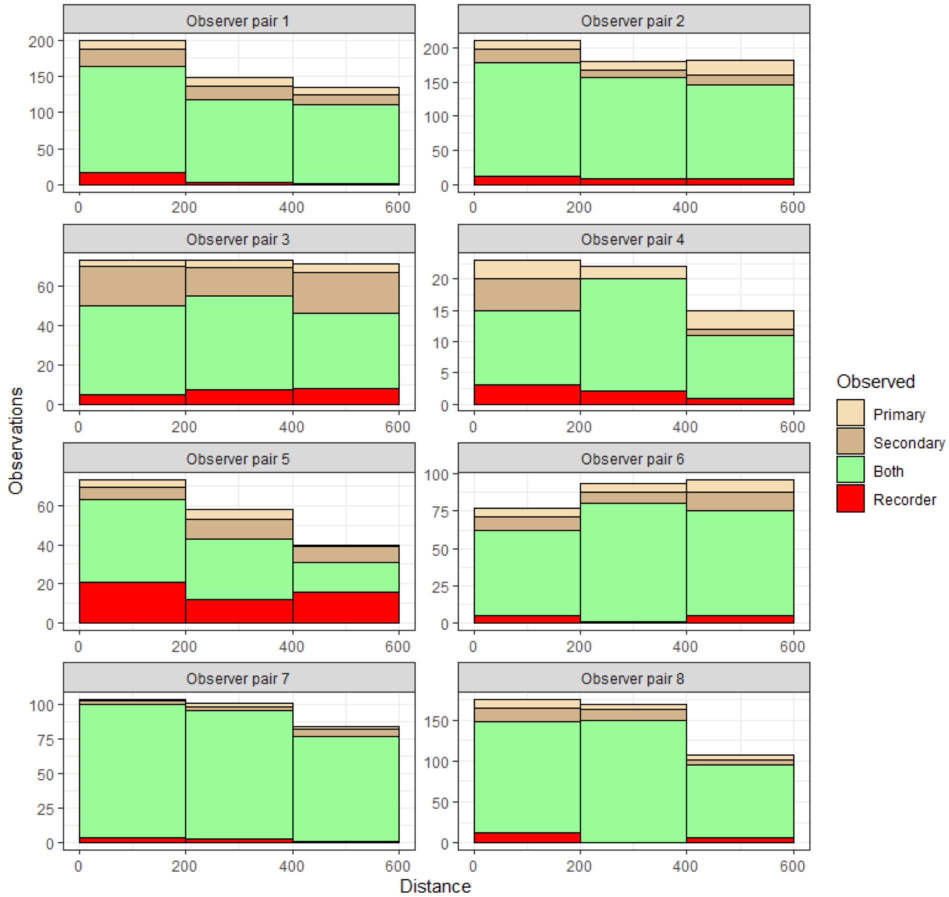

Figure 9: Detection frequencies for distance bins for observer pairs. Sub-bars indicate whether the caribou group was observed by the primary, secondary or both observers, or only by the data recorder. Note different y-axis scales.

Group size also influenced the shape of distance detection frequencies with decline by distance bin most notable for group sizes of 5 caribou or less (Figure 10).

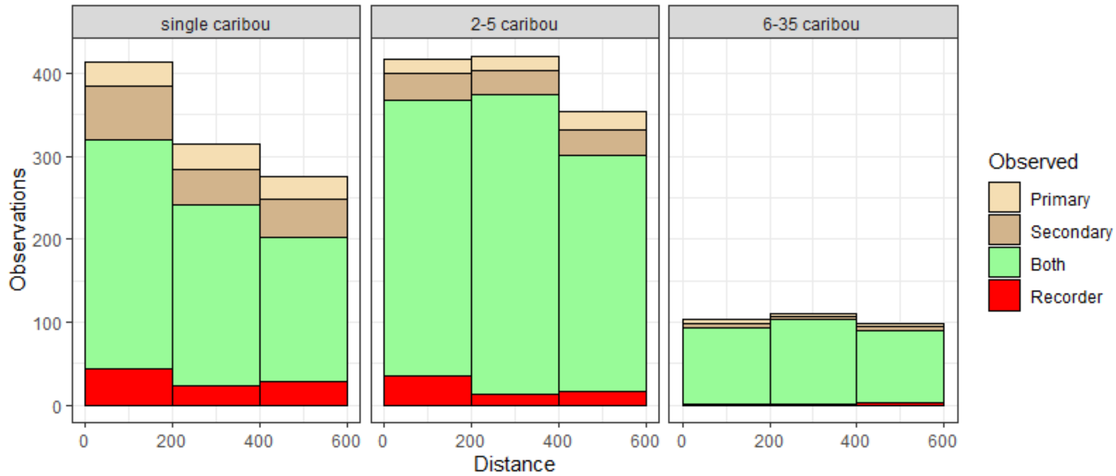

Figure 10: Detection frequencies a function of ranges of group size. Sub-bars indicate whether the caribou group was observed by the primary, secondary or both observers, or only by the data recorder.

Cloud cover and snow cover were at low levels for the majority of the survey and therefore did not influence survey conditions greatly (Figure 11 and Table 7).

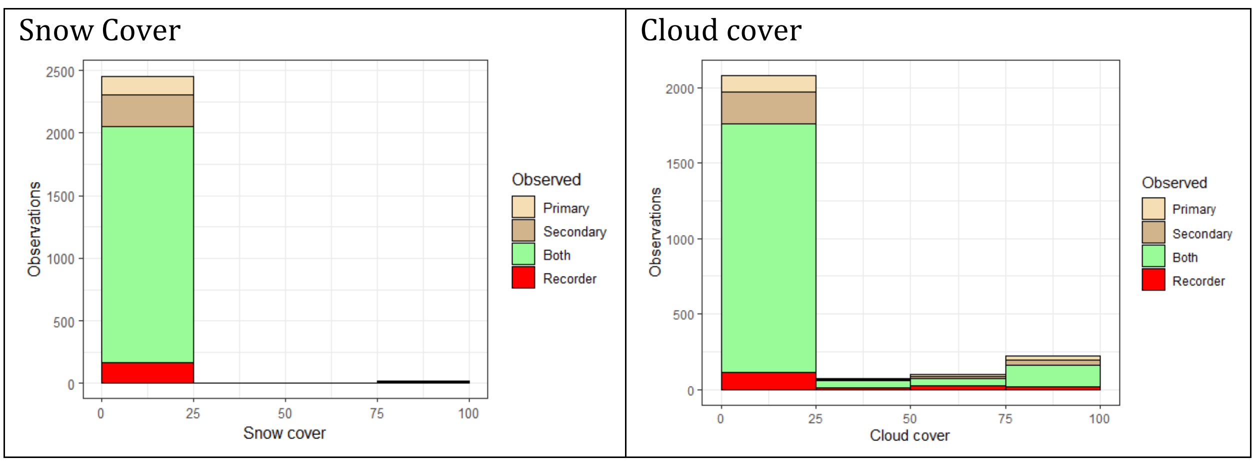

Figure 11: Distribution of snow and cloud cover during the survey. Sub-bars indicate whether the caribou group was observed by the primary, secondary or both observers, or only by the data recorder.

Distance sampling (DS) model comparison

The DS data sets for modelling included all observations, including data recorder observations. Model selection focused initially on determining optimal detection functions with half-normal, hazard rate, and uniform detection functions considered (Table 7). Of models considered, the hazard rate (Model 18) and half normal (Model 17) models both showed support with the difference in AICc values being 2 while the uniform detection functions which assume detection probabilities of 1 regardless of distance were not supported. For half normal and hazard rate, models were further considered with group size and log of group size covariates which were likely to influence detection function shape. The scale parameter for the hazard rate model had issues with convergence (scale parameter close to 0) which was likely due to roughly equal decrease in detection across all survey bins which made estimation of a shoulder area problematic. For this reason, the half-normal detection function with size as a covariate was used for subsequent model building (Model 9). Sensitivity of estimates to half normal and hazard rate was minimal with differences of <1% in estimates for Bluenose-East and Bathurst abundance estimates (Model 6 vs Model 9). A final model (Model 1) that allowed detection functions to vary by the group size, observer pair, and an interaction between survey phase and herd was most supported. A model with the log of group size was also supported (Model 2) with minimal difference in estimates. A strata-specific model (Model 4) was less supported suggesting that strata-specific variation could be explained by observer-specific, group size, and survey phase covariates. A model with covariates from Model 1 but with hazard rate as the detection function was slightly more supported than Model 1 but did not converge adequately which was likely due to lack of data to estimate the scale parameter (as discussed later). Abundance estimates were relatively stable for the Bluenose-East herd for each model. Estimates were less stable for the Bathurst herd with an increase when PhaseXHerd was a covariate in the model. This was likely due to the difference in detection histogram shape between the Bathurst recon and visual surveys (Figure 8).

Table 3: Model selection results for distance sampling analysis. Detection functions (DSDF) are given (hn=half-normal, hr=hazard rate, un=uniform). Model selection criteria are AIC_c, the difference in AIC_c values between the *i*th and most supported model 1 (ΔAIC_c), Akaike weights (w_i), and number of parameters (K), and log-likelihood. Total visual estimates are also given for each herd.

No	DSDF	DS model	Model selection					BNE N		Bathurst N	
			AIC_c	ΔAIC_c	wi	K	LL	N	CV	N	CV
1	hn	size + Obs + Phase * Herd	5388.9	0	0.49	12	-2682.3	15165	0.058	458	0.157
2	hn	logsize + Obs + Phase * Herd	5389.2	0.25	0.43	12	-2682.5	15260	0.060	449	0.159
3	hn	size + Obs	5392.8	3.90	0.07	9	-2687.3	15276	0.058	390	0.143
4	hn	size + Strata	5399.3	10.37	0.00	9	-2690.6	15030	0.059	450	0.155
5	hn	size + Phase * Herd	5401.7	12.71	0.00	5	-2695.8	14797	0.059	455	0.158
6	hr	logsize	5403.2	14.25	0.00	3	-2698.6	15119	0.060	382	0.144
7	hr	logsize + Phase * Herd	5403.4	14.48	0.00	6	-2695.7	14833	0.060	486	0.227
8	hr	size	5403.6	14.63	0.00	3	-2698.8	15032	0.058	381	0.143
9	hn	size	5403.7	14.71	0.00	2	-2699.8	15094	0.058	380	0.143
10	hn	logsize	5403.8	14.86	0.00	2	-2699.9	15198	0.060	381	0.144
11	hn	logsize + snow	5404.0	15.03	0.00	3	-2699.0	15127	0.060	378	0.144
12	hn	logsize * Phase * Herd	5404.1	15.12	0.00	8	-2694.0	14572	0.060	445	0.169
13	hn	logsize + Phase	5405.0	16.04	0.00	3	-2699.5	15079	0.060	376	0.146
14	hn	logsize + cloud	5405.8	16.82	0.00	3	-2699.9	15227	0.060	380	0.145
15	hn	logsize * Phase	5406.0	17.05	0.00	4	-2699.0	14774	0.061	370	0.145
16	hn	logsize + Phase + Herd	5406.7	17.79	0.00	4	-2699.4	15050	0.060	387	0.151
17	hn	constant	5409.7	20.70	0.00	1	-2703.8	15893	0.060	387	0.149
18	hr	constant	5411.6	22.63	0.00	2	-2703.8	15858	0.066	386	0.151
19	un	constant	5431.5	42.59	0.00	0	-2715.8	14165	0.056	345	0.147

There were insufficient degrees of freedom to test goodness of fit for the fully parameterized model 1 with an overall chi-square value of 2.12. Goodness of fit to the base half normal (model 17), which had adequate degrees of freedom suggested reasonable fit (χ^2=2.1, df=1, p=0.14). A plot of detection functions with various covariates demonstrates fit of the model to various sources of variation (Figure 12). For herd X phase detection functions for the Bathurst recon and Bluenose-East visual there were reduced slopes compared to Bathurst visual and Bluenose-East recon as also suggested by frequency plots (Figure 8). The slope of detection functions was reduced with increasing group size as suggested by detection frequencies (Figure 10). The slopes of observer pair functions also varied with detection function for 6 and 3 being relatively flat compared to other observer pairs as also suggested in frequency plots (Figure 9).

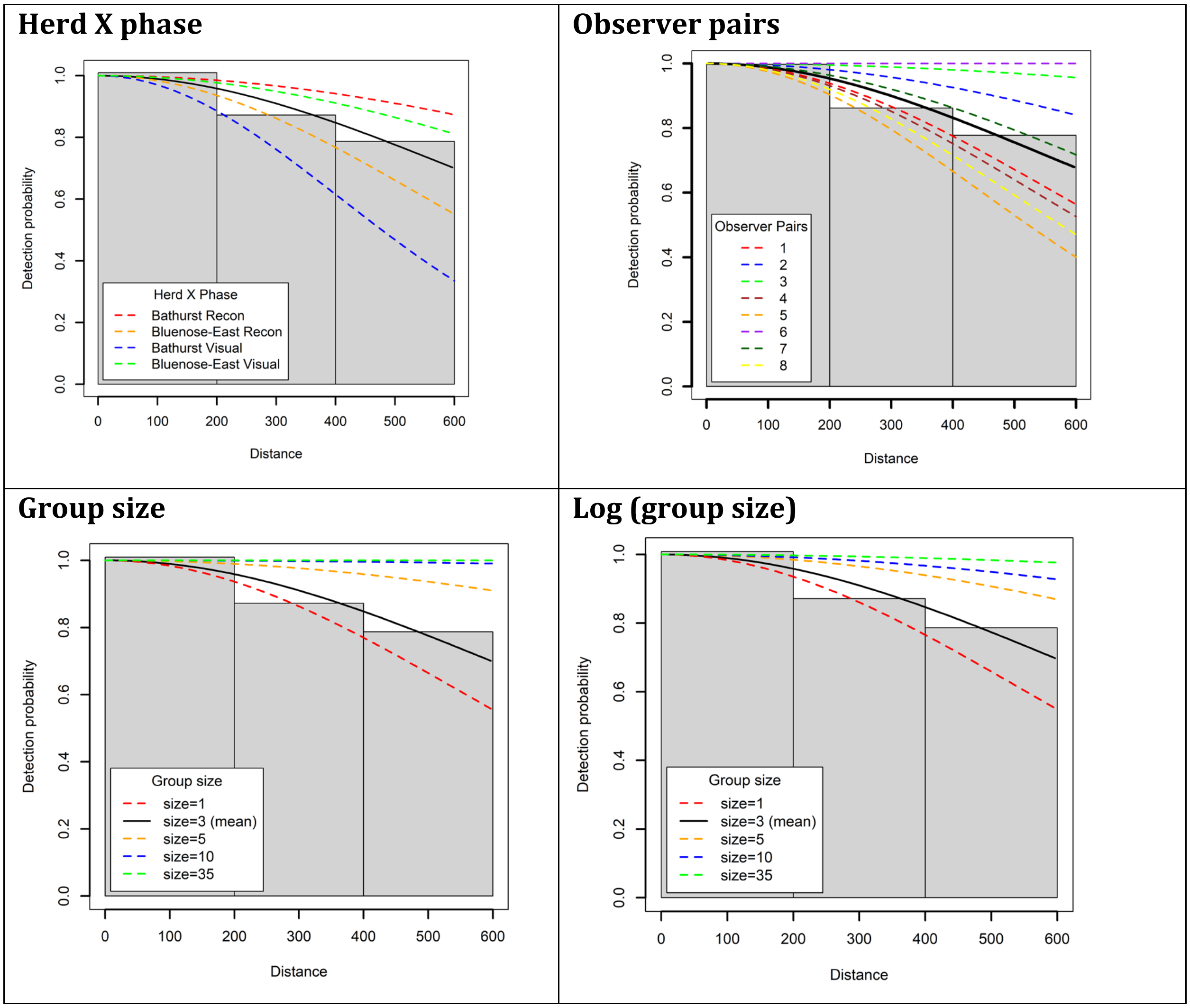

Figure 12: Predictions of model 1 (Table 7) for each covariate in the model. The black line in each plot is the mean intercept detection function. Group size and log of group size (Model 2) are also shown for comparison.

Mark-recapture/distance sampling (MRDS) results

Model selection for the combined analysis focused on the most supported distance sampling and mark-recapture models with the most supported distance sampling and mark-recapture models being also supported in the combined analysis (Table 8, Model 1). In each case, distance sampling and mark-recapture detection probabilities varied by the log of group size, observer pair, survey phase, and herd. A hazard rate (Model 2) and half normal detection function (Model 1) were considered with half normal showing higher support with estimates being relatively close (<1% difference) for both herds. The log of group size was more supported than size for both hazard rate and half-normal models.

Table 4: Model selection results for mark-recapture distance sampling (MRDS) analysis. Detection functions (DSDF) are given (hn=half-normal, hr=hazard rate, un=uniform). Model selection criteria are AIC_c, the difference in AIC_c values between the *i*th and most supported model 1 (ΔAIC_c), Akaike weights (w_i), and number of parameters (*K*), and log-likelihood. Total visual estimates are also given for each herd.

				Model selection					Bluenose-East		Bathurst	
No	DSDF	DS model	MR Model	AICc	ΔAIC_c	wi	K	LL	N	CV	N	CV
1	hn	logsize + Obs + Phase * Herd	logsize + Obs + Phase + Herd	6537.2	0.00	0.42	23	-3245.4	14,950	0.060	402	0.164
2	hr	logsize + Obs + Phase * Herd	logsize + Obs + Phase + Herd	6538.9	1.67	0.18	24	-3245.2	14,491	0.058	410	0.164
3	hr	logsize + Obs	logsize + Obs + Phase + Herd	6539	1.81	0.17	21	-3248.3	14,690	0.062	348	0.168
4	hn	logsize + Obs + Phase * Herd	logsize + Obs + Phase * Herd	6539.3	2.06	0.15	24	-3245.4	14,951	0.060	402	0.164
5	hr	logsize + Obs + Phase * Herd	logsize + Obs + Phase * Herd	6541	3.73	0.07	25	-3245.2	14,491	0.058	410	0.165
6	hn	logsize	logsize + Obs + Phase + Herd	6546.7	9.51	0.00	13	-3260.3	14,823	0.060	347	0.165
7	hr	logsize	logsize + Obs + Phase + Herd	6547	9.73	0.00	14	-3259.4	14,735	0.067	346	0.165
8	hn	logsize + Phase * Herd	logsize + Obs + Phase + Herd	6548.8	11.62	0.00	16	-3258.3	14,710	0.061	398	0.160
9	hn	logsize * Herd	logsize + Obs + Phase * Herd	6551.7	14.46	0.00	16	-3259.7	15,015	0.060	345	0.163
10	hn	logsize	logsize + Obs + Phase	6551.9	14.72	0.00	12	-3263.9	14,831	0.060	342	0.165
11	hn	logsize	logsize + Obs	6574.3	37.11	0.00	11	-3276.1	14,893	0.060	346	0.164
12	hn	logsize	logsize	6633.6	96.37	0.00	4	-3312.8	14,864	0.060	341	0.165
13	hr	logsize	logsize	6633.8	96.58	0.00	5	-3311.9	14,776	0.067	341	0.166
14	hn	logsize	Constant	6688.3	151.07	0.00	3	-3341.1	14,939	0.060	342	0.166
15	hn	size	Constant	6688.3	151.10	0.00	3	-3341.2	14,834	0.059	341	0.165
16	hr	logsize	Constant	6688.5	151.28	0.00	4	-3340.2	14,849	0.067	341	0.167
17	hr	size	Constant	6689	151.79	0.00	4	-3340.5	14,783	0.066	340	0.166
18	hn	Constant	Constant	6694	156.76	0.00	2	-3345.0	15,633	0.061	348	0.171
19	hr	Constant	Constant	6696	158.83	0.00	3	-3345.0	15,523	0.066	346	0.173
20	un	Constant	Constant	6711.4	174.13	0.00	1	-3354.7	14,040	0.056	313	0.169

Plots of predicted detection probabilities (Figure 13) suggest reasonable fit of the distance sampling and mark-recapture portions for Model 1 (Table 8). For the mark-recapture portion, the conditional detection probabilities (based on the data) show no decline in detection probability with distance which contrasts with detection probabilities from the distance curve. As discussed later, this is likely due to heterogenous double observer detection probabilities obscuring decline in detection probability with distance. There were no degrees of freedom available for chi-square goodness of fit tests. However, the overall chi-square value for the distance sampling portion was less than 1 (0.759), suggesting adequate fit. The chi-square value for the mark-recapture portion was 10 suggesting lesser fit with the main deviance occurring at the outer survey bin. It is suspected that heterogeneity of detection probabilities is causing lack of fit of the MR model.

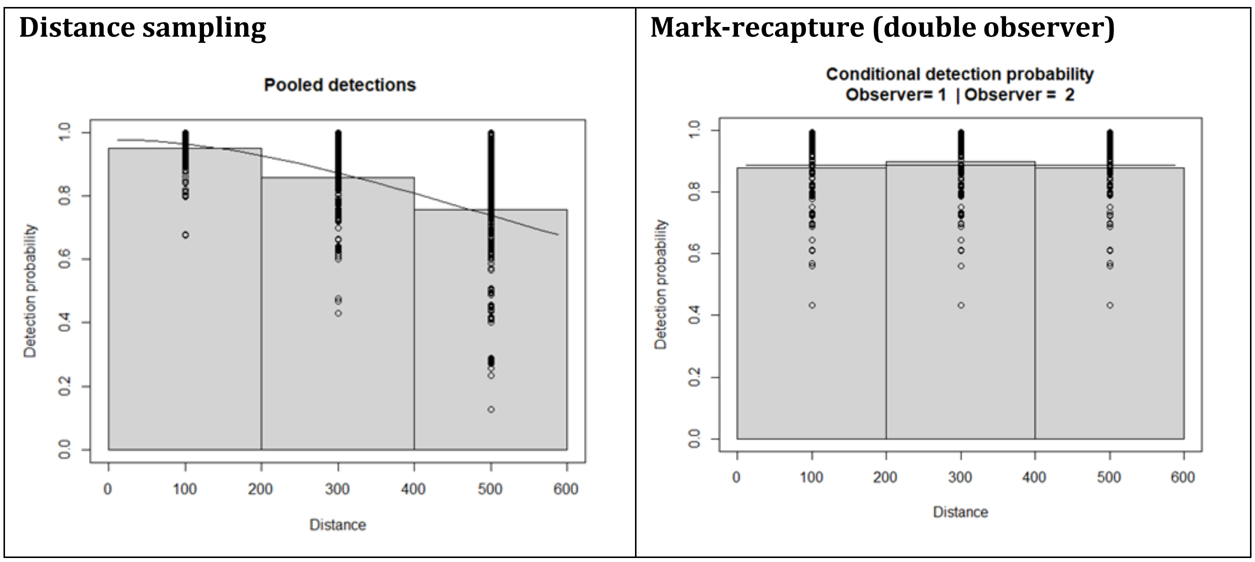

Figure 13: Plots of predicted detection probabilities for Model 1 (Table 8) for the distance sampling (left) and Mark-recapture portions of the MRDS model (right). Points are estimated detection probabilities for each of the observations in the data set.

Comparison of abundance estimates

Visual abundance estimates from the most supported model for each method were relatively similar for the Bluenose-East herd and less similar for the Bathurst herd, though all abundance estimates for each herd had overlapping confidence intervals (Table 9 and Figure 14). The probable reason for the greater variation among Bathurst herd estimates was lower overall precision of the Bathurst estimates creating more variability in estimates regardless of method. A few general patterns are evident. First, addition of a third observer on each side of the aircraft (data recorder observations) for strip transects increased the number of groups observed with a large increase in the Bathurst herd estimate for both strip transect and distance sampling approaches. The increase is less for the Bluenose-East herd which is likely due to relatively higher sightability during the Bluenose-East survey. Second, strip transect estimates with data recorder observations were higher than the double observer mark-recapture observations suggesting that inclusion of the 3rd data recorder observation was as efficient if not more efficient than modelling detection variation using double observer methods. Third, MRDS and distance sampling estimates were higher than strip transect or mark-recapture estimates suggesting distance sampling was accounting for heterogeneity of detection probabilities. Finally, distance sampling estimates, with data recorder observations were the highest which suggested that inclusion of the data recorder observations was able to improve the assumption of sightability=1 on the line with distance sampling accounting for variation in detection within the 600m survey strip.

Table 5: Total visual estimates for each herd. Included are whether data recorder observations were included (DR obs), the number of caribou groups, the number of caribou in the analysis data set (n), abundance estimates (N), confidence limits, coefficient of variation (CV) and the % difference of estimates from strip transects.

Method	DR obs	groups	n	N	SE	95% Conf. Limit		CV	% diff from ST
Bathurst									
Strip transect (400m strip)	N	57	115	319	48.1	236	432	0.151	0.0%
Strip transect (400m strip)	Y	69	131	367	48.9	281	480	0.133	14.9%
Mark-recapture (400m strip)	N	57	115	330	61.9	227	481	0.187	3.4%
Mark-recapture (600m strip)	N	74	163	317	50.9	230	437	0.161	-0.7%
Distance sampling	Y	88	181	458	71.8	334	627	0.157	43.3%
MRDS	N	74	163	402	62.4	295	548	0.155	25.9%
Bluenose-East									
Strip transect (400m strip)	N	806	2,452	14,238	996.0	12,386	16,367	0.070	0.0%
Strip transect (400m strip)	Y	832	2,497	14,525	1017.9	12,631	16,702	0.070	2.0%
Mark-recapture (400m strip)	N	806	2,452	14,347	1095.6	12,323	16,703	0.076	0.8%
Mark-recapture (600m strip)	N	1,166	3,566	13,927	775.4	12,438	15,594	0.056	-2.2%
Distance sampling	Y	1,211	3,641	15,260	908.4	13,536	17,204	0.060	6.5%
MRDS	N	1,166	3,566	14,950	900.5	13,249	16,871	0.060	5.0%

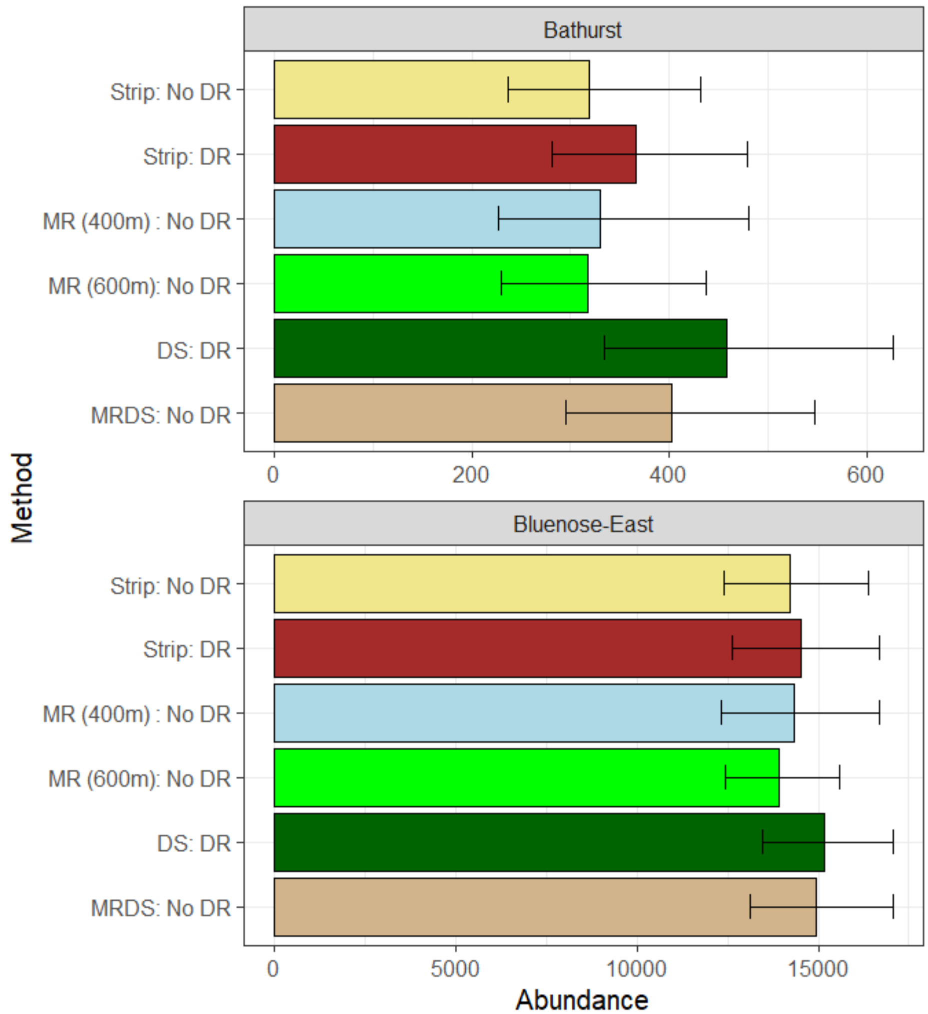

Figure 14: Comparison of estimates using different data set formulations and estimation methods as listed in Table 9.

Precision of estimates was high for the Bluenose East (CV=6-7%) and reasonable for the Bathurst herd (CV=13-19%). For MRDS it was possible to dichotomize sources of variance in the analysis. For the Bluenose-East herd, 81% and 19% of variation was due to encounter rate (variation between transect) and detection estimation. For the Bathurst herd, 95% and 5% of the variation was due to encounter rate and detection estimation. These summaries suggest that the majority of variation in the estimates was due to uneven densities as opposed to estimation of detection. Inspection of relative densities from the 600m strip transects shows that there was a fair amount of variation in the Bathurst herd with very low densities (Figure 15). In contrast, densities were more even for the Bluenose-East herd. Variation in density will be proportional to encounter rate variation (caribou detected per km), and in this context we can conclude that the very low and spotty distribution of the Bathurst herd was the primary reason for reduced precision.

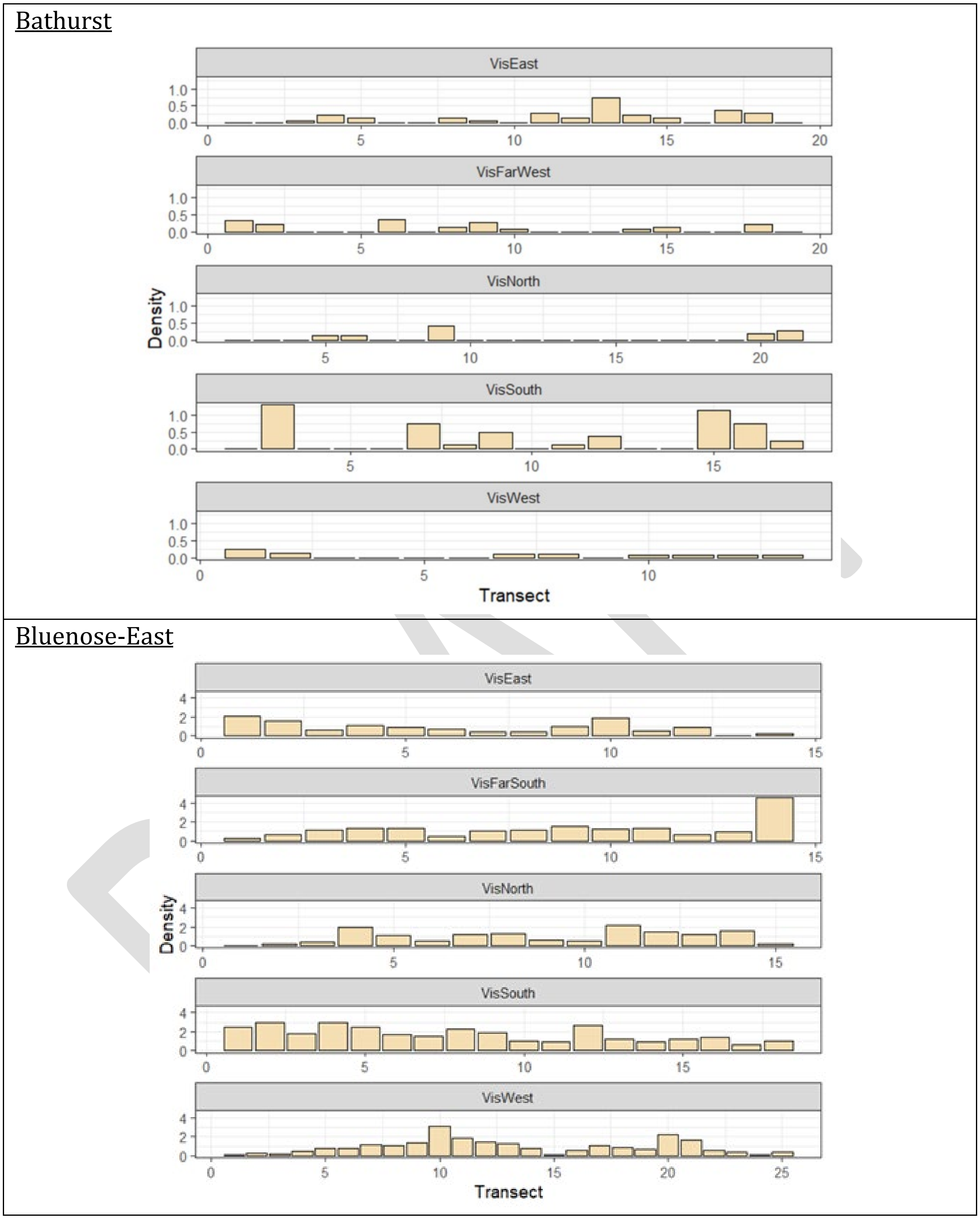

Figure 15: Relative densities on visual transects for Bluenose-East and Bathurst visual surveys based on the 600m strip. Densities are based on strip-transect methods (no correction for sightability) and therefore should be interpreted relatively.

Overall evaluation of methods

Table 10 provides a summary of hypothesized and actual results from analyses. The main conclusions are as follows: 1) Strip transect surveys with no correction for sightability are likely negatively biased, however, the degree of bias can be lessened by having 3 observers (2 observers + data recorder observer) per side of plane. 2) Double observer (mark-recapture) methods likely overestimate sightability (i.e. estimate that observers detect more of the caribou present in the survey area than they actually do). This result is based on comparison of counts with 3 observers to estimated counts using mark-recapture methods (using 2 observers) which result in a negatively biased estimate. 3) Distance sampling with 3 observers (including data recorder observations) provides the most robust estimate. Having 3 observers likely meets the assumption of sightability=1 on the survey line and distance sampling is most efficient at modelling the decay in detection relative to distance from the survey plane.

Table 6: Summary potential results from comparison of estimators with implications with survey results.

Hypothesized result	Implications/biases	Actual Result
Distance sampling/2x observer methods estimates of sightability are less than 1	Strip transects with 2 observers are negatively biased	Strip transect estimates < all other estimators
Distance sampling detects decline in detection probability that is not detected when distance is used as a covariate in double observer models	Increasing heterogeneity with distance from plane prevents the detection of decline in detection probabilities using double observer methods suggesting double observer methods are negatively biased	Decline in detection with distance not detected by double observer methods (Figure 15)
Strip transect estimates with data recorder observations included are higher than strip transect estimates using 2 observers	3 observers detect more caribou than 2 observers	Strip transects with no data recorder (2 observers) were less than strip transects with data recorder (3 observers)
Strip transect estimates with data recorder included are higher than estimates from double observer models	Double observer models are negatively biased	Strip transect estimates with data recorder included were higher than double observer strip transect estimates
Distance sampling estimates with data recorder observations are higher than strip transect and/or double observer models using 2 observers	Distance sampling estimates using 3 observers provides a less biased estimate due to robustness to unequal sighting probabilities. Inclusion of data recorder observations helps meet the assumption of sightability=1 on survey line	Distance sampling estimates with data recorders were the highest of all estimates.
MRDS (Mark-recapture/distance sampling) estimates are lower than distance sampling estimates with data recorder	Including data recorder observations (as a 3rd observer) provides a more efficient means to mitigate imperfect observer detection than double observer (mark-recapture) methods.	MRDS (Mark-recapture distance sampling) estimates were lower than distance sampling estimates (with data recorder observations)

DISCUSSION

The results of the 2025 visual surveys provided additional insight into factors affecting the robustness of visual survey estimates. Survey conditions in 2025 were ideal with minimal to no snow cover and good visibility. Using the distance sampling method (including observations from the data recorder) provided the most robust and highest estimates of all methods for each herd. It is likely that future calving ground surveys may have more challenging survey conditions than experienced in 2025, which could lead to greater differences between methods used to estimate abundance from visual surveys, and therefore continued testing of distance sampling and other approaches will provide further tools to ensure robustness of estimates.

A key take away from this comparison is that all methods provided similar abundance estimates of caribou in the visual strata. For each herd the 95% confidence intervals for each method's abundance estimate overlapped and the variance around the estimates was similar, particularly for the Bluenose-East herd, where the coefficient of variation (CV) ranged from 0.06 to 0.07 (with a range in CV of 0.13 to 0.19 for the Bathurst herd). This provides support for comparing calving ground visual survey results using newer methods to visual survey results from prior surveys that used alternative methods compared here.

Differences between estimates were not large for the Bluenose-East herd: the largest estimate (17,204) was 10% higher than the lowest estimate (15,594), though the estimates were not significantly different. In contrast, differences between estimates were greater for the Bathurst herd, where the largest estimate (627) was 45% higher than the lowest estimate (432), though the estimates were not significantly different. When densities are very low with patchy caribou distribution, as occurred in the Bathurst visual strata, it is challenging to provide precise estimates regardless of the method used, and the numerical difference between the estimates is small (195 caribou). Overall sightability was lower for the Bathurst herd. Differences in sightability were potentially due to different survey conditions and lower densities in the Bathurst strata that may have affected sightability (Campbell et al. 2014).

Results suggest that use of data recorder observations as a third observer is as effective as double observer estimation of missed groups. In this analysis, in all cases the number of additional groups counted by data recorders exceeded the estimated number of groups using double observer methods. This resulted in estimates from strip transects (that utilized data recorder observations) being equivalent to slightly higher than the equivalent 400m double observer (mark-recapture) estimates. The data recorder is only able to observe the survey area when they are not entering data from primary and secondary observers, so they are not a fully dedicated 3rd observer. Under conditions of higher caribou densities, the data recorder will necessarily spend more time recording and may be a less effective 3rd observer. However, the tablet interface makes data entry much faster than taking a GPS waypoint and recording it along with each caribou observation on a paper datasheet, which allows the data recorder to spend more time focusing on observing the survey area.

Comparison of double observer 400m strip transect estimates, the method used for visual survey estimates from 2011 to 2023, suggests that the double observer approach alone does not adequately estimate sightability. It provides estimates that are lower than strip transect estimates using 3 observers, as noted, and lower than distance sampling (DS) and mark-recapture distance sampling (MRDS) estimates (Table 9). The likely reason for this is that heterogeneity (unequal detection probabilities) of caribou groups within the survey strip. This creates a positive bias in estimated detection probabilities from double observer surveys (over-estimating the detection rate) leading to a negative bias in estimates (under-estimating the true abundance). One example of the effect of heterogeneity is seen in the comparison of detection probability estimates with distance (Figure 13). Distance sampling estimates a decline in detection probabilities as seen in the detection frequency histograms (i.e. the numbers of caribou observations declined with distance from plane). In contrast, the double observer method does not detect this decline. The reason for this is that observers both detect the easy to see (larger) groups at further distances and miss the hard to see groups (e.g., single caribou). The detection probability (based on ratios of detections by both or single observers) therefore stays constant (since they are both seeing the same groups) even when it should be declining. The mean detection probability for the survey strip is then dominated by groups that are easier to observe closer to the plane and is biased high.

The issue of heterogeneity is especially relevant for the dependent double observer method used in many surveys that allows observers to communicate findings but also places restrictive assumptions on detection probability variation such as each observer in a pair having equal detection probability (Buckland et al. 2004). The general mechanism for bias in double observer detection probability with distance is that as distance from the plane increases and detection probability decreases observers become more likely to just observe the more obvious caribou groups and both miss the less obvious groups leading to non-independence of observer detection probabilities (Buckland et al. 2010). This non-independence basically creates heterogeneity variation leading to a positive bias in detection probability (only the easy to observe groups are detected usually by both observers) and a resulting negative bias in abundance.

Another issue with double observer methods is the restrictive assumptions of the removal method and lack of flexibility in modelling individual differences in observation ability. Even with these limitations, results suggested that each observer pair had a unique set of mean detection probabilities. Even with its current limitations, the dependent double observer method provides a good framework for observers to work together as well as a means to evaluate observer pair efficiency.

An additional benefit to the dependent double observer method, unrelated to the statistical analysis of the resulting data, is that it provides accountability where observers are motivated to observe carefully, and the ongoing conversations can help observers to maintain focus during long stretches of on-transect survey time. It also facilitates training as new observers can learn from more experienced observers. Each of these factors likely improves detection of caribou during visual surveys.

Distance sampling estimates, with data recorder observations included, were the highest of all estimates. One unique attribute of distance sampling is that it is relatively robust to heterogeneity variation as long as sightability on the survey line is 1 or is estimated using double observer methods due to pooling robustness of the distance sampling approach (Buckland et al. 1993, Rexstad et al. 2023, Laake and Collier 2024). Rexstad et al (2023) comments that robustness of distance sampling estimates can be affected by strata-specific detection function variation. We tested this and found that estimates with strata included were minimally changed compared to other covariate models. Use of data recorder observations likely increases the likelihood that the assumption of detection probability being 1 at 0 distance is met. Use of multiple observers is permitted in distance sampling and in this context, it could be argued that a sole distance sampling approach could be used due to the challenges of dependent double observer methods. However, the main risk with a distance sampling alone approach would be a case where survey conditions are poor leading to reduced sightability on the survey line that would potentially be accounted for by the double observer method. For this scenario, the mark-recapture distance sampling (MRDS) would likely be more robust.

One challenge to the application of distance sampling methods was the high sightability of groups which extended the shoulder area of constant sightability to the outer range of the survey data (600m) for larger group sizes and with some observer pairs (Figure 12). The 600m survey width was based on winter surveys of caribou in Nunavut where sightability was likely lower. To effectively model a shoulder area the shoulder area should be less than the total width of survey data allowing explicit modelling of the shoulder area using the hazard rate model. In general, the hazard rate model is preferable to the half normal model due to its flexibility and greater robustness. In the case of this analysis, estimates did not vary greatly between half-normal and hazard rate detection functions. However, other analyses and simulations suggest that half normal detection functions may exhibit a positive bias when a significant shoulder area is present (Boulanger and Campbell 2025).

Ideal survey conditions reduced the decline in detection with distance in the 3 survey bins compared to the distance sampling data sets used in caribou surveys in Nunavut. Even with this limitation, it was still possible to fit distance sampling models to data sets, however use of the hazard rate model was limited due to lack of a defined shoulder area. We provide considerations on how modifying the distance bins, or adding a fourth bin, may allow for better application of the distance sampling methodology. Further improvements to distance sampling could be:

1. Moving the inner strip boundary away from the tire of the airplane or adding a narrower bin closest to the aircraft. Because observers typically focus farther out than the area immediately closest to the aircraft (e.g. closer to 200 m than to 0 m) this may better define the area where detection is greatest.
2. Adding an additional bin and modifying bins to make the observation areas for each bin a bit more equal. Bins of equal ground distance are progressively narrower from the observer's perspective with distance from the plane (see Figure 2). For example, one modification might be strips of 0-150, 150-350, 350-550, and 550-800, where bin width increases with distance from the plane. This would make the viewing areas more equal.
3. Adding an additional bin with a slight expansion of overall width would allow a better estimate of shoulder area (where sightability is constant) when conditions are ideal as they were in 2025 as well as if distance sampling methods are to be used in areas of higher density (e.g., with larger group sizes that would likely increase the shoulder area of the detection function). In this analysis fitting of the hazard rate model (which estimates a survey strip of constant sightability) was challenged by lack of decline in the third survey bin. Using a fourth bin would improve the overall robustness of estimates.
4. In studies where observers can stop moving to measure distance (e.g. walking or helicopter-based transects), observers have used a rangefinder and angle measures to measure distances rather than binning (e.g. Montalvo et al. 2021, Wingard et al. 2011). This approach is difficult to apply from a moving fixed-wing aircraft, but would give a much more robust data set than bins. Advances in technology might occur to make this approach more viable in the future.

A simulation study like that of Boulanger and Campbell (2025) detailed in Appendix B could be applied to explore strategies to improve the distance sampling approach. For this study detection function parameters estimated from this study would be used to simulate a range of detection functions with varying shoulder areas simulating low to high sightability. The 3 and 4 distance bin designs would then be applied to further assess the robustness of each approach.

LITERATURE CITED

Borchers, D. L., W. Zucchini, and R. M. Fewster. 1998. Mark-recapture models for line-transect surveys. Biometrics **54**:1207–1220.

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

Boulanger, J., J. Adamczewski, J. Williams, D. Cluff, K. Clark, S. Goodman, K.-S. Chan, and R. Abernathy. 2022. Estimates of Breeding Females & Adult Herd Size and Analyses of Demographics for the Bluenose-East Herd of Barren-Ground Caribou: 2021 Calving Ground Photographic Survey Environment and Natural Resources, Yellowknife, NWT.

Boulanger, J., and M. Campbell. 2025. Use of a reduced distance sampling and double observer methodology to improve robustness of calving ground survey estimates: Appendix in Campbell, M., J. Boulanger, J. Ringrose, J. Danahy, R. Kite, and A. Roberto-Charron. 2025. Abundance and Trends of the Beverly Mainland Migratory Subpopulation of Barren-Ground Caribou (Rangifer tarandus groenlandicus) June 2011 – June 2023. Department of Environment, Government of Nunavut, Arviat, Nunavut. Government of Nunavut, Arviat, Nunavut.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993. Distance Sampling. Estimating Abundance of Biological Populations. Chapman & Hall, London.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2004. Advanced Distance Sampling - Estimating abundance of biological populations. Oxford Press.

Buckland, S. T., J. Laake, and D. L. Borchers. 2010. Double-observer line transect methods : levels of independence Biometrics **66**:169–177.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: A practical information theoretic approach. Springer, New York, New York, USA.

Campbell, M., J. Boulanger, and D. Lee. 2020. Long-term trends in abundance and distribution of the Southampton Island caribou herd: 1978 – 2019 Government of Nunavut, Department of Environment, Arviat, NU.

Campbell, M., J. Boulanger, D. Lee, M. Dumond, and J. McPhearson. 2012. Calving Ground Abundance Estimates of the Beverly and Ahiak Subpopulations of Barren-Ground Caribou (Rangifer tarandus groenlandicus) – June 2011, Technical Summary. Department of Environment, Government of Nunavut.

Campbell, M., J. Boulanger, J. Ringrose, J. Danahy, R. Kite, and A. Roberto-Charron. 2025. Abundance and Trends of the Beverly Mainland Migratory Subpopulation of Barren-Ground Caribou (Rangifer tarandus groenlandicus) June 2011 – June 2023. Department of Environment, Government of Nunavut, Arviat, Nunavut.

Campbell, M., J. Goorts, D. S. Lee, J. Boulanger, and T. Pretzlaw. 2015. Aerial Abundance Estimates, Seasonal Range Use, and Demographic affiliations of the Barren-Ground Caribou (Rangifer tarandus groenlandicus) on Baffin Island – March 2014. Department of Environment, Nunavut Wildlife Research Division, , Arviat, Nunavut.

Campbell, M., D. Lee, and J. Boulanger. 2019. Calving Ground Abundance Estimates of the Beverly Subpopulations of Barren-Ground Caribou (Rangifer tarandus groenlandicus) – June 2018. Technical Summary. Government of Nunavut.

Campbell., M., J. Ringrose, J. Boulanger, A. Roberto-Charron, K. Metheun, C. Mutch, T. Davison, and C. Wray. 2021. An Aerial Abundance Estimate of the Dolphin and Union Caribou (Rangifer tarandus groenlandicus x pearyi) herd, Kitikmeot Region, Nunavut – Fall 2020, Government of Nunavut, Technical Report Series No 01-2021. 146 pp. .

Cook, D. R., and J. O. Jacobsen. 1979. A design for estimating visibility bias in aerial surveys. Biometrics **35**:735–742.

Elphick, C. S. 2008. How you count counts: the importance of methods research in applied ecology. Journal of Applied Ecology **45**:1313–1320.

Grolemund, G., and H. Wickham. 2011. Dates and Times Made Easy with lubridate. Journal of Statistical Software **40**:1 – 25.

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

Jolly, G. M. 1969. Sampling methods for aerial censuses of wildlife populations. East African Agricultural and Forestry Journal **34**:46–49.

Kelly, A., J. Boulanger, J. Williams, R. Abernethy, J. Hodson, C. Modeste-Burgin, S. Goodman, E. McHugh, and K. Clark. In Prep. June 2025 Calving Ground Survey and 2025 Population Modelling of the Bluenose-East Barren-ground Caribou Herd. Draft Manuscript Report, Environment and Climate Change, Government of the Northwest Territories, Yellowknife, NT.

Kelly, A., J. Boulanger, J. Williams, R. Abernethy, J. Hodson, C. Modeste-Burgin, S. Goodman, E. McHugh, and K. Clark. In prep. June 2025 Calving Ground Survey and 2025 Population Modelling of the Bathurst Barren-ground Caribou Herd. DRAFT Manuscript Report, Environment and Climate Change, Government of the Northwest Territories, Yellowknife, NT.

Krebs, C. J. 1998. Ecological Methodology (Second edition). Benjamin Cummins, Menlo Park, California.

Laake, J., D. L. Borchers, L. Thomas, D. Miller, and J. Bishop. 2012. Mark-recapture distance sampling (MRDS) 2.1.0. R statistical package program.

Laake, J., M. J. Dawson, and J. Hone. 2008a. Visibility bias in aerial survey: mark-recapture, line-transect or both? Wildlife Research **35**:299–309.

Laake, J., R. J. Guenzel, J. L. Bengtson, P. Boveng, M. Cameron, and M. B. Hanson. 2008b. Coping with variation in aerial survey protocol for line-transect sampling. Wildlife Research **35**:289–298.

Laake, J. L., and B. A. Collier. 2024. Understanding implications of detection heterogeneity in wildlife abundance estimation. The Journal of Wildlife Management **88**:e22516.

Mazerolle, M. J. 2016. AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). . R package version 2.1-0.: https://cran.r–project.org/package=AICcmodavg.

Montalvo, A., L.A. Brennan, M.L. Morrison, E.D. Grahmann, and A.N. Tri. 2022. Distance sampling survey effort to improve density estimates for northern bobwhites. Wildlife Society Bulletin (East Foundation Manuscript 065).

Norton-Griffiths, M. 1978. Counting Animals. 2nd edition. African Wildlife Leadership Foundation, Nairobi.

Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal **10**:439–446.

QGIS Foundation. 2020. QGIS Geographic Information System. QGIS Association. http://www.qgis.org.

R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rexstad, E., S. Buckland, L. Marshall, and D. Borchers. 2023. Pooling robustness in distance sampling: Avoiding bias when there is unmodelled heterogeneity. Ecology and Evolution **13**:e9684.

Thomas, D. C., S. T. Buckland, E. A. Rexstad, J. Laake, S. Strindberg, S. L. Hedley, J. R. B. Bishop, T. A. Marques, and K. P. Burnham. 2009. Distance software: Design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology **47:** 5–14.

Wickham, H. 2009. ggplot2: Elegant graphics for data analysis. Springer, New York.

Wickham, H. 2011. The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software **40**:1 – 29.

Wingard, G.J., R.B. Harris, S. Amgalanbaatar and R.P. Reading. 2011. Estimating abundance of mountain ungulates incorporating imperfect detection: argali Ovis ammon in the Gobi Desert, Mongolia. Wildlife Biology, 17: 93-101.

APPENDICES

Appendix A. Double observer (MR) model selection results using 400m strip width

Table A-1: Model selection results for double observer mark recapture analysis using 400m strip width. Detection functions (DSDF) are given (hn=half-normal, hr=hazard rate, un=uniform). Model selection criteria are AIC_c, the difference in AIC_c values between the *i*th and most supported model 1 (ΔAIC_c), Akaike weights (w_i), and number of parameters (K), and log-likelihood. Total visual estimates are also given for each herd.

No	MR model	Model selection					BNE N		Bathurst N	
		AIC_c	ΔAIC_c	wi	K	LL	N	CV	N	CV
1	logsize + Obs + Phase + Herd	1047.64	0.00	0.47	11	-512.7	14,347	0.075	330	0.204
2	logsize + Obs + Phase + Herd + cloud	1048.24	0.60	0.35	12	-512.0	14,345	0.075	331	0.204
3	logsize + Obs + Phase * Herd	1049.68	2.04	0.17	12	-512.7	14,347	0.075	330	0.205
4	logsize + Obs + Phase	1053.79	6.16	0.02	10	-516.8	14,357	0.075	324	0.205
5	logsize + Obs	1067.45	19.81	0.00	9	-524.7	14,404	0.075	327	0.204
6	logsize + Phase * Herd	1084.64	37.00	0.00	5	-537.3	14,336	0.075	329	0.205
7	logsize + Herd	1084.74	37.10	0.00	3	-539.4	14,356	0.075	331	0.204
8	logsize + cloud	1090.17	42.53	0.00	3	-542.1	14,362	0.075	325	0.205
9	logsize + Phase	1093.39	45.75	0.00	3	-543.7	14,348	0.075	322	0.206
10	logsize + snow	1097.13	49.49	0.00	3	-545.6	14,378	0.075	323	0.206
11	logsize	1099.23	51.59	0.00	2	-547.6	14,383	0.075	323	0.206
12	logsize + distance	1100.36	52.72	0.00	3	-547.2	14,383	0.075	323	0.206
13	logsize * distance	1102.12	54.48	0.00	4	-547.0	14,383	0.075	323	0.206
14	size	1106.86	59.22	0.00	2	-551.4	14,390	0.075	323	0.206
15	distance* Phase	1130.77	83.14	0.00	4	-561.4	14,396	0.075	322	0.207

Appendix B. Use of a reduced distance sampling and double observer methodology to improve robustness of calving ground survey estimates

This appendix is an excerpt from an appendix in the 2023 Beverly calving ground survey report (Campbell et al. 2025). It should be cited as the appendix of the 2023 Beverly calving ground survey report:

Boulanger, J., and M. Campbell. 2025. Use of a reduced distance sampling and double observer methodology to improve robustness of calving ground survey estimates: Appendix in Campbell, M., J. Boulanger, J. Ringrose, J. Danahy, R. Kite, and A. Roberto-Charron. 2025. Abundance and Trends of the Beverly Mainland Migratory Subpopulation of Barren-Ground Caribou (Rangifer tarandus groenlandicus) June 2011 – June 2023. Department of Environment, Government of Nunavut, Arviat, Nunavut. Government of Nunavut, Arviat, Nunavut.

Introduction

This analysis explores the use of a reduced distance sampling combined with double observer methodology to improve the estimation of sightability and subsequent abundance estimates from calving ground surveys. Some of the recent calving ground surveys have been challenged by poor survey conditions (Boulanger et al. 2019, Campbell et al. 2019, Boulanger et al. 2022) which created potential challenges with comparison of surveys such as the 2023 calving ground surveys that had ideal survey conditions. In this paper, we assess potential negative bias in double observer methods and propose a reduce distance sampling / double observer methodology that should improve the robustness of calving ground survey abundance estimates.

The current calving ground survey methodology uses a 400 meter strip transect combined with a double observer mark-recapture modelling effort method to estimate sightability and abundance. The two observers in this case represent an initial detection of caribou (the mark) and a potential re-detection of the caribou group (the recapture) hence double observer methods are analyzed using a mark-recapture analysis approach. This approach was adopted in 2011 to accommodate higher densities of caribou observed on calving grounds which prevented the use of full distance sampling methods, which require greater observer training and the ability to observe caribou across a relatively wide survey strip (1500m) with classification of groups into 5 distance bins (Campbell et al. 2012).

Distance sampling methods have been applied in Nunavut for caribou surveys (Campbell et al. 2015, Campbell et al. 2020, Campbell. et al. 2021) that occurred in spring or fall periods and therefore had lower densities of caribou making it possible to apply the distance sampling method. Results from all of these surveys estimate that detection probability declines with distance within the 400m survey strip. Comparison of distance sampling estimates with double observer 400m strip transect estimates suggest that that the double observer approach does not adequately estimate sightability with estimates being 5-11% lower than double observer distance sampling estimates (Campbell et al. 2015, Campbell et al. 2020). The likely reason for this is that heterogeneity (unequal detection probabilities) of caribou groups within the survey strip creates a positive bias in estimated detection probabilities from double observer surveys leading to a negative bias in estimates. This issue is especially relevant for the dependent double observer method used in many surveys that allows observers to communicate finding but also places restrictive assumptions on detection probability variation such as each observer in a pair having equal detection probability (Buckland et al. 2004). The general mechanism for bias in double observer detection probability with distance is that as distance from the plane increases and detection probability decreases observers become more likely to just observe the more obvious caribou groups and both miss the less obvious groups leading to non-independence of observer detection probabilities (Buckland et al. 2010). This non-independence basically creates heterogeneity variation leading to a positive bias in detection probability (only the easy to observe groups are detected usually by both observers) and a resulting negative bias in abundance. The amount of bias likely varies with overall sightability. If survey conditions are good (low snow years such as 2023) the bias is likely minimal, however, in years of poor sightability such as in some previous calving ground surveys then higher levels of bias are possible.

One unique attribute of distance sampling is that it is relatively robust to heterogeneity variation as long as sightability on the survey line is 1 or is estimated using double observer methods due to pooling robustness of the distance sampling approach (Buckland et al. 1993, Rexstad et al. 2023, Laake and Collier 2024). For this reason, the general approach used in distance sampling surveys is to focus mark-recapture/double observer methods on estimating detection on the survey line with distance sampling used to model and estimate the decline in detection probability with distance (Laake et al. 2008a, Laake et al. 2008b). We suggest that it is advantageous to collect and include distance sampling data during aerial surveys as it provides a way to confront likely heterogeneity bias that challenges double observer methods.

Methods

The general approach we propose is the use of distance sampling with a reduced number of bins and overall strip distance surveyed. This approach should be less overwhelming than a full distance sampling approach but still may allow a better estimate of sightability and more robust estimates. To assess this approach, we used empirical comparisons and a simulation study.

For the empirical comparisons, we subset distance sampling data sets from the Baffin 2024 (Campbell et al in prep) and South Hampton Island surveys (Campbell et al. 2020) to create reduced data sets and reran the analyses with the reduced data sets to produce estimates. These estimates were then compared with full data set estimates as well as double observer estimates that do not include distance sampling using the *mrds* R (Laake et al. 2012) package. Detailed analysis results are given in survey reports from each project and are not detailed in this report.

We also conducted a brief simulation study to assess whether unbiased estimates could be derived using reduced bins. For these analyses, we simulated the detection function from the Baffin 2024 and then analyzed the simulated data using the 2 and 3 bin designs. Simulations were focused on the distance sampling component of the analysis and assumed detection on the survey line was 1 with no variation in detection functions. We considered hazard rate and half normal models for estimation and assessed whether there was adequate power to select the appropriate detection function model with each design. Simulations were evaluated in terms of overall estimator bias, confidence interval coverage (proportion of simulations where confidence limits from the estimate overlapped the true value), and precision. Sample sizes of observations were generated to roughly approximate sample sizes from the Baffin Island survey. Five hundred simulations were conducted for each design. Simulations were conducted using the *dsims* R package (Marshall 2022).

Results

<u>Southampton Island</u>

The 2017 Southampton Island caribou survey employed a double observer distance sampling approach with 5 survey bins as shown in Figure 1. The data set consisted of 962 observations used to fit the detection function. A hazard rate detection function was selected which allows a “shoulder” of constant detection which in this case extended for the 0-200m survey bin before declining in subsequent bins. Detection also varied by observer order, snow cover and the log of group size for the distance detection function. For the double observer model detection varied by observer order and snow cover. Detection on the survey line was estimated to be 0.94 (CV=0.013).

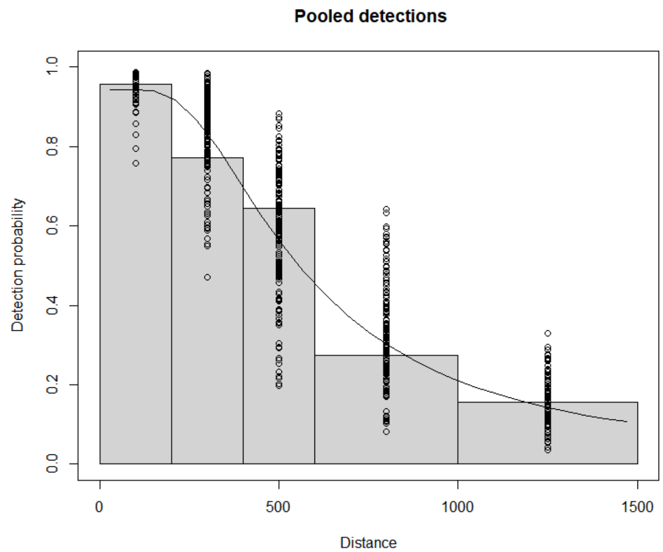

Figure 1. Distance sampling detection function for the full MRDS analysis of Southampton Island.

For the mark-recapture portion of the analysis it is possible to plot double observer sightability estimates for each distance bin to assess if double observer detection probabilities were reduced with distance (Figure 2). It can be seen that there is a slight decline up to the 600m bin but minimal change overall. This plot illustrates the effect of heterogeneity variation. With no heterogeneity variation this plot would be similar to Figure 1 with a decline in detection with distance. Instead, observers both only observe the more sightable groups with increasing distance leading to an overestimate of detection probabilities and a negative bias in double observer-only estimates.

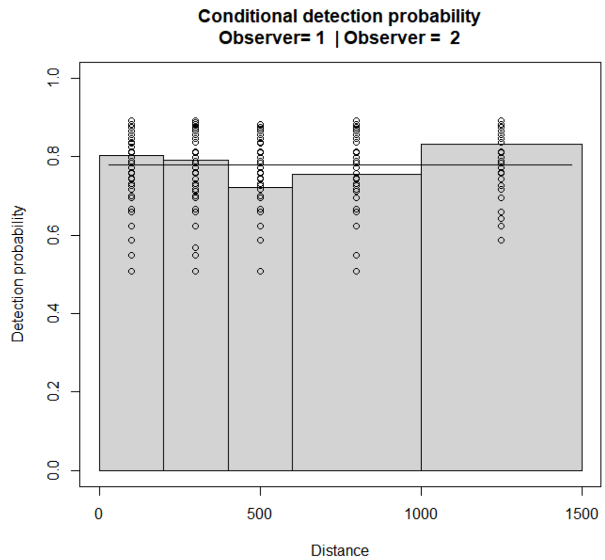

Figure 2. Distance sampling detection function for the full MRDS analysis of Southampton Island. Each bar is the conditional double observer detection probability based on the ratio of caribou groups observed by the secondary observer to overall groups detected.

Figure 3 shows estimates based on subsetting the full data set. The first bar represents abundance from the full data set for comparison with the 1, 2, and 3 bin estimates. Hazard rate (the detection function that allows a larger shoulder area) and half normal (a simpler detection function that restricts the shoulder distance) were considered along with mark-recapture/double observer models with 1 bin (as used in calving ground surveys) and 2 bins (with distance as a covariate to estimate decline in detection across the 2 bins). In general, reasonably similar estimates occur when 2 and 3 bins were used with the best agreement for the hazard rate (HR) model with 3 bins. The hazard rate model did not converge when only 2 bins were used, however, in this case the half-normal still provided an estimate of reasonable agreement. The double observer/mark-recapture only estimates were 5% lower than the full MRDS estimates for this survey.

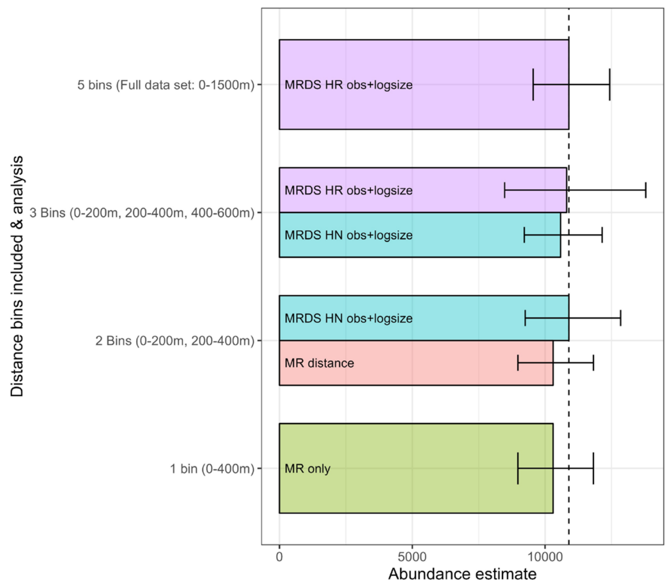

Figure 3. Southampton Island 2017 analysis. The same double observer/mark-recapture model (obs+snow) was used for all analyses shown. The dashed line indicates the estimate using the full 5 bin (0-1500m) data set.

<u>Baffin 2024 fixed wing survey</u>

The fixed wing only data set (a portion of the survey was conducted via helicopter which was less relevant) which consisted of 827 observations was used for this analysis. The detection function for the South Baffin Island 2024 survey was relatively similar to the Southampton Island survey with a hazard rate detection function estimating a shoulder area of approximately 200m before declining at successive distances (Figure 4). Sightability on the survey line was estimated as 0.96 (CV=0.029). Plotting double observer estimates of sightability with distance revealed a decline in estimated detection probabilities, however, detection probability estimates at further survey bins were still much higher than distance sampling estimates.

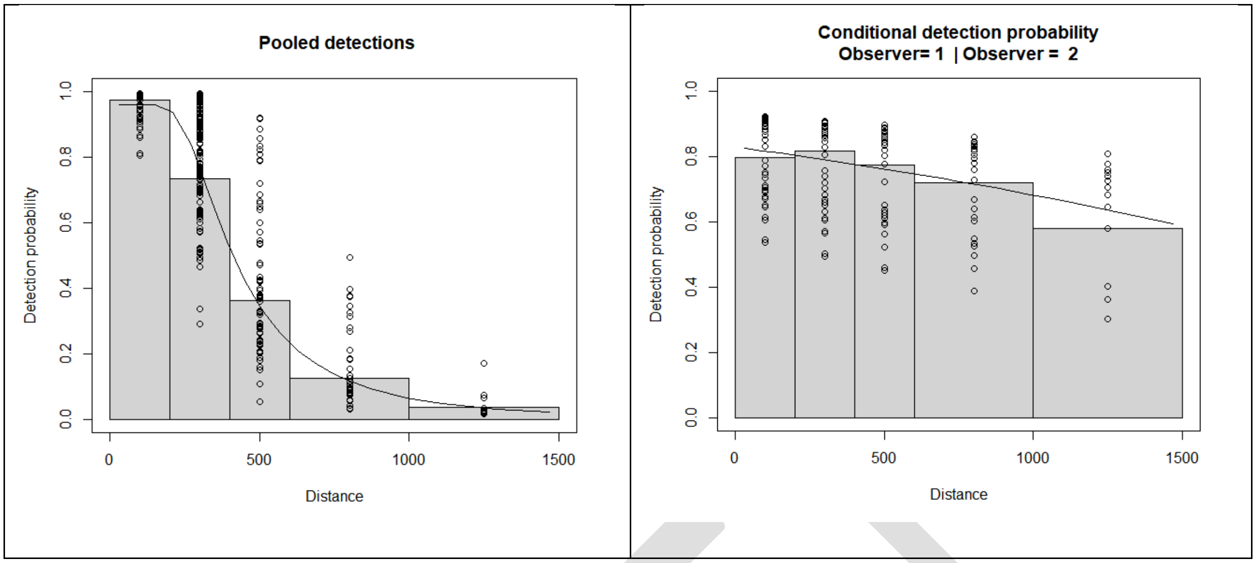

Figure 4. Distance sampling detection function for the full MRDS analysis of Baffin Island (left) and double observer estimates of detection with distance (right).

At the time of this report, estimates for the Baffin Island project had not been released so we use relative abundance, which is the estimate for each model divided by the estimate of the full MRDS model (Figure 5) for comparisons. Estimates when the data was subsetted to reduced numbers of bins were different than South Hampton Island with 3 bins required to obtain an estimate using a hazard rate model that was similar to the full estimate. At 2 survey bins, only the half-normal model converged leading to higher (+10%) estimates presumably due to the half-normal model underestimating the area of constant sightability in the vicinity of the plane. If only double observer model methods were used, estimates were 10% lower than the full MRDS model estimates.

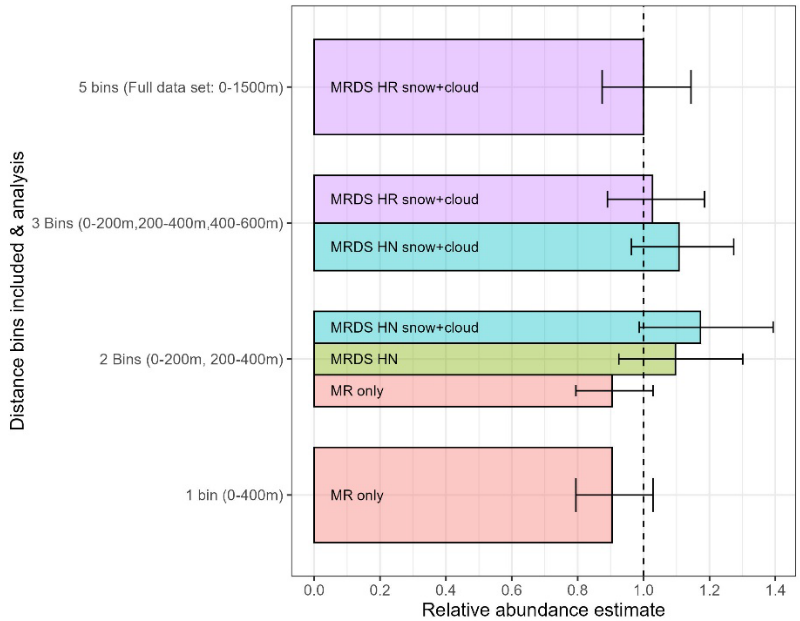

Figure 5. Baffin Island 2024 analysis. Estimates are given as relative abundance, which is the estimate for each model divided by the estimate of the full 5 bin MRDS model. The same double observer/mark-recapture model (ob6+ob11+distance+logsize) was used for all analyses shown. The dashed line indicates the estimates using the full 5 bin (0-1500m) data set.

<u>Simulation study</u>

One of the main questions addressed by the simulation study was whether the reduced bin designs could produce unbiased estimates of abundance. The main source of bias would be the inability of distance models to adequately estimate the shoulder area of constant sightability near the plane and/or the reduction in detection across the range of distances used in the survey. An additional source of bias could be lack of resolution in the data to allow selection of the hazard rate model compared to the simpler half-normal model. For simulations a hazard rate generating model used was based on the Baffin Island detection function with a shoulder of approximately 200m of constant sightability in the vicinity of the plane (Figure 4). A hypothetical study area based on the High East Beverly strata was simulated with a population of 3000 caribou. Figure 6 shows an example of one run of the simulations. For estimation, the data was binned in 5 distance bins (0-200, 200-400, 400-600, 600-1000, and 1000-1500m). Single caribou were simulated with no group size variation in detection probabilities.

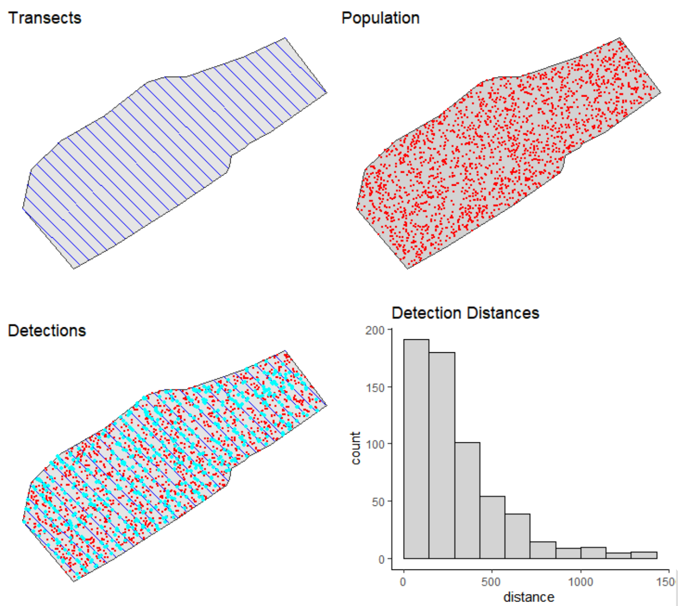

Figure 6 Example of one simulation run with transects, simulated caribou (red dots), detections (blue dots) and simulated distance data. The actual bins used for analyses were (0-200,200-400,400-600,600-1000, and 1000-1500m).

Simulations generated 686 to 919 detections of caribou depending on the number of bins included in the simulation (Table 1). These sample sizes were approximately similar to the South Hampton and Baffin Island data sets. Simulation study results suggested that the hazard rate model could provide unbiased estimates with 97% confidence interval coverage when 3 bins of sampling were used for the analysis (Figure 7 and Table 1). When 2 bins were used, estimates were slightly positively biased with occasional large estimates with non-model convergence in 3% of simulations, however confidence interval coverage was still adequate (94%). In contrast, estimates from the half-normal detection function were positively biased in most simulations with 2 or 3 bins of sampling with lower confidence interval coverage. When AIC model selection was used to choose between the hazard rate and simpler half normal model the hazard rate model was chosen in 61% of simulations with 2 or 3 bins sampled. As a result, estimates from simulations using AIC model selection resulted in an overall positive bias due to the half normal model being selected in simulations. If 5 bins were simulated, the hazard rate model was selected in 100% of simulations and estimates were unbiased.

Table 1. Simulation results. The mean number of caribou detected, method used to obtain estimates (AIC selection of HR or HN, or HN and HR models), % mean bias, confidence interval coverage (% of simulations where the confidence interval included the true value), mean CV of estimates, proportion of simulations using AIC where the hazard rate (HR) model was selected, and proportion of simulations where models did not converge.

Distance Bins used	Sampling width (m)	Mean caribou detected	Model method	Bias (%)	Conf. Interval Coverage	Mean estimate CV	Proportion AIC HR selected	Proportion failures
5	1500	919.3	AIC	0.2	0.96	0.046	1.00	0.00
3	600	803.2	AIC	4.4	0.73	0.079	0.61	0.00
2	400	684.9	AIC	3.5	0.90	0.055	0.61	0.04
3	600	805.5	HN	9.0	0.57	0.094		0.00
2	400	686.0	HN	5.0	0.87	0.074		0.00
3	600	805.0	HR	0.3	0.97	0.049		0.00
2	400	689.2	HR	3.4	0.93	0.063		0.03

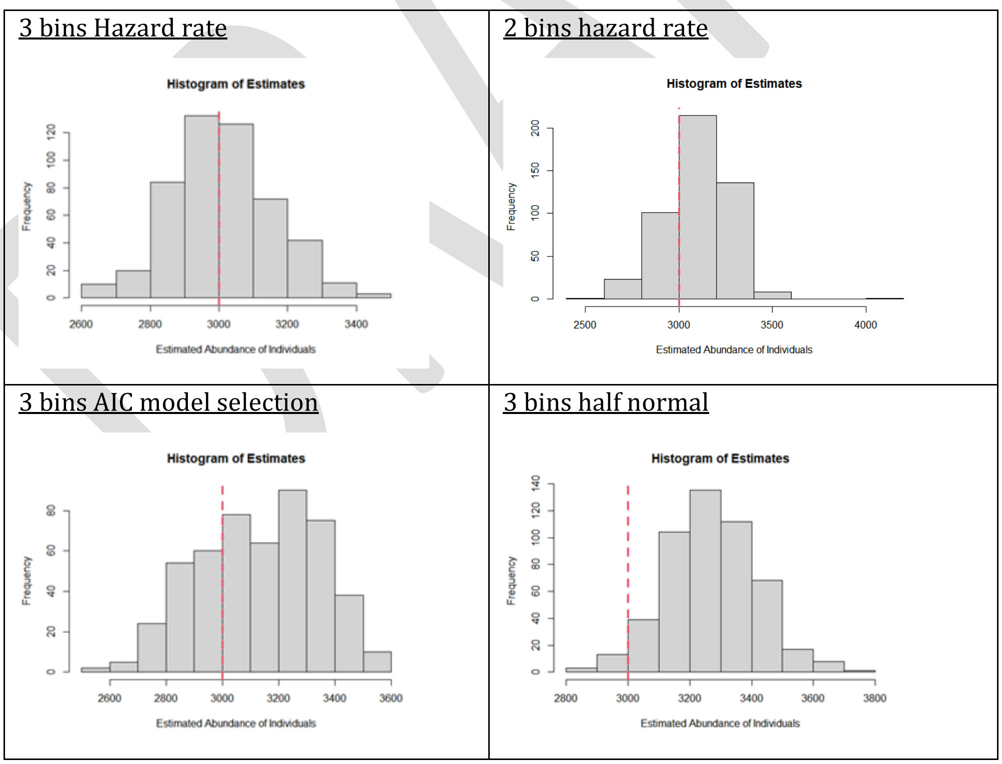

Figure 7. Examples of the distributions of simulated estimates compared to true value (red line) for each simulation run.

Discussion

Results from this analysis suggest that it is possible to get robust estimates using a reduced distance sampling and double observer approach. Of the data formulations considered the 3-bin (0-200m, 200-400m, 400-600m) provided the most reliable estimates based on both empirical comparisons and the simulation study. One reason for this is that it was possible to fit a hazard rate model that assumes an area of constant sightability near the plane (which is likely) with 3 bins of data. If 2 bins are used, then the half-normal model had to be used which is more restrictive in shoulder width leading to a reduced shoulder area and an overestimate of abundance.

Empirical comparisons provided a useful first step in assessing this methodology, however they were limited in that only one outcome of the survey is considered, and the true abundance was not known. The simulation study allowed an estimate of actual bias which further supported empirical comparisons. The simulation study was a simplification of field efforts given that detectability on the line was assumed to be 1 and group size variation was not simulated. Therefore, the levels of precision from the simulation study will be optimistic. However, simulation results provide the same general conclusion as empirical comparisons which is reassuring.

Empirical comparison of double observer and distance sampling estimates from past surveys suggest that double observer approach may underestimate abundance by 5-10% which is a reasonable difference but not enough to change general survey conclusions. Other comparisons also suggest that double observer estimates may be biased low. For example, IPM analysis of the Beverly 2023 survey also suggests that adult female estimates were approximately 5% lower in 2018 that IPM predictions. A comparison of paired photo and double observer survey lines during the Bluenose-East surveys in 2013 and 2023 surveys suggested that photo estimates were 7 to 14% higher, however, further analysis suggested that the main bias occurred on lines with higher densities of caribou where photo counts were higher with similar estimates at moderate densities (Boulanger et al. 2024). Regardless, negative biases in strip transect/double observer estimates cannot explain the increase documented in the 2023 survey or the increase documented in between 2018 and 2023 for the Bluenose-East herd. However, it is possible that double observer estimates showed a greater degree of negative bias in years such as 2021 Bluenose-East survey when survey conditions were poor. In this case, the proposed approach may provide a better estimate of abundance when survey conditions are challenging.

One additional consideration with visual survey estimates is potential counting bias when densities of caribou are higher. In the Beverly 2023 survey it was found that visual estimates may be biased high when densities are very high due to challenges of determining whether caribou were within the survey strip. In this case, the video-assisted approach provides a secondary way to count caribou and can complement the use of double observer and distance methods. In general, it is assumed that visual surveys will be applied at moderate densities of caribou (<10 caribou/km^2) with video or photo plane methods used when densities are higher.

Simulation results suggest that unbiased estimates can be obtained with 2 or 3 bin sampling, however the 3-bin approach (0-200,200-400, 400-600m) is preferable to reduce the risk of not having enough data to adequately determine the shoulder area of constant sightability. For example, the 2-bin design only selected the hazard rate model in 61% of simulations leading to a positive bias if the half-normal model was used for estimates. One strategy to confront this issue would be to only consider a hazard rate model for the analysis. The hazard rate model is considered to be "model robust" in that it can still produce unbiased estimates when the width of the shoulder of constant sightability near the survey line is not large (Buckland et al. 1993). Further simulations could be conducted to explore a greater range of survey conditions.

The main difference between the South Hampton and Baffin Island surveys in this analysis and calving ground surveys is the reduced densities of caribou which made it easier to apply distance sampling methods. For example, on the Baffin Island survey densities were less than 1 caribou per km^2 and therefore would be considered low density in a calving ground survey. During calving ground surveys observers and data recorders are more preoccupied with counting caribou groups and therefore successful application of this approach requires minimal additional complexity for observers and data recorders. However, we suggest the only additional piece of information that would be required is what survey bin the majority of any group occurs in. Also, observers would need to be reminded to pay the most attention to the closest survey bins (if the 3 bin approach is used). We propose that this additional complexity is worth considering given the likely robustness of this approach compared to just using the double observer method for calving ground survey estimates.

Literature cited

Borchers, D. L., W. Zucchini, and R. M. Fewster. 1998. Mark-recapture models for line-transect surveys. Biometrics **54**:1207–1220.

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

Boulanger, J., J. Adamczewski, J. Williams, D. Cluff, K. Clark, S. Goodman, K.-S. Chan, and R. Abernathy. 2022. Estimates of Breeding Females & Adult Herd Size and Analyses of Demographics for the Bluenose-East Herd of Barren-Ground Caribou: 2021 Calving Ground Photographic Survey Environment and Natural Resources, Yellowknife, NWT.

Boulanger, J., J. Adamczewski, J. Williams, S. Goodman, K. Clark, R. Abernathy, and L. LeClerc. 2024. June 2023 Calving Ground Surveys: Bluenose-East and Bathurst Barren-Ground Caribou Herds Environment and Climate Change, Govt of Northwest Territories, Yellowknife NWT.

Boulanger, J., and M. Campbell. 2025. Use of a reduced distance sampling and double observer methodology to improve robustness of calving ground survey estimates: Appendix in Campbell, M., J. Boulanger, J. Ringrose, J. Danahy, R. Kite, and A. Roberto-Charron. 2025. Abundance and Trends of the Beverly Mainland Migratory Subpopulation of Barren-Ground Caribou (Rangifer tarandus groenlandicus) June 2011 – June 2023. Department of Environment, Government of Nunavut, Arviat, Nunavut. Government of Nunavut, Arviat, Nunavut.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993. Distance Sampling. Estimating Abundance of Biological Populations. Chapman & Hall, London.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2004. Advanced Distance Sampling - Estimating abundance of biological populations. Oxford Press.

Buckland, S. T., J. Laake, and D. L. Borchers. 2010. Double-observer line transect methods : levels of independence Biometrics **66**:169–177.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: A practical information theoretic approach. Springer, New York, New York, USA.

Campbell, M., J. Boulanger, and D. Lee. 2020. Long-term trends in abundance and distribution of the Southampton Island caribou herd: 1978 – 2019 Government of Nunavut, Department of Environment, Arviat, NU.

Campbell, M., J. Boulanger, D. Lee, M. Dumond, and J. McPhearson. 2012. Calving Ground Abundance Estimates of the Beverly and Ahiak Subpopulations of Barren-Ground Caribou (Rangifer tarandus groenlandicus) – June 2011, Technical Summary. Department of Environment, Government of Nunavut.

Campbell, M., J. Boulanger, J. Ringrose, J. Danahy, R. Kite, and A. Roberto-Charron. 2025. Abundance and Trends of the Beverly Mainland Migratory Subpopulation of Barren-Ground Caribou (Rangifer tarandus groenlandicus) June 2011 – June 2023. Department of Environment, Government of Nunavut, Arviat, Nunavut.

Campbell, M., J. Goorts, D. S. Lee, J. Boulanger, and T. Pretzlaw. 2015. Aerial Abundance Estimates, Seasonal Range Use, and Demographic affiliations of the Barren-Ground Caribou (Rangifer tarandus groenlandicus) on Baffin Island – March 2014. Department of Environment, Nunavut Wildlife Research Division, , Arviat, Nunavut.

Campbell, M., D. Lee, and J. Boulanger. 2019. Calving Ground Abundance Estimates of the Beverly Subpopulations of Barren-Ground Caribou (Rangifer tarandus groenlandicus) – June 2018. Technical Summary. Government of Nunavut.

Campbell, M., D. Lee, J. Boulanger, A. Kelly, M. Dumond, and McPherson. 2014. Calving Ground Abundance Estimates of the Beverly and Ahiak Subpopulations of Barren-Ground

Caribou (Rangifer tarandus groenlandicus) – June 2011,Government of Nunavut ,Department of Environment Technical Report Series – No: 01-2013.

Campbell., M., J. Ringrose, J. Boulanger, A. Roberto-Charron, K. Metheun, C. Mutch, T. Davison, and C. Wray. 2021. An Aerial Abundance Estimate of the Dolphin and Union Caribou (Rangifer tarandus groenlandicus x pearyi) herd, Kitikmeot Region, Nunavut – Fall 2020, Government of Nunavut, Technical Report Series No 01-2021. 146 pp. .

Cook, D. R., and J. O. Jacobsen. 1979. A design for estimating visibility bias in aerial surveys. Biometrics **35**:735–742.

Elphick, C. S. 2008. How you count counts: the importance of methods research in applied ecology. Journal of Applied Ecology **45**:1313–1320.

Grolemund, G., and H. Wickham. 2011. Dates and Times Made Easy with lubridate. Journal of Statistical Software **40**:1 – 25.

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

Laake, J., D. L. Borchers, L. Thomas, D. Miller, and J. Bishop. 2012. Mark-recapture distance sampling (MRDS) 2.1.0. R statistical package program.

Laake, J., M. J. Dawson, and J. Hone. 2008a. Visibility bias in aerial survey: mark-recapture, line-transect or both? Wildlife Research **35**:299–309.

Laake, J., R. J. Guenzel, J. L. Bengtson, P. Boveng, M. Cameron, and M. B. Hanson. 2008b. Coping with variation in aerial survey protocol for line-transect sampling. Wildlife Research **35**:289–298.

Laake, J. L., and B. A. Collier. 2024. Understanding implications of detection heterogeneity in wildlife abundance estimation. The Journal of Wildlife Management **88**:e22516.

Marshall, L. 2022. dsims: Distance Sampling Simulations R package.

Mazerolle, M. J. 2016. AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). . R package version 2.1-0.: https://cran.r–project.org/package=AICcmodavg.

Pebesma, E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal **10**:439–446.

QGIS Foundation. 2020. QGIS Geographic Information System. QGIS Association. http://www.qgis.org.

R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rexstad, E., S. Buckland, L. Marshall, and D. Borchers. 2023. Pooling robustness in distance sampling: Avoiding bias when there is unmodelled heterogeneity. Ecology and Evolution **13**:e9684.

Thomas, D. C., S. T. Buckland, E. A. Rexstad, J. Laake, S. Strindberg, S. L. Hedley, J. R. B. Bishop, T. A. Marques, and K. P. Burnham. 2009. Distance software: Design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology **47:** :5–14.

Wickham, H. 2009. ggplot2: Elegant graphics for data analysis. Springer, New York.

Wickham, H. 2011. The Split-Apply-Combine Strategy for Data Analysis. Journal of Statistical Software **40**:1 – 29.