

NARWHAL ABUNDANCE IN THE EASTERN CANADIAN HIGH ARCTIC IN 2013

Thomas Doniol-Valcroze^{1,2}, Jean-François Gosselin¹, Daniel G. Pike³, Jack W. Lawson⁴, Natalie C. Asselin⁵, Kevin Hedges⁵ & Steve H. Ferguson⁵

¹ Maurice-Lamontagne Institute, Fisheries and Oceans Canada, Mont-Joli, Canada. Corresponding author: thomas.doniol-valcroze@dfo-mpo.gc.ca

² Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, Canada.

³ Esox Associates, North Bay, Canada.

⁴ Northwest Atlantic Fisheries Centre, Fisheries and Oceans Canada, St. John's, Canada.

⁵ Freshwater Institute, Fisheries and Oceans Canada, Winnipeg, Canada.

ABSTRACT

In summer, narwhals (*Monodon monoceros*) migrate from Baffin Bay to northeastern Canada and northwest Greenland, where they are hunted by Inuit for subsistence. To prevent localized depletion, management of narwhals is based on summer stocks. The High Arctic Cetacean Survey (HACS), conducted in August 2013, was the first survey to estimate abundance of all 4 Canadian Baffin Bay narwhal summer stocks, as well as putative stocks in Jones Sound and Smith Sound, in the same summer. Narwhal abundance was estimated using a double-platform aerial survey. Distance sampling methods were used to estimate detection probability away from the track line. Mark-recapture methods were used to correct for the proportion of narwhals missed by visual observers on the track line (i.e., perception bias). We used a data-driven approach to identify single and duplicate sightings, using 4 covariates to compare differences in sightings made by front and rear observers based on: time of sighting, declination angle, group size, and species identity. Abundance in fjords was estimated using density surface modelling to account for their complex shape and uneven coverage. Estimates were corrected for availability bias (narwhals that are not available for detection because they are submerged when the aircraft passes overhead) using a new analysis of August dive behaviour data from narwhals equipped with satellite-linked time depth recorders. Corrected abundance estimates were 12,694 (95% CI: 6,324–25,481) for the Jones Sound stock; 16,360 (95% CI: 3,833–69,836) for the Smith Sound stock; 49,768 (95% CI: 32,945–75,182) for the Somerset Island stock; 35,043 (95% CI: 14,188–86,553) for the Admiralty Inlet stock; 10,489 (95% CI: 6,342–17,347) for the Eclipse Sound stock; and 17,555 (95% CI: 8,473–36,373) for the East Baffin Island stock. Total abundance for these 6 stocks was estimated at 141,908 (95% CI: 102,464–196,536). Sources of uncertainty arise from the high level of clustering observed, in particular in Admiralty Inlet, Eclipse Sound, and East Baffin Island, as well as the difficulty in identifying duplicate sightings between observers when large aggregations were encountered.

Keywords: narwhal, Baffin Bay, summer stocks, abundance, aerial survey, double-platform, distance sampling, fjords, density surface modelling.

INTRODUCTION

A large meta-population of narwhals (*Monodon monoceros*) overwinters in Baffin Bay and Davis Strait (Heide-Jørgensen, Hansen, Westdal, Reeves, & Mosbech, 2013). In late spring, these narwhals migrate to the fjords and inlets of northeastern Canada and northwest Greenland, where they spend the summer before migrating back to their wintering grounds in Davis Strait and Baffin Bay in late fall. Inuit from across Nunavut hunt narwhals for subsistence on the whales' summering grounds or along their spring and fall migration routes, particularly in the Qikiqtani and Kitikmeot regions, which include the Boothia Peninsula, Baffin Island, and the Canadian High Arctic communities (Priest & Usher, 2004). In addition, Greenlandic Inuit also hunt Baffin Bay narwhals in winter. The sustainable management of this harvest activity in the Canadian High Arctic, which is of great economic and cultural importance, relies on obtaining up-to-date estimates of abundance.

Narwhals appear relatively sedentary in summer, with telemetry studies indicating that tracked animals tend to remain in the aggregation areas where they were tagged and rarely visit other summering areas (Dietz, Richard, &

Acquarone, 2001; Dietz et al., 2008). Narwhals are also believed to exhibit inter-annual site fidelity by returning to the same summering areas every year (Dietz et al., 2008; Heide-Jørgensen et al., 2003), although there have been exceptions (Watt, Orr, Leblanc, Richard, & Ferguson, 2012; Heide-Jørgensen, Richard, Dietz, & Laidre, 2013). It is possible to have localized depletions or extinctions if site fidelity is not considered when harvesting occurs, even if harvesting is confined to a single, more widespread biological stock (Cope & Punt, 2009). The population in Baffin Bay and surrounding waters has therefore been divided into smaller management units that represent seasonally spatially discrete stocks, believed to have little or no exchange during summer. The use of summering stocks as management units in this way is considered precautionary.

Narwhals in Baffin Bay are divided into 8 stocks, or summer aggregations (Figure 1), that migrate between, and are susceptible to hunting in, Greenland and Canada (North Atlantic Marine Mammal Commission [NAMMCO], 2018). Two of these stocks occur in Greenland: Melville Bay and Inglefield Bredning. However, Inglefield Bredning narwhals have not been tracked

out of their summering area and their wintering area is unknown. Therefore, their relationship to the Baffin Bay population remains unclear. Six of these stocks occur in Canadian waters: Somerset Island (SI), Admiralty Inlet (AI), Eclipse Sound (ES) and East Baffin Island (EB), as well as 2 putative stocks for which migration patterns are unknown and that may represent separate population(s): Smith Sound (SS) and Jones Sound (JS). Some stocks have been surveyed numerous times and have recent abundance estimates, while others have only been surveyed once or have dated estimates (Table 1, Higdon & Ferguson, 2017). Several stocks cover large areas and may have further sub-structuring. For instance, the SI stock has the largest areal extent of the Baffin Bay population summering stocks and includes Prince Regent Inlet and the Gulf of Boothia, Peel Sound, Barrow Strait, and northern Foxe Basin. Systematic surveys of this summer stock, or parts thereof, were conducted in 1981 (Smith, Hammill, Burrage, & Sleno, 1985), 1984 (Richard, Weaver, Dueck, & Barber, 1994), 1996 (Innes et al., 2002), and 2002-2004 (Richard et al., 2010). For JS and SS stocks, information on summer abundance was lacking altogether. Narwhals are also known to occur elsewhere in small numbers in the Canadian High Arctic during summer (e.g., Parry Islands, Cambridge Bay), but no narwhal surveys have been conducted in these areas. On the Greenland side, the Melville Bay summering stock was estimated to number 3,091 (CV=0.50; 95% CI: 1,228–7,783) in 2014 (NAMMCO, 2015) and the Ingfield Bredning summering stock was estimated to number 8,368 (95% CI: 5,209–13,442) in 2007 (Heide-Jørgensen et al., 2010).

Previous surveys (Asselin & Richard, 2011; Richard et al., 2010), land-based observations (Marcoux, Auger-Méthé, &

Humphries, 2009) and telemetry studies (Dietz et al., 2001), as well as a DFO reconnaissance survey flown along the coast of Ellesmere Island in 2012, showed that narwhals in some areas spend a large proportion of their time inside narrow inlets and fjords on their summer range. Thus, any survey effort must include these areas to provide a credible abundance estimate. Moreover, fjords are often where Inuit hunters observe and harvest narwhals, and therefore a better understanding of the occurrence and space use of narwhals within these areas is particularly relevant to management efforts.

However, the estimation of abundance of narwhals in fjords presents several challenges. First, most fjords cannot be surveyed with systematic lines because they are often too narrow and steep-walled, making it impossible or unsafe to maintain a constant altitude near the edges. Second, fjords vary in width, which means that in some cases, the entire fjord can be seen from the aircraft and clipping of the observation field-of-view occurs at various points along the shoreline. Standard distance analysis is then made complex because of unequal coverage probability for different segments of areas surveyed. In addition, these fjords are numerous and sometimes separated by large distances. For these reasons, fjord surveys require specialized field and analytical methodologies.

The objective of the High Arctic Cetacean Survey (HACS), a large-scale aerial survey conducted in August 2013, was to obtain new abundance estimates, in the same year, for all 6 of the Baffin Bay narwhal summering stocks present in the Canadian High Arctic. Up-to-date stock abundance information can then be used for management advice.

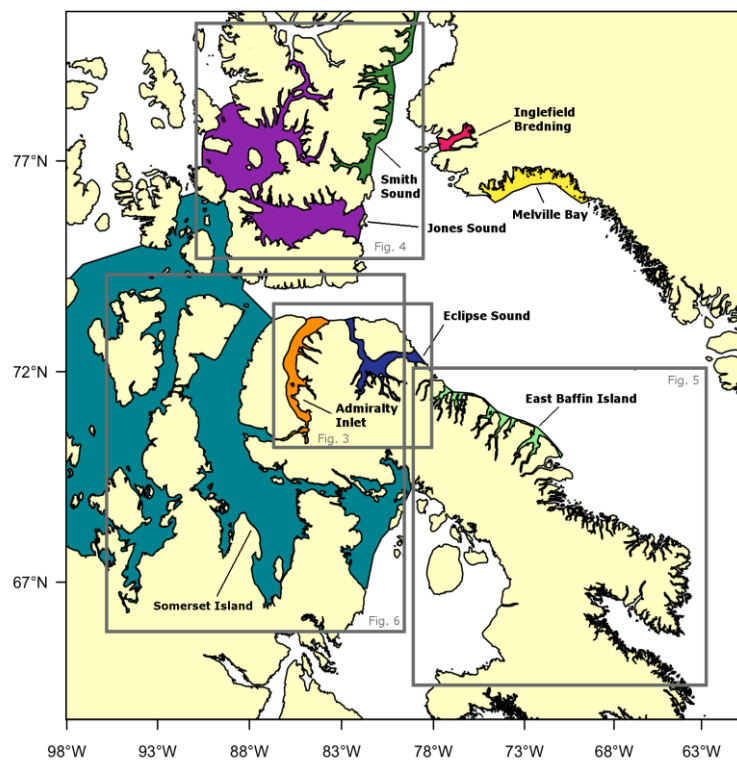


Figure 1. Map of High Arctic narwhal summer stocks (adapted from NAMMCO 2018). Melville Bay and Ingfield Bredning are summering aggregations in Greenland recognized as part of the Baffin Bay population by the Joint Commission of Narwhal and Beluga, and were not considered in this study. Boxes show the areas covered by Figures 3, 4, 5 and 6.

Table 1. Estimated numbers of narwhal in recognized Baffin Bay summer stocks. All estimates were fully corrected for perception and availability bias in the original studies. Alternative estimates are noted in the comments where appropriate. Adapted from Higdon & Ferguson (2017).

Summer stock	Survey year	Abundance estimate	CV	Source	Comments
Admiralty Inlet	2010	18,049	0.25, 0.40	Asselin & Richard 2011	Two complete surveys (mean estimates 24,398, 13,729), averaged using effort-weighted mean. Photographic survey of aggregations.
Somerset Island					
<i>Barrow Strait</i>	2004	2,925	0.46	Richard et al., 2010	Higher estimate in 1996 survey (Innes et al., 2002), mean = 5,898, but less precise (CV=0.75).
<i>Prince Regent Inlet</i>	2002	20,871	0.71	Richard et al., 2010	Estimate from 1996 survey (Innes et al., 2002) larger and more precise (mean = 34,159; CV=0.35).
<i>Peel Sound</i>	1996	5,240	0.6	Innes et al., 2002	Similar to 1984 estimate of 1701 on surface (CV=0.25) (Richard et al. 1994) when corrected by 3.1 (Asselin & Richard 2011) (5,273).
<i>Gulf of Boothia</i>	2002	6,770	0.3	Richard et al., 2010	First time surveyed.
<i>Sum</i>		35,806			Total = 52,057 if higher estimates from 1996 (Innes et al., 2002) are used for Barrow Strait and Prince Regent Inlet.
Eclipse Sound	2004	20,225	0.36	Richard et al., 2010	
East Baffin Island	2003	10,073	0.31	Richard et al., 2010	
Total		84,153			Total estimate of 100,404 narwhals if 1996 results for Somerset Island stock are used.

MATERIALS AND METHODS

Study area and survey timing

The extent of the HACS area was based on the aerial abundance surveys done since the 1970s, telemetry tracking studies (Dietz et al., 2001), reports of Inuit Traditional Knowledge (ITK) and recent observations by Inuit hunters (Priest & Usher, 2004). The 6 major summering stocks in the Canadian High Arctic cover a large area from the east coast of Baffin Island to the Central Arctic Archipelago, and possibly farther west (Richard, 2010). Because of assumed inter-annual fidelity to summering sites and logistical challenges, previous surveys of Baffin Bay narwhal stocks were conducted by covering summering areas separately over several years (Richard et al., 2010). However, recent concerns about potential exchanges among neighbouring summering areas (Dietz et al., 2001; Heide-Jørgensen, Dietz, Laidre, & Richard, 2002; Watt et al., 2012) have made it desirable to attempt to survey all stocks over as short a period

as feasible in order to avoid biases that may arise from animals moving between strata over the course of the survey.

Therefore, it was decided that the HACS would attempt to cover the summering areas of each stock in the same year, with a priority on the Jones Sound, Smith Sound, and Somerset Island stocks (Figure 1). Since little was known of the distribution of narwhals in the waters around Ellesmere Island (JS and SS stocks), a reconnaissance aerial survey was performed in late August 2012 with members of the Grise Fjord community. Although fog prevented coverage of the offshore areas, numerous narwhals were observed in coastal waters and deep inside fjords, as far north as Alexandria fjord, which was taken into account when planning the extent of the 2013 survey (Figure 2). For the Somerset Island stock, it was considered unrealistic to cover the entire known distribution range, which potentially extends at low densities far to the west and southwest of Prince of Wales Island. It was decided to focus instead on the presumed core areas of Prince Regent Inlet, Peel Sound, and the Gulf of Boothia.

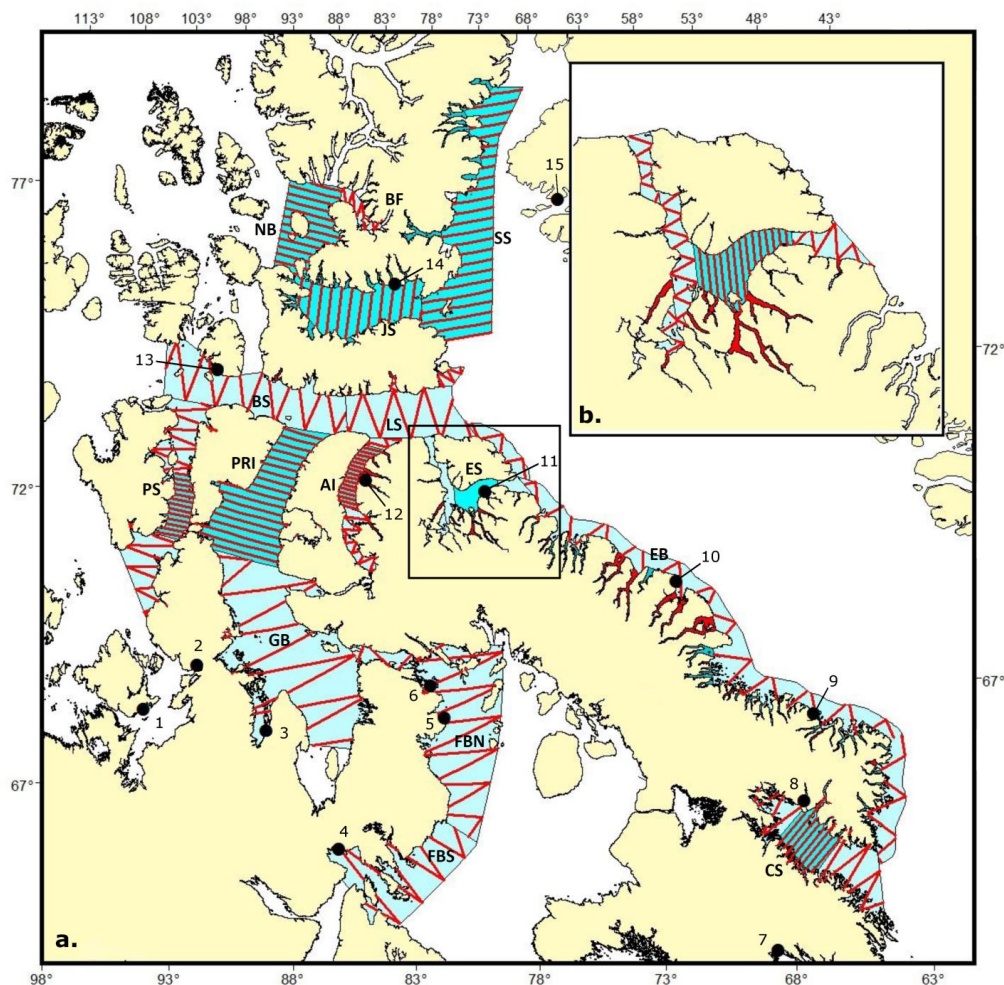


Figure 2. a. Map of planned survey strata (blue polygons), transect lines (red lines), and fjord strata (red areas). AI: Admiralty Inlet. BF: Baumann Fjord. BS: Barrow Strait. CS: Cumberland Sound. EB: East Baffin. ES: Eclipse Sound. FBN: Foxye Basin North. FBS: Foxye Basin South. GB: Gulf of Boothia. JS: Jones Sound. LS: Lancaster Sound. NB: Norwegian Bay. PRI: Prince Regent Inlet. PS: Peel Sound. SS: Smith Sound. Communities (black dots): 1. Gjoa Haven; 2. Taloyoak; 3. Kugaaruk; 4. Repulse Bay; 5. Hall Beach; 6. Igloodik; 7. Iqaluit; 8. Pangnirtung; 9. Qikiqtarjuaq; 10. Clyde River; 11. Pond Inlet; 12. Arctic Bay; 13. Rolute Bay; 14. Grise Fjord; 15. Qaanaaq (Greenland). b. Inset: enlargement of the Eclipse Sound stratum (boxed area).

Dates for the survey were established based on the short period of relatively ice-free waters in the Arctic Archipelago and the historical timing of narwhal aggregations in their summering areas. The best time was determined to be August, when telemetry studies show that whales are relatively sedentary within their summering range and the weather is most favourable. Later than the third week of August, there was a risk that some narwhals would have begun migrating and either moved between strata or outside of the study area (Watt et al., 2012).

Survey design

Design-based strata

The survey was designed to balance 2 conflicting objectives: to cover the largest possible proportion of the summering areas of narwhal stocks while at the same time improving on the precision of past estimates, which required higher coverage over a larger area than previous surveys. In order to minimise the sampling error, we divided each stock range in several strata (Figure 2) based on geographic boundaries, telemetry studies, and observed densities of narwhals from past surveys. For instance, the Somerset Island stock comprised the strata Prince Regent Inlet, Peel Sound “high intensity”, Peel Sound “low intensity”, and Gulf of Boothia. Where data from previous studies were not available, we relied on traditional knowledge and the observations made during the 2012 reconnaissance survey.

Transect design was performed in Distance 6 (Thomas et al., 2010) and the projection of each stratum was selected to minimize distortions of area using Young’s rule (Maling, 1992). The first transects of each stratum were chosen at random and the others were spaced at regular intervals (i.e., the design was systematic with a random start). As much as possible, transect lines were oriented in a direction perpendicular to the longest axis of the stratum to provide a maximum number of lines (sampling units) per stratum. In most areas, this also meant that transects were perpendicular to the presumed density gradient of narwhals, which often aggregate along coastlines. For presumed high-density strata, we used systematic parallel transects with greater coverage (7–15%) than had been used in the past. Areas where we expected lower densities of narwhals were covered using zigzag transects with equally spaced endpoints. Parallel line transects are preferred over zigzag, especially in high coverage area, as they maintain uniform coverage probability. However, zigzag transects are more efficient to cover wide areas as transit time between transects is reduced. Some low-coverage strata had complex geographic shapes that were divided in subareas where equally spaced zigzag designs were created using the same spacing for the whole stratum. Using convex hull shapes and the longest axis of the subareas as the main design axis allowed us to maintain a relatively equal coverage probability within these strata.

The sequence of stratum coverage was designed to survey High Arctic narwhal summer aggregation areas in order of decreasing priority, with high priority given to strata with presumed highest densities, older survey estimates, and high management concerns, and with the condition that adjacent strata considered to be part of the same summering stock be covered within a short time window to avoid the possible influence of animal movements among strata. Weather permitting, the strata were to be surveyed in the following order: 1) Jones

Sound/Smith Sound/Norwegian Bay (high densities and numbers, never surveyed systematically, and a high management priority); 2) Prince Regent Inlet (highest density and numbers of all stocks, old estimate) and the neighbouring Peel Sound (high narwhal numbers seen in the past, exchange of narwhals from Prince Regent Inlet through Bellot Strait); 3) Admiralty Inlet and Eclipse Sound (high density and numbers, more recent estimate; ideally both would be surveyed simultaneously as it was recently demonstrated that there can be exchange of whales in August, (Watt et al., 2012)); 4) East Baffin Island (high densities, old estimate); and 5) Gulf of Boothia (presumed low-density area for the Somerset Island stock).

To avoid the effect of potential directed movements of narwhals within strata, attempts were made to survey each stratum in a day or two. Unpredictable weather also makes single-day stratum coverage desirable. For large or remote areas, this often required the use of more than 1 aircraft. All 3 survey aircrafts initially combined their effort to complete surveys of priority strata 1 and 2 in the list above. Each of the 3 aircraft then deployed north (priority 3), east (priority 4), or west (priority 5) of Baffin Island.

Fjord strata

Effort allocation among fjords

Fjords were sampled separately from the main water bodies in 6 areas, each considered a distinct stratum: West Ellesmere (WEF), Jones Sound (JSF), Smith Sound (SSF), Admiralty Inlet (AIF), Eclipse Sound (ESF) and East Baffin Island (EBF). Ideally, every fjord in a given stratum would be surveyed. This was possible in AIF and ESF. However, in the other strata, the number of fjords and distances between them made it impossible to survey all of them. Therefore, following Thomas, Williams, & Sandilands (2007), we used a 2-stage sampling design. At stage 1, each fjord was considered a primary sampling unit (PSU) and a custom algorithm was used to select a subset of PSUs where each fjord had a probability of being selected proportional to its area in an attempt to maintain equal coverage probability within fjord strata. At stage 2, distance sampling was conducted within each selected fjord.

A GIS was used to select and clip sections of the shoreline that were considered separate fjords. This process relied on published nautical charts and local knowledge, and was somewhat arbitrary when fjords had complex shapes with multiple openings into a larger body of water, or when it was difficult to distinguish a fjord/inlet from a bay. Any fjord smaller than a cut-off value of 20 km² was excluded from the design. This process yielded a total of 111 fjords, ranging in area from 21 to 1,236 km². Ideally, 10–20 fjords (PSUs) per stratum should be sampled to obtain a reliable abundance estimate using the distance sampling approach (Buckland et al., 2001), but when logistics preclude this, it is advisable to have at least 5 (Thomas et al., 2007).

The algorithm to select the fjords (PSUs) required the following properties:

1. The probability of selecting a fjord should be proportional to its area, so that each part of the stratum would have the same chance of being in a sampled fjord;

2. There should be a broad geographic range of fjords (e.g., from north to south or east to west), which necessitated the use of a systematic survey scheme;
3. No fjord should be selected twice.

This last property could be achieved by sampling without replacement (i.e., removing each fjord from the pool of potential samples once it is selected), but a disadvantage of this type of algorithm is that variance estimation is greatly complicated. Instead a systematic algorithm developed by Thomas et al. (2007) was used that samples with replacement, but fulfils the first 2 of the above criteria and has little chance of sampling the same fjord twice if sampling intensity is not too high. Each fjord identified as a PSU had a probability of being selected that was proportional to its area, giving equal coverage probability across the fjord strata.

All 14 fjords in Admiralty Inlet and Eclipse Sound (7 in each) could be surveyed (Figure 3). For the other strata, we decided to sample 5 out of 15 fjords in Smith Sound (Figure 4), 5 out of 17 in Jones Sound (Figure 4), and 10 out of 54 in East Baffin Island (Figure 5). Due to exceedingly large transit times, only 3 out of 11 fjords could be surveyed in West Ellesmere (Figure 4).

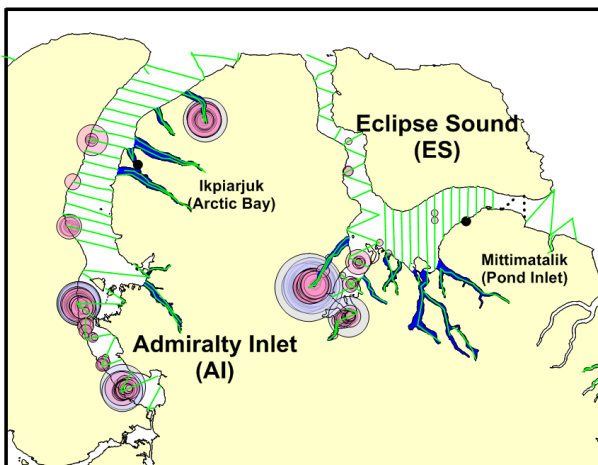


Figure 3. Admiralty Inlet and Eclipse Sound survey, showing completed transects (green), planned un-completed transects (black dashed) and fjord strata (blue). Sightings of narwhals are shown with symbol size proportional to group size from 1–10 (pink) and 11–600 (light blue).

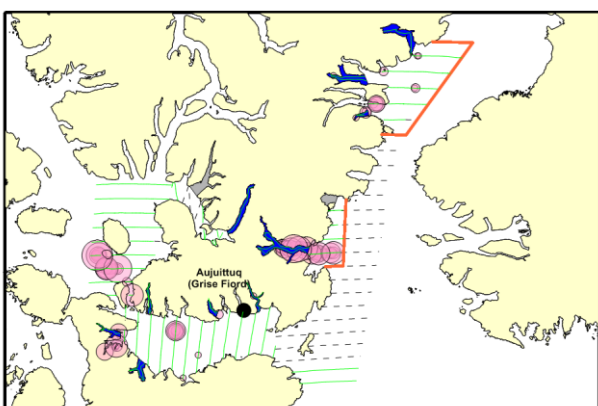


Figure 4. Norwegian Bay, Jones Sound and Smith Sound survey showing completed transects (green), un-completed planned transects (black dashed), surveyed fjord strata (blue), un-surveyed planned fjord strata (grey) and sightings of narwhals with symbol size proportional to group size (1–10). Orange box: post-stratified Smith Sound stratum.

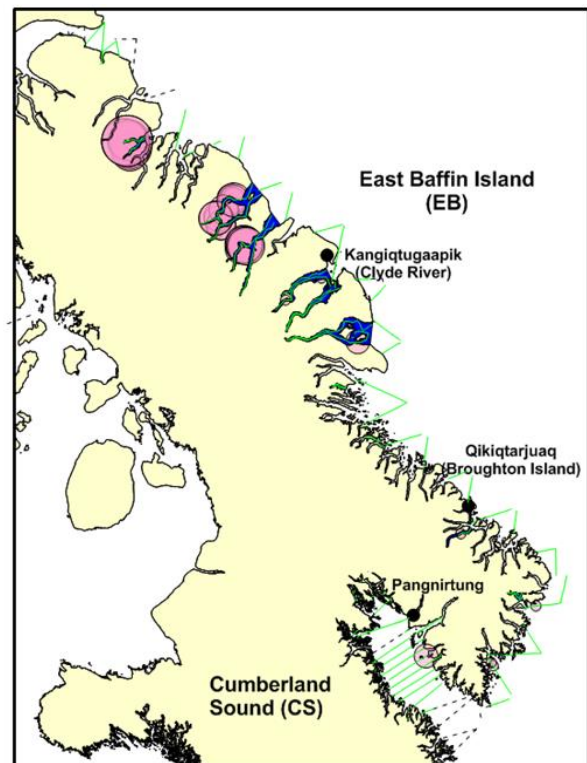


Figure 5. East Baffin and Cumberland Sound survey, showing completed transects (green), un-completed planned transects (black dashed), completed fjord strata (blue) and sightings of narwhals with symbol size proportional to group size (1–4).

Effort allocation within fjords

Unlike the survey described by Thomas et al. (2007), an equal coverage design could not be maintained within fjords. Many of the fjords were non-convex with long, thin sections, and cliffs on the sides that rose higher than the target survey altitude (305 meters). Therefore, it was impractical and unsafe to use systematic or random transects within the selected fjords, precluding the use of a standard line transect analysis for these strata. Flights were planned as continuous tracks and adjusted on site by the navigator to follow the main axis of each fjord, while aiming to spread coverage uniformly according to distance from the shore when the fjords were wide enough, and to minimise duplicate coverage of any area.

Survey methodology

The aerial survey was flown at an altitude of 305 m and a target speed of 100 knots (185 km/h) using 3 de Havilland Twin Otter 300 aircraft, each equipped with 4 bubble windows on the sides that allowed the observers to view the track line directly below the aircraft, and a large belly window used for cameras. Four observers were stationed at the front and rear bubble windows, with a fifth team member acting as a navigator and camera operator. The visual surveys were conducted as a double-platform experiment with independent observation platforms at the front (primary) and rear (secondary) of the survey plane. The 2 observers stationed on the same side of the aircraft were separated visually and acoustically to ensure independence of their conditional detections.

Observer training was conducted over 2 days, beginning 1st August in Resolute, NU. These sessions included classroom

presentations, on-the-ground training in the aircraft using moving practice targets, and practice flights around Resolute.

Observers used hand-held Sony PCM-D50 voice recorders to record their observations. A separate recording was made for each transect, and each recording was time-calibrated at its initiation, allowing it to be correlated with GPS time and position. Observers recorded the following data for each cetacean observation: time at which they sighted the group (“spot time”); time at which the group passed abeam (“beam time”); perpendicular declination angle of each sighting relative to the horizontal plane at beam time, measured using Suunto inclinometers; species; group size; direction of movement; number of tusked narwhal; and number of calves. A group was defined as 2 or more animals that were within 3 body lengths of each other and oriented or moving in a similar direction. Observers were instructed to give priority to the estimation of group size, especially when densities were high, followed by declination angle and then other variables if time permitted. Position and altitude of the plane were recorded every 2 s using a GPS connected to a laptop running electronic mapping software.

Primary observers recorded weather and sighting conditions at the beginning and end of transects and whenever conditions changed. The conditions noted included sea state (Beaufort scale), ice concentration (in tenths), cloud cover (%), fog (% cover and intensity), and angle of searching area affected by sun glare along with glare intensity (3 levels: “intense” when animals were certainly missed in the centre of glare angle, “medium” when animals were likely missed in the centre of glare angle, “low” when animals were likely detected in centre of glare).

In addition to visual observations, the 3 aircraft collected continuous photographic records below the aircraft using dual oblique cameras (Nikon D-800, lens Zeiss Distagon-T 35 mm) pointing downwards through a belly window, aimed towards either side of the track line. A camera imaging interval of 3 s allowed a target overlap of 20% between successive photographs along the direction of travel. A GPS unit was connected to each camera, which was in turn connected to a laptop. Geo-referenced images were thus saved on the laptop in real time. The photographs were oriented with their longest side perpendicular to the track line, at an angle of 27 degrees. At an altitude of 305 m, the swath width of the pictures taken was 420 m, for a total strip width of 840 m at the surface of the water.

Data management and photo verification

Whale sighting and other observations were geo-referenced by matching the recording time with the GPS time to the nearest second and the corresponding location. Narwhal sightings and aircraft flight tracks were mapped using ArcGIS 9.2 (ESRI Inc.). Transect lengths and stratum areas were determined in ArcGIS using an Albers Equal Area conic projection with 71.75°N as the reference latitude, 85°W as the reference longitude, and 66°N and 77.5°N as standard parallels, and using the 1980 Geodetic Reference System. Declination angles of abeam sightings were transformed into perpendicular distances by dividing the recorded altitude by the tangent of the angle.

Sightings made directly underneath the aircraft (i.e., on the track line) were examined to avoid duplication between right and left observers (and were assigned randomly to one side

when necessary). Sightings where declinations had not been recorded, or were coded as “uncertain”, were compared to the photographic records. The perpendicular distance was retrieved from the pixel position of the sighting on the photo if a visual sighting could be identified without ambiguity on the corresponding photo. If the sighting was not made within the swath width of the picture, could not be found, or could not be discriminated from other sightings unambiguously, it was coded as having a missing distance. These sightings were not used in fitting the detection function, but were added to the total count per transect, as described below. Sightings where group size had not been recorded or was coded as uncertain were also compared to the photographic records, and group size was retrieved if a match could be made based on perpendicular distance. Otherwise, sightings with missing group size were given the average group size in that stratum (posterior to estimation of the expected group size so that it did not affect the estimation of its variance).

Data analysis

Duplicate identification

Between-platform duplicate sightings are usually identified in aerial surveys by coincidence in sighting time and distance from the track line. While most previous studies have used simple thresholds in sighting time and/or declination angle to identify duplicates (reviewed in Pike & Doniol-Valcroze, 2015), the HACS dataset presented numerous problems for duplicate determination:

1. With 1,553 primary sightings and 2,565 potential duplicates with beam time differences of 10 seconds or less, an algorithm-based method of duplicate determination was required.
2. The groups of the 2 most common species seen, narwhal and beluga (*Delphinapterus leucas*), were highly aggregated, meaning that many sightings were often made in quick succession, often with beam time differences of as little as 1 s. Observer performance became degraded at such high encounter rates, and we expected the accuracy of measurement of beam time, declination, group size, and species identity to be affected.
3. For the same reason, there was a great deal of missing covariate data. Of the 2,630 sightings, 470 were missing valid declinations, 122 were missing group sizes and 18 were missing both. This problem is amplified because in any potential duplicate pair, the covariate was unusable if 1 member of the pair lacked a valid covariate value. Thus, of the 2,565 potential duplicates, 20% lacked a valid declination difference.
4. Group size for narwhal was tightly clustered in the range of 1–4. Therefore, we expected, on average, little difference in absolute group size between 2 sightings, even if they were not duplicates.

Several data-driven approaches have been explored to identify duplicate sightings during surveys. Hiby & Lovell (1998) used 2 aircraft in tandem formation about 9 km apart during harbour porpoise surveys. This spatial separation meant that the diving status of a group seen by one aircraft was independent from its status when passed by the other aircraft. Similarly, Stevenson, Borchers, & Fewster (2018) developed an estimator for 2 aircraft detecting individuals via high-definition cameras and

modelled detection locations as a clustered point process to estimate animal density without requiring individual identification. However, these approaches are not applicable to our HACS observers searching from the same aircraft. Instead, in this analysis, we adapted and extended the methodology developed by Southwell, de la Mare, Underwood, Quartararo, & Cope (2002) to identify duplicates from the HACS visual survey data. The dataset was segregated by side of the plane (R or L) and all sightings for each side were sorted by plane, date and beam time. For each sighting, the dataset was scanned forward 10 s, and all sightings made within this time range by the other station on the same side (rear for front sightings, front for rear sightings) were identified and paired with the initial sighting as a potential duplicate. Therefore, a sighting could have several potential duplicate matches within the 10 s interval. Potential duplicate pairs for the R and L sides were then combined into a single dataset incorporating all relevant data for each sighting.

For each potential duplicate pair, the following covariates, based on variables directly recorded by the observers, were derived:

- *T*: Difference in beam time, in seconds;
- *D*: Difference in declination angle, in degrees;
- *C*: Difference in group size.
- *S*: Difference in species identity (see Table 2).

Table 2. Difference in species certainty *S* for hypothetical species A and B, where U = Unknown. Certainty: 1 = certain, 0 = uncertain. Adapted from Southwell et al. (2002).

OBSERVER 1		OBSERVER 2		<i>S</i>
ID	CERT	ID	CERT	
A	1	A	1	0
A	1	A	0	0.25
A	0	A	0	0.25
A	0	U	1	0.25
U	1	U	1	0.25
A	0	B	0	0.5
A	1	U	1	0.5
A	1	B	0	0.75
A	1	B	1	1

Other potential covariates, such as direction of travel, number with tusks and number with no tusks, were recorded too infrequently to be of use in a statistical analysis.

The dataset described above contained sightings pairs that were duplicates and others that were not, in unknown proportion. Duplicate sightings pairs were expected to have a particular range and mix of covariate levels that differentiated them from non-duplicate sightings pairs. Some covariates may be more important than others in differentiating duplicates.

Choice of threshold levels for covariates used to identify duplicates is usually subjective. Here, we followed Southwell et al. (2002) in estimating threshold levels by examining graphs

showing the number of identified duplicates as covariate levels changed. It was expected that such curves would show a sharp initial increase followed by a levelling-off, with the inflection point being roughly equivalent to the covariate value below which most duplicates could be found.

In an effort to determine which covariates were most useful to identify duplicates, we created a second dataset that contained sighting pairs between front and rear observers that occurred close together in time but were certainly not duplicates. This was done in exactly the same way as the dataset described above, except that observer sightings were paired with stations on the opposite side of the plane. For example, a Front Right sighting would be paired with a Back Left sighting that occurred within the 10 s interval. This opposite-side dataset should be similar to the same-side dataset except that it could not possibly contain duplicate sightings pairs.

Logistic regression was used to determine which covariates were important in identifying data that contained duplicate pairs. The response variable was same-side (1) vs. opposite-side (0). Candidate logistic regression models were fit using all combinations of individual covariates *T*, *D*, *C* (if $C > 30$, it was considered missing) and *S*. For each case, the model with the highest Area Under Curve (AUC), representing the best reclassification performance, was chosen.

Separate models were created for:

1. Pairs for which all 4 covariates were available.
2. Pairs with missing declination angle differences *D*.
3. Pairs with missing count differences *C*.
4. Pairs missing both *D* and *C*.

Using the coefficients from the best model in each of these situations, regressions produced *p* values (coded here as *d*) corresponding to the probability that a particular sightings pair belonged to the same-side dataset, as opposed to the opposite-side dataset. Because the same-side dataset contained a mix of duplicate and non-duplicate pairs, these *d* values did not correspond directly to the probability that a pair was truly a duplicate. However, since the main difference between these 2 datasets was the presence of duplicate pairs in the same-side dataset, these scores were interpreted as a relative index of the probability that a particular sightings pair was a duplicate.

For each of these models, we calculated $d(0)$ the value of *d* when all covariates were at 0 (i.e., the maximum value of *d*). To be able to pool scores from the 4 models, we scaled *d* values by dividing them by $d(0)$ and used $1-d$ to obtain a dissimilarity index ranging between 0 (most likely to be a duplicate) and 1. We then substituted the covariate thresholds obtained by graphical methods (see above) into the regression equations to obtain a threshold value for *d*, identified as $d(T)$. This meant that a duplicate could exceed a threshold for a given covariate if the other covariate values were low, for example a candidate pair with a declination difference of 12 degrees might have been considered a duplicate even if the threshold was 10 degrees, if its time, count and species differences were low enough to result in a score below $d(T)$. To avoid extremely unlikely values, however, we placed a limit of 20 degrees on *D*.

Duplicates were selected by ranking all potential duplicate pairs with a score below $d(T)$ by their *d* value from lowest to highest, then selecting those with the lowest scores while removing selected sightings from the list.

Design-based strata

Uncorrected estimates

Mark-recapture distance sampling (MRDS) was used to estimate narwhal abundance in the design-based strata. This approach involves fitting 2 models, each with potentially different covariates: a distance sampling model fitted to all unique sightings, and a mark-recapture model fitted to the double observer data. Estimates of the density (\hat{D}) and abundance (\hat{N}) of narwhals at the surface during systematic survey of each design-based stratum, uncorrected for visible animals missed by observers (perception bias) or animals that were submerged and not visible during the passage of the aircraft (availability bias), were calculated using multiple covariate distance sampling (MCDS) techniques (Buckland et al., 2001). Analyses were carried out using the *mrds* package in R (Laake, Borchers, Thomas, Miller, & Bishop, 2019). A single, global detection curve was fitted to narwhal sightings from all non-fjord strata. The analyses were performed on the perpendicular distances of unique sightings (i.e., duplicate sightings, plus sightings made only by observer 1 plus sightings made only by observer 2). For duplicate sightings, we used the distance recorded by observer 1 (the most experienced observers), unless it was only available from observer 2. Distances were not binned prior to analysis. The overall distribution of perpendicular distances was examined for right truncation to remove outliers at great distances. Preliminary examination of detection curves suggested that some narwhals were missed close to the track line despite the bubble windows. Therefore, there was a risk that hazard-rate and half-normal distributions would overestimate the probability of detection and the resulting effective strip width. However, almost a hundred narwhal sightings were made within 100 m of the track line and therefore it seemed inappropriate to lose a large amount of data by left-truncating (i.e., discarding sightings close to the track line). The shape of the histogram suggested that, in addition to half-normal and hazard rate keys, a unimodal detection curve like a gamma distribution might provide a better fit (Quang & Lanctot 1991). Because a gamma distribution usually takes the value zero at zero distance, we fitted the gamma distribution with an offset term from 0 to 500 m. We used Akaike's Information Criterion (AIC) (Buckland et al., 2001) to select the best-fitting detection function among half normal, hazard rate, gamma and uniform models, with and without adjustment series.

AIC was also used to select among models with covariates. In addition to the environmental covariates ice cover, cloud cover, sea state and glare, we tested the impact of a "sighting rate" covariate, which was computed as a rolling average of the number of sightings made by the observer in a 30 s window prior to each sighting.

Because observers were instructed to give priority to group size estimation, some observations were lacking a perpendicular distance measurement (usually when high densities of narwhals were encountered). These observations were not included in the selection of the detection function. However, these observations were all assumed to be within truncation distances as we expect that the effective searching width was narrowed in higher densities. Therefore, these observations were included in the estimation of encounter rates and expected group size for the estimation of density and abundance.

The expected group size in each stratum was estimated using the size bias regression method of the natural log of group size, s , against the probability of detection, that is, $\ln(s)$ versus $g(x)$ with the *lm* function in R. The regression was used if significant at $\alpha = 0.10$, otherwise the mean group size was used (Buckland et al., 2001).

Mark-Recapture Distance Sampling

We used MRDS to estimate the proportion of available narwhals that were seen at distance 0, and thereby provide a correction for perception bias (Burt, Borchers, Jenkins, & Marques, 2014; Laake & Borchers, 2004). As our front and rear platforms were symmetrical, we used the Independent Observer configuration. Although observers 1 and 2 were acting independently, detection probabilities of observers can be correlated because of factors such as group size (for example, both observers are more likely to see only large groups at long distances). We therefore used "point-independence" models, which relax the independence assumption such that independence is assumed only at distance 0 (Laake & Borchers, 2004). A point-independence model requires the estimation of 2 detection functions: the MCDS detection function described above, and a MRDS detection function to estimate the probability of detection on the track line.

MRDS models were built with different combinations of covariates, fitted to the data and compared using AIC. By definition, all point-independent models included perpendicular distance as a covariate. Candidate covariates were the same as those described for the MCDS analysis, with the addition of observer platform (1 or 2). The best-fitting MRDS model yields an estimate of detection probability on the track line, $p(0)$, for both observers combined, which is used as a correction factor for the perception bias.

There were a few segments of the survey during which only 1 observer was active on either side of aircraft, due to logistic issues (e.g., recorder failure). On these segments, observations were recorded by a single platform, and accordingly they were corrected for perception bias using the $p(0)$ for the corresponding platform (1 or 2) instead of the $p(0)$ of the combined observers.

As outlined above, identification of duplicates in the HACS dataset was not a straightforward process due to a large proportion of missing measurements and highly aggregated narwhal groups. Choice of thresholds beyond which pairs of sightings would not be considered potential duplicates was somewhat arbitrary and had an effect on the resulting number of unique sightings. In turn, this has an impact on abundance estimates by affecting the MCDS estimate and the estimate of $p(0)$. To take into account the effect of choosing thresholds for duplicate identification, we performed a sensitivity analysis and quantified the resulting uncertainty. We estimated surface abundance of narwhals in all non-fjord strata for each of the sets of unique sightings obtained with threshold values of 3–7 s for time difference T (i.e., 5 possible values) and 5°–15° for declination angle difference D (i.e., 11 possible values), because these 2 covariates had the largest effect on duplicate identification. We then calculated the variance in the abundance estimates resulting from these multiple analyses. This allowed us to include an additional variance component in the surface abundance estimate with a CV equal to that of the sensitivity analysis (CV_{dup}), thus leaving the point estimates unchanged but increasing the range of uncertainty.

Availability bias

Experiments with narwhal-shaped models have shown that narwhals could be seen and identified by observers (i.e., were available) to depths of about 2 m in clear water (Richard, Hall, Dueck, & Barber, 1994) and therefore this is the depth threshold that has been used to correct for availability bias in most past narwhal surveys (e.g., Asselin & Richard, 2011).

The proportion of time P_a that narwhals spend within 2 m of the surface was estimated at $31.4 \pm 1.06\%$, based on data from 24 narwhals fitted with satellite tags near the communities of Arctic Bay and Pond Inlet every August from 2009 to 2013 (Watt, Marcoux, Asselin, Orr, & Ferguson, 2015). The correction factor for availability bias when sightings are instantaneous is given by $C_I = 1/P_a = 3.18$ (CV=3.37%). C_I is an appropriate correction factor when sightings are instantaneous (e.g., for photographic surveys). If sightings are not instantaneous, this correction factor positively biases the abundance estimate. McLaren (1961) developed a correction factor that incorporates the dive cycle of the animal and the search time of the observer. The model has 2 components: the probability that an animal is at the surface when entering the observer's view, expressed as $t_s/(t_s + t_d)$, with t_s being the time the animal can be seen at the surface and t_d the period when animals are submerged, and the probability that an animal is in a dive while entering the viewing area $t_d/(t_s + t_d)$ multiplied by the probability of surfacing within the viewing area, which Laake et al. (1997) proposed expressing as $(1 - e^{-\theta/t_d})$, where θ is the time available for an observer to see a group (i.e., "Time in View"). Therefore, for a given time-in-view θ , the correction factor for availability is given by:

$$C_M(\theta) = \frac{t_s}{t_s + t_d} + \frac{t_d \cdot [1 - e^{-\theta/t_d}]}{t_s + t_d}$$

The 24 satellite tags used for estimating P_a could not be used to estimate t_d and t_s . Instead, we used values $t_s = 43$ s and $t_d = 87$ s from Richard et al. (2010), which are based on data from 3 archival time-depth recorders (ATDRs) deployed on narwhals in Tremblay Sound in August 1999 (n=1) and in Creswell Bay in August 2000 (n=2).

To estimate the empirical "Time in View" of the HACS sightings, we examined the length of time from the initial recording of a detection (i.e., spot time) to the recording of the abeam declination angle measurement (i.e., beam time). Following the technique proposed by Richard et al. (2010), we used a weighted availability bias correction factor C_a :

$$C_a = C_I \cdot \frac{\sum_{i=1}^n f_i (1 - b_i)}{\sum_{i=1}^n f_i}$$

where n is the maximum time in view, f_i is the frequency of times in view of duration i seconds and b_i is the percent bias of an instantaneous correction C_I for i seconds:

$$b_i = \frac{C_M(0) - C_M(i)}{C_M(0)}$$

We considered that only C_I contributed to the variance of C_a and that therefore their CVs were identical.

The surface abundance estimate of each stock \hat{N}_s was then multiplied by C_a to give a total abundance estimate \hat{N}_c . The variance was calculated using the delta method (Buckland et al., 2001).

$$\text{var}(\hat{N}_c) = \hat{N}_c \cdot \left\{ \frac{\text{var}(\hat{N}_s)}{\hat{N}_s^2} + \frac{\text{var}(C_a)}{C_a^2} \right\}$$

Fjord Strata

Data in fjord strata were collected using the same protocol as described for the design-based strata and analysed separately. We used a density surface modelling framework to model spatially-referenced count data with the additional information provided by collecting distances to account for imperfect detection. Modelling proceeds in 2 steps. First, an MCDS detection function is fitted to the perpendicular distance data as described above. Second, counts are then summarised per segment (contiguous transect sections), and a generalised additive model (GAM, (Wood, 2006)) is then constructed with the per-segment counts as the response with segment areas corrected for detectability. GAMs provide a flexible class of models that include generalized linear models but extend them with the possible addition of splines to create smooth functions of covariates.

Exploratory analyses showed that the histogram of detection distances in fjords differed in shape and extent from that of the design-based strata. We therefore used a separate detection function and mark-recapture analysis using data from all fjord strata.

Transect lines were split into contiguous segments (indexed by j), which were of length l_j . The maximum segment length was such that neither the density of objects nor covariate values varied appreciably within segments: for this survey we divided the transects into segments corresponding to about 10 s of flying time or 514 m. The area of each segment j is $A_j = 2 \cdot w \cdot l_j$ (where w is the truncation distance). For each segment and observation, we extracted 2 spatial covariates: distance from the nearest shore and distance from the nearest mouth of the fjord (into the adjacent open-water stratum). Distance from the fjord mouth was calculated as the shortest path not intersecting land, therefore taking into account the complex shapes and multiple branching of some fjords.

Count per segment was then modelled as a sum of smooth functions of covariates (e.g., location, depth, distance to shore, distance to fjord mouth measured at the segment level) using a GAM. Smooth functions were modelled as splines, providing flexible curves and surfaces to describe the relationship between the covariates and the response.

The general model for the count per segment is:

$$E(n_j) = \hat{p}_j \cdot A_j \cdot \exp \left[\beta_0 + \sum_k f_k(z_{jk}) \right]$$

where z_{jk} represents the value of the k^{th} explanatory spatial variable in the j^{th} segment, and function, f_k is a smooth function of the covariate z_{jk} and β_0 is an intercept term (Hedley & Buckland, 2004). Multiplying the segment area (A_j) by the probability of detection (\hat{p}_j) gives the effective area for segment j . If there are no covariates other than distance in the detection function, then the probability of detection is constant for all segments.

After examination of the data, we decided against using a Poisson distribution (where the variance is assumed to be equal to the mean) because narwhal sightings were clearly clustered

(i.e., overdispersed) and instead we modelled counts as a negative binomial distribution (Forney, 2000).

Spatial modelling requires the abundance of groups to be predicted throughout the survey area. To do so, a grid of cells of resolution 250 m x 250 m was constructed to cover the whole area of each fjord. Cell size was arbitrary but constrained by 2 requirements: resolution had to be coarser than segment lengths, yet small enough to minimize variability of the explanatory variables within each cell. Values for the explanatory variables were calculated from the midpoint of each grid square. Extensive simulations revealed that model descriptions and predictions were robust to variation in choice of grid size (Hedley & Buckland, 2004).

Spatial model fitting

The number of narwhals seen in each segment was described by a GAM with a spatial smoother and by the effective area of each segment (i.e., the product of its effective strip width and its length). Density surface models (DSMs) are typically fitted with thin-plate regression splines (Wood, 2003). However, previous work has highlighted that in some cases, the fitted surface tends to increase unrealistically as predictions are made further away from the locations of survey effort (Miller, 2012). This problem can be alleviated using a generalization of thin plate regression splines called Duchon splines (Miller & Wood, 2014). Problems can also occur when smoothing over areas with complicated boundaries (Wood, Bravington, & Hedley, 2008). If 2 parts of the study area are linked by the model without taking into account obstacles, then some boundaries (e.g., peninsula, island) can be “smoothed across”. Therefore, we also fitted a “soap film” smoother (Wood et al., 2008), which usually performs better for complex study regions by reducing smoothing of density contours across land boundaries and minimising edge effects. The soap film is a bivariate smooth of spatial coordinates only and cannot include covariates such as distance to fjord mouth and to shore.

We fitted models with and without covariates, for each of the 3 types of spatial smoothers, in the package *dsm* (Miller, Rexstad, Burt, Bravington, & Hedley, 2013). Within each model, flexibility (estimated degrees of freedom) and removal of model terms were based on functions from the *mgcv* package (Wood, 2001), which uses restricted maximum likelihood to choose a statistically defensible degree of smoothing, with penalties for unnecessary flexibility. Then, the best model (choice of smoother and covariates) was selected based on AIC. Goodness of fit was examined with random-quartiles Q-Q plots on the residuals (Dunn & Smyth, 1996).

Variance in spatial models of abundance can be estimated by resampling through the use of moving-block bootstraps, but our survey design precluded easy identification of an independent resampling unit (Williams et al., 2006). Therefore, we used an alternative Bayesian approach that simulates replicate parameter sets from the posterior distribution of the estimated parameters of the spatial model to obtain a measure of the variance in the spatial model (Wood, 2006).

Spatial models also include variability that comes from estimating the parameters of the detection function because the effective area of each cell is based the estimated strip half-width. The total variance of abundance of each fjord (PSU) was estimated using the delta method (Seber 1982) to combine the variance of the effective area (detection model) with the

variance from the spatial component (Hedley & Buckland, 2004).

If no model could be fitted for a given fjord (i.e., no significant coefficients) or if there was only 1 sighting made, then a DSM was not used for that PSU and we used instead a “naïve” abundance estimate, equal to the number of individuals seen divided by the effort (length multiplied by effective strip width) and multiplied by fjord area. The CV of such estimates was calculated using an empirical estimate of variance. For the special case of fjords with only 1 sighting, we arbitrarily assumed a CV of 1.00.

Stratum-wide estimates of density and abundance

Within each fjord stratum, we used a 2-stage sampling design in which the first stage consisted of sampling with replacement among potential fjord candidates (PSUs). Appropriate estimators for the density \hat{D} and total surface abundance \hat{t} in a stratum were given by a ratio estimator (Cochran 1977):

$$\hat{D} = \frac{\sum_{i=1}^n \hat{y}_i}{\sum_{i=1}^n A_i}$$

$$\hat{t} = A_t \cdot \frac{\sum_{i=1}^n \hat{y}_i}{\sum_{i=1}^n A_i}$$

where A_i and \hat{y}_i are respectively the area and the estimated abundance in each of the n surveyed fjords (PSU) in the stratum, and A_t is the total stratum area. Note that this is equivalent to averaging the estimated narwhal densities of all fjords, weighted by their respective areas. Note also that when all N possible fjords in a given stratum are surveyed (as in Admiralty Inlet and Eclipse Sound), $\sum_i^n A_i = \sum_i^N A_i = A_t$, and thus the equation simplifies to the sum of the abundance estimates, as per a standard stratified design (Buckland et al., 2001).

The sample variance of the estimated density among fjords, with fjords of unequal areas, was adapted from the formula proposed by Innes et al. (2002):

$$\widehat{Var}(\hat{D}) = s_a^2 = \frac{n}{A^2(n-1)} \cdot \frac{\sum_{i=1}^n A_i^2 \left(\frac{\hat{y}_i}{A_i} - \frac{\sum_{i=1}^n \hat{y}_i}{A} \right)^2}{n-1}$$

where $A = \sum_i^n A_i$ is the sum of the areas of the surveyed fjord.

Because we used a 2-stage sampling scheme, the variance of the total abundance has 2 components: among-fjord variance and within-fjord variance. The among-fjord variance is equal to $A_t^2 s_a^2$ with the addition of a finite population correction $(1-f)$ and the within-fjord component is the sum of the variances of each surveyed fjord, multiplied by the inverse of the sampling fraction. Thus, the estimator of the variance is:

$$\widehat{Var}(\hat{t}) = A_t^2 (1-f) \cdot s_a^2 + \frac{1}{f} \cdot \sum_{i=1}^n s_i^2$$

where $f = \frac{n}{N}$ is the sampling fraction and s_i^2 is the variance of the i^{th} fjord (obtained from the DSM) and s_a^2 is the among-fjord variance of the estimated density. Note that when all the fjords within a stratum are sampled (i.e., $f = 1$), the first term disappears and the multiplier of the second term becomes unity, that is, the total variance is the sum of the within-fjord variances, as per a standard stratified design (Buckland et al., 2001).

Availability bias

Estimates were corrected for availability bias in the same manner as was done for the design-based strata. However, all narwhal sightings in the fjords of the East Baffin Island stratum were reported by observers as having occurred in murky or opaque waters, which was confirmed by examination of the photographs taken underneath the plane. This suggests that observers would not have been able to detect and identify narwhals as deep as 2 m, as was assumed for clearer waters. Therefore, for this stratum, a correction factor was calculated based on the assumption that narwhals could only be seen between 0 and 1 m. Watt et al. (2015) found that narwhals spend an estimated 20.4 ± 0.78 % of their time in the 0 to 1 m depth interval.

Abundance estimates by stock

Total abundance estimates for each stock were obtained by the addition of the estimated abundances of all the strata that were part of that stock's summer range, including results from fjord strata. Variance for the stock-wide abundance estimate was calculated by summing the encounter rate and group-size components of the variances of each stratum. Since the design-based estimates all share the same detection function, the detection component of the variance was incorporated only once in each stock estimate (but the detection function of model-based estimates was different and therefore added to the total variance for stocks that included a fjord stratum). The same approach was used to provide a total abundance estimate and variance for the sum of all 6 Canadian summer stocks in 2013. Confidence intervals were derived assuming a log-normal distribution and degrees of freedom were calculated using the Satterthwaite approximation (Buckland et al., 2001).

RESULTS

Survey coverage and narwhal sightings

The timing of the ice break-up in the northern parts of the survey range during the summer of 2013 affected the timing and coverage of portions of the survey areas. At the beginning of the survey period, several areas were still completely (Norwegian Bay, Peel Sound) or partially (Jones Sound, Barrow Strait) covered with ice. Contingency days had been planned to allow for poor weather conditions. In the end, the aircraft were able to survey for about 40% of the available time. Weather conditions deteriorated substantially towards the end of the survey period. Some areas were characterized by poor weather during the entire survey period (e.g., strong winds and thick fog in Smith Sound).

The strata believed to constitute the main aggregation areas of the putative Jones Sound and Smith Sound narwhal stocks had been given the highest level of priority (Figure 4). However, heavy ice conditions imposed delays. Norwegian Bay was flown in good weather, but its northern part and several of its fjords were still frozen. Narwhals were observed in its southern half. Jones Sound and its fjords were flown in excellent conditions in a single day although few narwhals were observed. Grise Fjord community members reported during the survey that narwhals arrived later than usual in 2013. Consequently, we attempted to fly this stratum again on the last day of the survey (August 26), although winds were stronger than desirable. From this second coverage, only effort and sightings from the more

sheltered fjord strata, combined with those from the earlier coverage, were used in the analysis. Fog and strong winds precluded complete coverage of Smith Sound. Several of the eastern Ellesmere fjords could be surveyed, however, and large numbers of narwhals and belugas were observed in Mackinson Inlet, in particular.

Strata of the Somerset Island stock were given the second highest priority ranking (Figure 6). By using all 3 aircraft simultaneously, both Peel Sound and Prince Regent Inlet were surveyed in a single day. The Gulf of Boothia was covered a week later over a 2-day period. Narwhals and bowhead whales were aggregated at the southern end of Prince Regent Inlet and in the northern part of the Gulf of Boothia. Despite heavy ice cover, numerous narwhals were observed in the central, high-density area of Peel Sound.

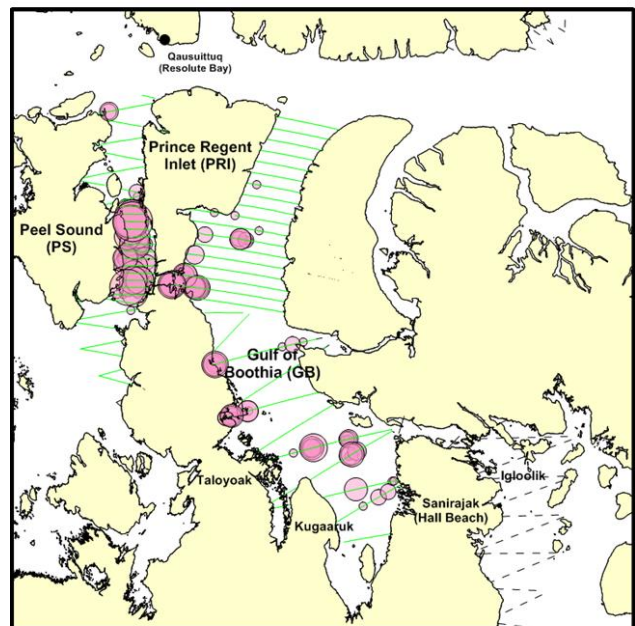


Figure 6. Somerset Island survey showing completed transects (green), un-completed planned transects (black dashed) and sightings of narwhals with symbol size proportional to group size (1–20).

We had planned to survey Admiralty Inlet and Eclipse Sound in quick succession. Admiralty Inlet was surveyed in 2 days, with a 4-day break in between due to bad weather (Figure 3). Eclipse Sound was covered immediately afterwards, in 2 successive days. Few narwhals were observed in the high intensity areas, but instead were highly aggregated in the southern ends of both areas, close to shore or within fjords.

The eastern coast of Baffin Island was surveyed by 1 aircraft over a 2-week period (Figure 5). Strong winds made it difficult to survey the offshore portion of the area and numerous attempts were necessary. In the end, about 90% of the planned transect lines were surveyed, and all but 1 of the planned fjords. Narwhals were seen predominantly in the fjords of the north-western half of the stratum. One narwhal was sighted in Cumberland Sound but was not included in any of the stock estimates due to uncertainty about its stock of origin (although it is likely it was from EBI).

Overall, there were 1,600 sightings of narwhal groups while on effort, of which 622 were made in fjord strata. After photo-

verification of sightings with missing measurements or coded as “uncertain”, 5 were missing group size information and 28 were missing a perpendicular distance measurement.

Duplicate identification

Visual and photographic data

The HACS dataset contains 2,553 on-effort visual sightings of cetaceans, of which 2,422 were made in double-platform configuration (1,428 made by the primary platform and 994 by the secondary platform). The observations include 1,600 narwhal, 224 bowhead, and 314 beluga sightings.

Of these, 431 visual sightings were missing declination angle measurements and 114 were missing group size counts. Using photographic data, we were able to recover 299 missing angles and 39 missing group sizes.

The differences between angles measured by visual observers and those determined from photographs were mostly distributed between 0 and 5°, although about 8% were between 5 and 10°, and 4% were beyond 10° (Figure 7). The majority of differences >10° corresponded to sightings that the visual observers had identified as “uncertain” measurements.

The difference between visual and photographic data sets for group size counts were tightly clustered around 0 and 1 (Figure 7), with few differences beyond 2 and almost none beyond 5. Measurements deemed “uncertain” by visual observers accounted for the majority of the cases where the difference was >3.

Covariate thresholds

There were 2,506 pairs of sightings made by same-side front and rear observers separated by 10 seconds or less, which we

considered to be potential duplicate candidates. Of these, after recovery of missing data from photographs, 2,140 pairs had data for all covariates for both sightings, but 222 were missing a declination value for at least 1 sighting in the pair, 89 were missing a count value, and 55 were missing both declination and count.

The opposite-side dataset contained 1,526 pairs of sightings made within 10 s of one another, including 1,364 with all covariate data; 80 missing a declination angle covariate, 75 missing a count covariate, and 7 missing both declination and count.

We used graphical methods to identify thresholds for duplicate identification but the inflection points observed by Southwell et al. (2002) were not as readily apparent in our data, except for the differences in counts C (Figure 8). The inflection point for time differences T was not clear but the rise in the number of duplicates (Dup) did slow at T greater than 5 seconds. The sensitivity of Dup to angle differences D appeared to lessen at D greater than 5° and slowed down further at 10°, but overall these decreases were gradual rather than abrupt. In contrast, the rise in the number of duplicates slowed rapidly at levels of count differences C greater than 3 or 4. There was little sensitivity of Dup to species differences S , especially at levels above 0.5. For the most part, these potential thresholds were not apparent in the opposite-side dataset (Figure 8, red lines), which confirmed that these inflexion points were due to the presence of true duplicates in the same-side dataset. Based on this we established conservative thresholds for duplicate identification of $T_{max}=5$ seconds, $D_{max}=10^\circ$, $C_{max}=3$ and $S_{max}=0.50$.

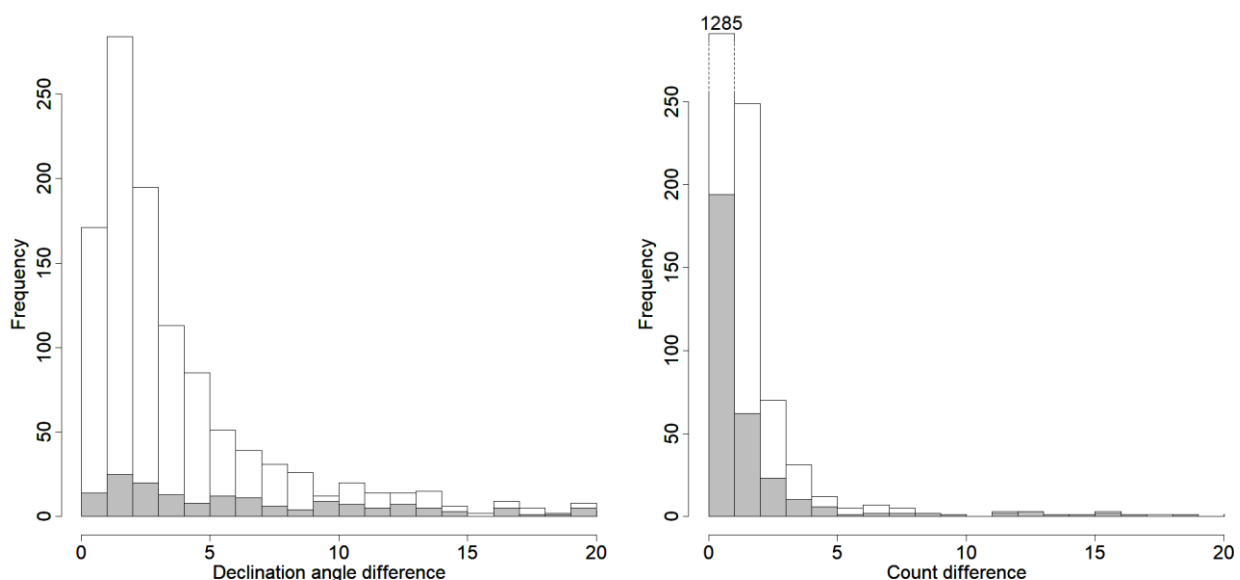


Figure 7. Distributions of the absolute differences in declination and group count between visual observations and the corresponding sightings found on photographs. White bars: all sightings. Grey bars: sightings coded as “uncertain” by visual observer for declination angle (left) or count (right).

Logistic regressions

Table 3 describes the equations of the best logistic regression models for sighting pairs with no missing data as well as each case of missing data. The magnitude of the standardized coefficients corresponds to the relative amount of variance accounted for by that coefficient, while the AUC represents the success of each model at correctly classifying pairs into each data set (same side vs. opposite). As expected, all coefficients for main effects were negative; that is, an increase in covariate differences T , D , C , or S caused a decrease in the probability that the pair was from the same side. Difference in declination (-0.56) was the most important factor in discriminating the same-side from the opposite-side cases when all covariates were available, with count differences C being slightly less useful (-0.37), and T and S of less importance. This was also the case when C was missing. When declination difference D was missing, the other 3 covariates became equally important, each explaining one third of the variance.

Using these thresholds, we identified 547 duplicates of all species combined (27% of 2,072 unique sightings) (Table 4). Bowhead whale sightings were duplicated at a rate of 50%, while narwhal and beluga whale sightings were duplicated at a rate of about 30%.

Figure 8 shows the sensitivity of the thresholds used in the logistic method on the numbers of duplicate and unique sightings for all data as well as broken down by species. Only the thresholds for time differences T and declination angle differences D are shown because they account for most of the variability (C and S have less effect). The minimum and maximum numbers of unique sightings, using extreme values of T_{max} and D_{max} were 146–148 for bowhead whales, 253–257 for beluga, and 1,316–1,335 for narwhals.

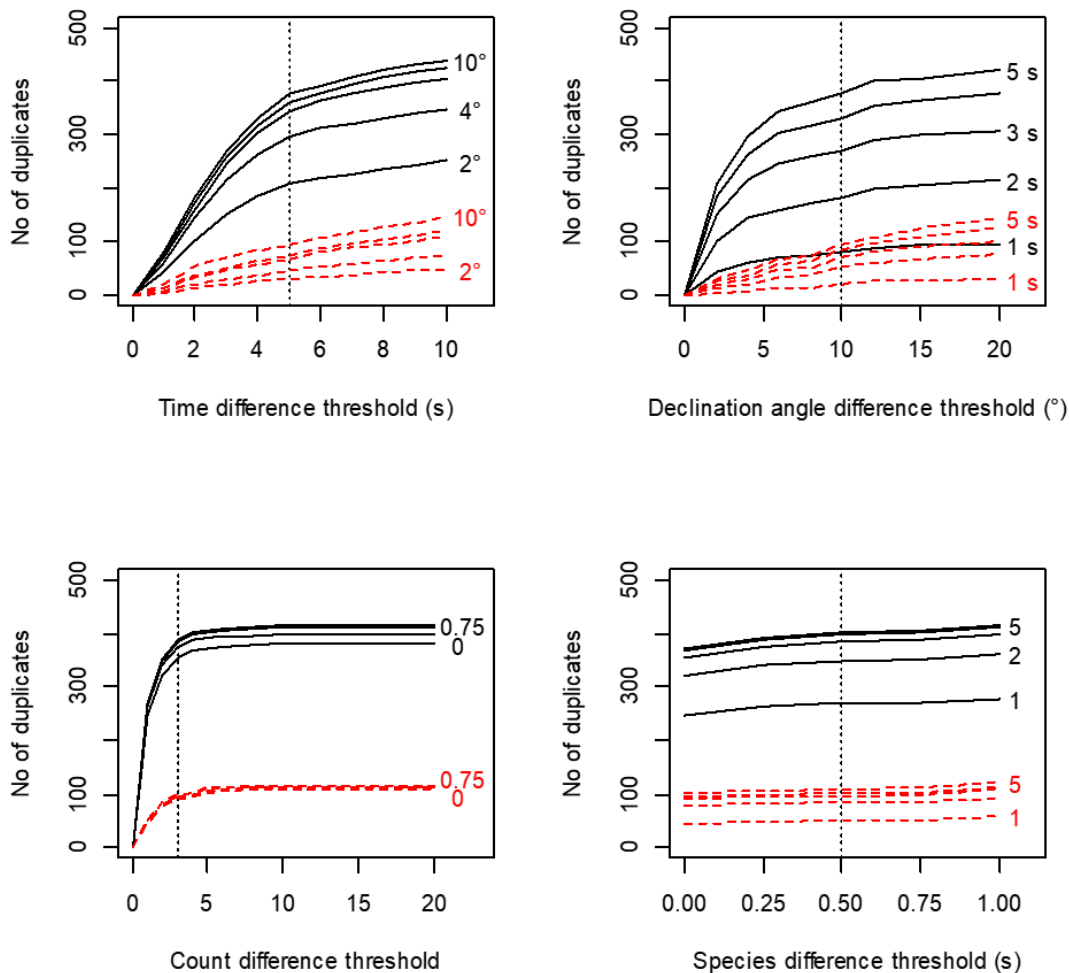


Figure 8. Sensitivity of the number of duplicate sightings to the between-observer difference in covariates. Solid black lines: same side dataset. Dashed red lines: opposite side dataset. Top: as a function of beam time difference T and declination D (each line corresponds to a different threshold value), while holding $C_{max}=3$ and $S_{max}=0.50$. Bottom: as a function of group count difference C (x-axis) and species identification S (each line corresponds to a different threshold), while holding $T_{max}=5$ s and $D_{max}=10^\circ$. Some labels omitted for clarity. Vertical dotted lines: proposed threshold value for each covariate.

Table 3. Standardized coefficients of logistic regression equations with “same” and “opposite” sides as response variable, with no missing data (full) and sets missing *D*, *C* or *D* and *C*. $P_{(T)}$: Probability of belonging to the same side dataset at threshold levels of *T*, *D*, *C*, and *S*. $P_{(0)}$: Probability of belonging to the same side dataset with all covariates set at 0. ROC(AUC): “Receiver Operating Characteristic, Area Under Curve” corresponds to the probability that a randomly chosen datum will be classified correctly.

MODEL	<i>T</i>	<i>D</i>	<i>C</i>	<i>S</i>	$P_{(T)}$	$P_{(0)}$	ROC(AUC)
Full	-0.24	-0.56	-0.37	-0.20	0.58	0.75	0.63
Missing <i>D</i>	-0.32	-	-0.33	-0.35	0.52	0.71	0.59
Missing <i>C</i>	-0.23	-0.54	-	-0.09	0.62	0.72	0.60
Missing <i>D</i> + <i>C</i>	-0.31	-	-	-0.25	0.56	0.68	0.56

Table 4. Duplicate assignments using the logistic method, for narwhals, belugas, and bowhead whales. N_p – number of groups detected by primary platform; N_s – number of groups detected by secondary platform; N – total primary and secondary detections from double platform effort; Dup – number of duplicate sightings; N_U – number of unique group sightings, i.e. $N_U = N_p + N_s - Dup$.

SPECIES	N_p	N_s	Dup	N_U	Dup %
Narwhal	1000	717	395	1322	30%
Beluga	187	174	77	284	27%
Bowhead	132	93	75	150	50%

Detection functions

Design-based strata

Right-truncation to 1,000 m reduced the number of sightings of narwhals to 762, of which 515 were seen by primary observers, 523 by secondary observers, and 276 by both. Sighting frequency was reduced within about 100 m of the track line (Figure 9a), suggesting that nearby sightings were likely missed by observers despite the use of bubble windows.

Model selection was performed on the 3 key functions and all the combinations of environmental covariates. The model with the lowest AIC was one with a truncated gamma key function (Figure 9a). Covariates “sighting rate”, “Beaufort” and “glare” were selected and had the effect of reducing detection distance at higher levels (Beaufort >3, Glare=intense, Sighting rate>10 in the last 30 seconds, Figure 10). This resulted in an average probability of detection within the truncation distance of 0.48 (CV=2.8%) and an estimated strip half-width (ESHW) of 481 m.

Selection among MR models was performed on all the combinations of environmental covariates as well as covariates “observer” and “group size”. The lowest AIC was a model with covariates “distance” and “sighting rate”. It resulted in a $p(0)$ for observers 1 and 2 of 0.58 (CV=7%), and a combined $p(0)$ of 0.82 (CV=3.4%). When the analysis was performed separately for each plane, the combined $p(0)$ for both platforms was relatively homogeneous: $p(0)$ =0.82 for plane 1 (426 sightings), $p(0)$ =0.91 for plane 2 (46 sightings), and $p(0)$ =0.83 for plane 3 (290 sightings).

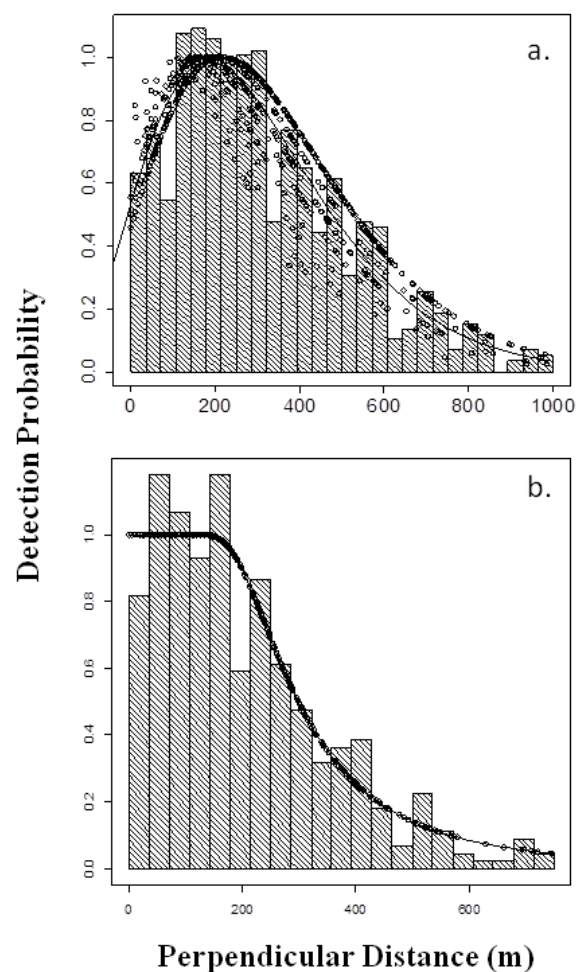


Figure 9. Detection functions for design-based (a.) and fjord (b.) strata. Circles are the probability of detection for each sighting given its perpendicular distance and other covariate values. Line is the fitted model. Probabilities are not corrected for perception bias.

When combining the effects of the decrease in detection probability with distance from the track line and of perception bias, the overall average probability of detecting a group of narwhals within 1000 m of the aircraft was 0.40. Due to variation in covariate levels across strata, the average probability of detection varies among stocks from 0.36 to 0.47.

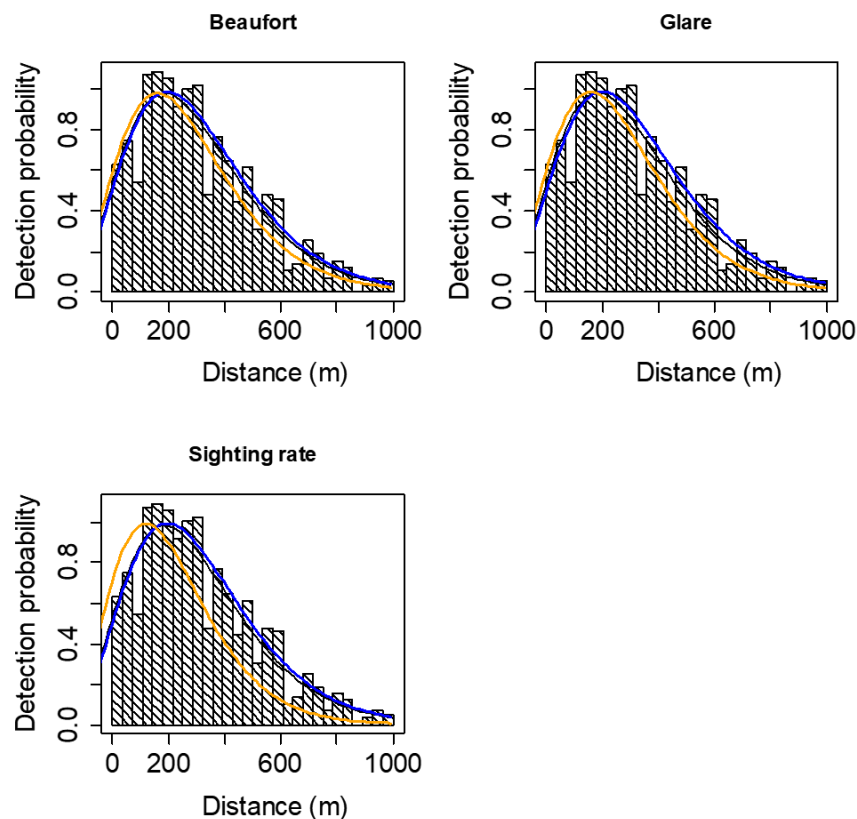


Figure 10. Predicted detection probabilities in design-based strata as a function of 3 environmental covariates in the detection function. Black lines: fitted global model. Blue line: “low” level of the covariate. Orange line: “high” level of the covariate (Beaufort >3, Glare=intense, Sighting rate>10 in the last 30 seconds).

Fjord strata

A total of 617 narwhal sightings were made in fjord strata. Of these, 76 were missing a declination angle, even after verification using photographic data. After duplicate identification, 386 sightings were seen by primary observers, 134 by secondary observers, and 97 by both, for a total of 521 unique sightings. Right truncation at 750 m removed 22 distant sightings. Left truncation did not improve the fit and therefore was not used. Best fit was achieved using a hazard-rate key function with no covariates, estimating ESHW as 343 m (CV=4.6%) (Figure 9b). The best MR model resulted in a $p(0)$ for observers 1 and 2 of 0.33 (CV=6%), and a combined $p(0)$ of 0.55 (CV=4%).

Encounter rates and group size

Encounter rates in design-based strata were highly variable (Table 5), reflecting strong differences among areas and the highly aggregated nature of narwhal groups across their summer range. There was no significant relationship between the probability of detection $g(x)$ and the natural log of group size $\ln(s)$. The global average group size was 2.76 (CV=3.8%), and stratum-wide mean group sizes ranged from 1.00 to 3.08.

Sensitivity to duplicate thresholds

As a sensitivity analysis, the abundance estimation process for the design-based strata was run multiple times with different values for the thresholds used in duplicate identification,

including fitting a new detection curve, estimating group size and encounter rates for every run. The resulting surface abundance estimates exhibited little variation, with CV ranging from 0% (Smith Sound) to 3% (Jones Sound, Admiralty Inlet) (Table 5). Therefore, for the final abundance estimates we retained the baseline threshold ($T=5$ seconds, $D=10^\circ$), and we applied a multiplier factor of 1 with a CV corresponding to the variation from the sensitivity analysis (CV_{dup}).

Spatial models in fjord strata

Spatial models were fit to the 10 fjords with more than 1 narwhal observation. Model selection for all PSU is summarized in Table 6. Duchon splines or soap films were always selected over the thin-plate regression splines. Spatial covariates were retained in 3 final models. A low degree of smoothing (i.e., high effective degrees of freedom) was often needed to fit to the high degree of clustering in narwhal sightings. Modelling attempts failed in EBF12, with no significant coefficients (spatial or covariates), and thus a naive estimate was computed. The naive approach was also used for SSF11, EBF14, and EBF36, each of which had 1 narwhal sighting.

Density surfaces for each fjord are shown in Figure 11. Sightings in AIF01 were located so close to the shoreline that a finer scale (i.e., a 100 x 100 m cell size) was needed to capture the relationship between sighting location and distance to shore. This was the fjord in which the 2 spatial covariates had the largest influence on the spatial model.

Table 5. Survey coverage, sightings, and surface abundance estimates by stratum and by summer stock. Encounter rates and group size means are given for each stratum, and average probabilities of detection are given for each stock. N_{groups} : number of unique sightings of narwhal groups. ER: Encounter rates. CV_{ER} : CV of encounter rates. $N_{\text{ind.}}$: number of individuals. GS: Mean group size. CV_{GS} : CV of group size. $p(x)$: Average probability of detection. CV_{DF} : CV of detection function (including perception bias).

Stock / Stratum	Area (km ²)	Effort (km)	Transects /PSU	N_{groups}	ER (km ⁻¹)	CV_{ER}	$N_{\text{ind.}}$	GS	CV_{GS}	$p(x)$	CV_{DF}	Surface abund.	CV
Jones Sound													
JS	19,231	930	13	16	0.017	1.08	38	2.38	0.15	0.39	0.04	940	1.08
NB	13,713	635	9	41	0.065	0.30	113	2.76	0.14	0.39	0.04	3,331	0.29
total (design-based)			22	57	0.036	0.51	151	2.65	0.11	0.39	0.04	4,271	0.33
JSF	2,413	888	6	24								45	0.94
Total	35,357											4,316	0.32
Smith Sound													
SS	10,861	546	11	76	0.139	0.98	171	2.24	0.07	0.46	0.04	3,647	0.97
SSF	4,237	413	4	165								1,916	0.45
Total	15,098											5,563	0.65
Somerset Island													
GB	63,178	1,627	11	81	0.050	0.26	158	1.95	0.09	0.42	0.04	7,335	0.28
PRI	29,178	1,888	18	88	0.047	0.58	208	2.36	0.08	0.42	0.04	3,627	0.63
PSHI	5,454	527	13	165	0.313	0.18	456	2.76	0.08	0.42	0.04	5,781	0.21
PSLO	17,499	860	16	4	0.005	0.39	8	2.00	0.20	0.42	0.04	179	0.38
Total	115,309		58	338	0.069	0.17	829	2.45	0.05	0.42	0.04	16,921	0.20
Admiralty Inlet													
AIH	3,981	500	16	26	0.052	0.64	61	2.35	0.18	0.36	0.04	535	0.64
AIL	4,526	387	18	220	0.568	0.34	678	3.08	0.09	0.36	0.04	11,237	0.45
total (design-based)			34	246	0.277	0.31	739	3.00	0.08	0.36	0.04	11,772	0.43
AIF	912	456	7	132								143	0.85
Total	9,419											11,915	0.42
Eclipse Sound													
ESH	2,839	460	14	2	0.004	0.77	2	1.00	0.00	0.42	0.04	16	0.77
ESL	4,334	310	29	68	0.220	0.33	142	2.09	0.11	0.42	0.04	2,415	0.34
total (design-based)			43	70	0.091	0.32	144	2.06	0.11	0.42	0.04	2,431	0.34
ESF	1,286	447	7	673								1,135	0.19
Total	8,459											3,566	0.24
East Baffin Island													
EB	43,419	1,140	28	3	0.003	0.63	3	1.00	0.00	0.47	0.04	122	0.63
EBF	10,091	1358	9	773								3,799	0.35
Total	53,510											3,921	0.34

Table 6. Spatial density models. For each surveyed fjord in which narwhals were sighted, the best spatial density model is shown. n_{groups} : number of unique narwhal groups sighted; n_{ind} : total number of individuals; \hat{N}_{naive} : naïve abundance estimate (no spatial model); xy smoother: best selected smoother among thin plate regression splines, Duchon splines and soap film (with effective degrees of freedom); covariates: distance to mouth and distance to shore (with effective degrees of freedom); Dev.: Deviance explained; \hat{N}_{dsm} : abundance estimate from spatial density model. CV has 2 components: distance detection function (ddf) and density spatial model (dsm). Note that when there was only 1 sighting in a fjord (SSF11, EBF14, EBF36) or when a spatial model with significant coefficients could not be fitted (EBF12), the naïve estimate was used.

Stratum	Area (km ²)	Effort (km)	PSU	n_{groups}	n_{ind}	\hat{N}_{naive}	Spatial model (edf)		Dev. (%)	\hat{N}_{dsm}	CV (%)		
							xy smoother	covariates			ddf	dsm	total
JSF	214	194	11	2	6	9	soap (1.6)	-	86.3	9	4.6	4.7	6.5
	219	122	12	5	13	34	duchon (4.7)	-	78.9	21	4.6	11.6	12.5
SSF	931	221	3	62	142	871	duchon (4.3)	dmouth (1.3) + dshore (1.0)	46.9	812	4.6	36.0	36.3
	88	31	8	3	6	21	duchon (0.96)	-	23.1	20	4.6	83.7	83.8
	416	98	11	1	1	6	-	-	-	-	4.6	100	100
AIF	97	61	1	32	124	286	duchon (0.32)	dmouth (1.58) + dshore (2.05)	96.2	143	4.6	84.5	84.6
ESF	161	96	1	111	1,411	3,438	soap (2.21)	-	68.8	1,135	4.6	18.8	19.3
EBF	151	63	3	237	544	1929	duchon (15.9)	-	78.6	969	4.6	17.9	18.5
	907	202	9	50	76	498	duchon (12.8)	dmouth (1.75) + dshore (1.26)	74.6	271	4.6	22.9	23.4
	738	205	10	51	105	550	duchon (8.75)	-	89.9	448	4.6	32.9	33.2
	577	162	12	2	2	10	-	-	-	-	4.6	100	100
	1247	288	14	1	2	13	-	-	-	-	4.6	100	100
	182	166	36	1	1	2	-	-	-	-	4.6	100	100

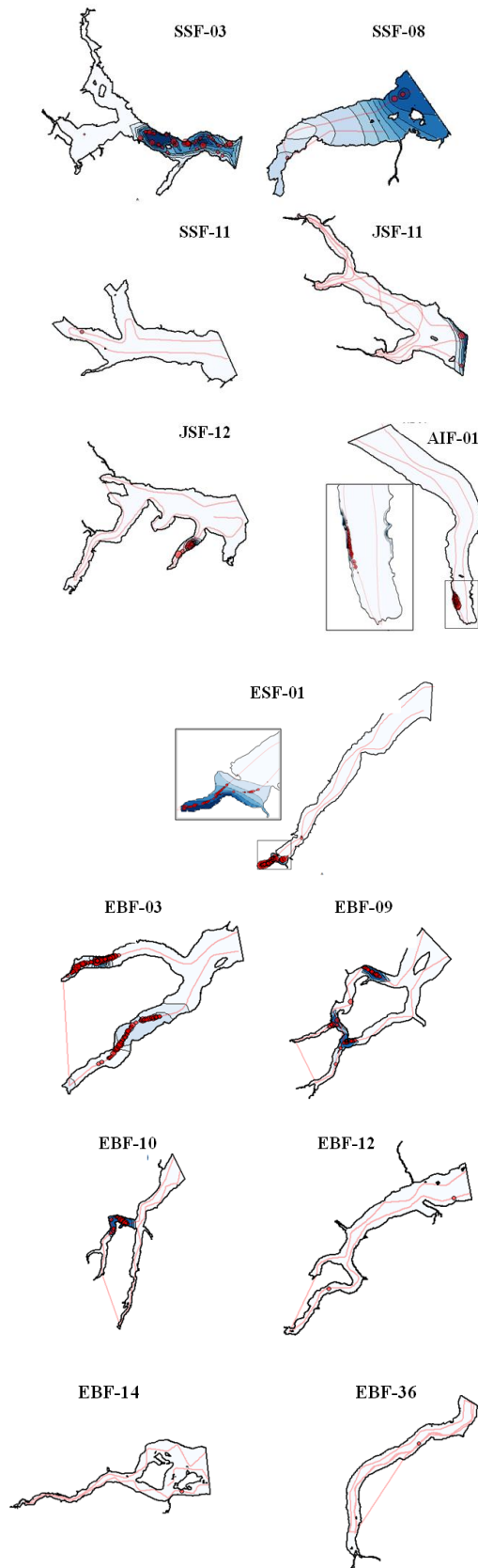


Figure 11. Spatial density surfaces of narwhal abundance in fjords. Only fjords with sightings are shown. Fjords are not on the same scale. Red line: track of aircraft. Red circles: sightings of narwhal groups. Darker shading indicates higher predicted density (different relative scales for different fjords). See Table 6 for model descriptions and estimates.

Availability and time-in-view correction

There were 527 narwhal sightings for which both a spot time and a beam time were available. Time in view ranged from 0 to 19 seconds, with an average time of 4.3 seconds (Figure 12). When applied to the instantaneous correction factors of 3.18 (narwhals assumed to be visible to a depth of 2 m) and 4.90 (assumed visible to 1 m depth), the resulting weighted availability correction factor C_a was equal to 2.94 (CV=3.4%) for the 0–2 m bin and 4.53 (CV=3.8%) for the 0–1 m bin. Based on observers’ assessment of water turbidity and examination of photographs, the 0–1 m correction was applied to East Baffin Island fjords, while all other strata used the 0–2 m correction.

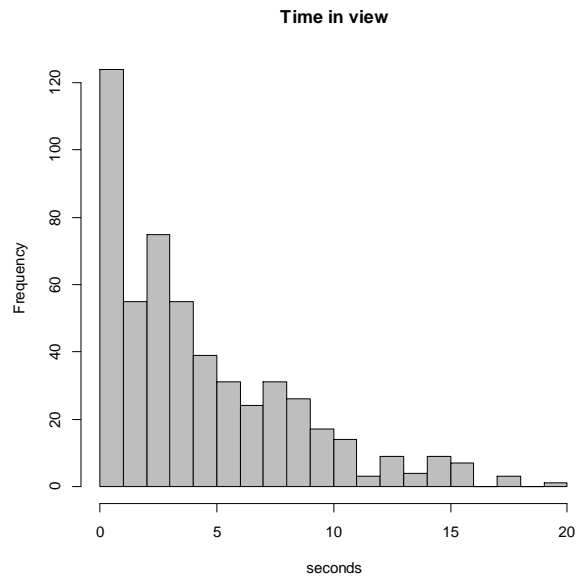


Figure 12. Histogram of time in view (i.e., difference between “spot time” and “beam time”) for 527 narwhal sightings.

Abundance estimates by stratum and stock

Fjord strata

Overall, DSM estimates were similar or slightly lower than naive estimates, except in ESF01 and EBF03 where there was a large difference between the 2 estimates (Table 6). In both cases, large aggregations of narwhals in narrow passages where the effective strip width overlapped significantly with the shoreline (Figure 11) caused the naive approach to overestimate the local density of narwhals. Moreover, in ESF01, the track overlapped itself when the aircraft turned around at the end of the fjord, resulting in double-coverage that was taken into account by the DSM but not by the naive approach.

The CVs around within-fjord abundance estimates ranged from 6.5% to 84% (not counting the fjords for which the naive CV of 100% was used), with higher CVs prevalent in fjords with few observations or where observations were clustered tightly. Jones Sound fjords that were surveyed twice were modelled once using the total effort and data of the 2 surveys (i.e. the resulting estimates reflect the average density weighted by effort).

Surveyed fjord strata estimates were applied to unsurveyed fjords (Table 5). Uncorrected abundance estimates for fjord strata ranged from 45 for JSF (CV=94%) to 3,799 (CV=35%) for EBF. Because no observations were made in WEF, no

abundance estimate was produced. Correction for perception bias based on the MRDS model would increase these estimates by an additional factor of 1.82; however, due to the challenges in identifying duplicates in fjord strata discussed below, this correction factor was not applied.

Design-based strata and summer stocks

Corrections for perception bias, availability bias and their associated variance, as well as that for sensitivity to duplicate thresholds, were applied at the stock level (Table 7). Note that because the coverage of the SS stratum was incomplete, we reduced the area over which density was calculated and abundance was extrapolated to coincide with the surveyed portion of the stratum (Figure 4). After adding fjord estimates to design-based strata, fully corrected abundance estimates were 12,694 (95% CI: 6,324–25,481) for the Jones Sound stock, 16,360 (95% CI: 3,833–69,836) for the Smith Sound stock, 49,768 (95% CI: 32,945–75,182) for the Somerset Island stock, 35,043 (95% CI: 14,188–86,553) for the Admiralty Inlet stock, 10,489 (95% CI: 6,342–17,347) for the Eclipse Sound stock, and 17,555 (95% CI: 8,473–36,373) for the East Baffin Island stock. Total abundance for these 6 stocks was estimated at 141,908 (95% CI: 102,464–196,536).

DISCUSSION AND CONCLUSIONS

This analysis provides updated, corrected abundance estimates for 6 narwhal summer stocks, which together comprise most of the narwhals found in Baffin Bay. It is the first time these stocks have been surveyed in the same summer, and also the first time that abundance estimates are available for the putative Jones Sound and Smith Sound summer stocks. These estimates result from the combination of a large-scale survey effort that mobilized resources from multiple DFO regions and Nunavut co-management partners, as well as information from concurrent projects such as satellite tracking studies operating in the same area and at the same time of year.

Potential biases

Coverage and timing

Accurate abundance estimates require that all individuals have a possibility of being sampled (Buckland et al., 2001), which implies that their entire distribution range must be surveyed. Using traditional knowledge and telemetry studies, we have tried to sample the entire areas known to be used by Admiralty Inlet, Eclipse Sound and East Baffin Island narwhals. However, complete coverage could not be achieved for the Somerset Island stock, which is believed to range farther west and south than it was possible to survey, and for the Smith Sound and Jones Sound stocks, for which exact distribution ranges are unknown. Instead we have focused on intensive coverage of the known core areas of these summer stocks.

Despite being one of the priority areas, Smith Sound could not be surveyed completely because of unfavourable weather conditions persisting throughout the survey period. We chose not to extrapolate the estimated density from the relatively small contiguous area that was surveyed to the rest of the stratum, as we have no prior knowledge of narwhal distribution in this area. The resulting estimate should be considered a conservative estimate for the entire stratum.

The Barrow Strait and Lancaster Sound strata, forming part of the Somerset Island summer stock, could not be surveyed as planned due to weather conditions. However, groups of narwhals were sighted during training flights in early August in Barrow Strait, and 1 narwhal was seen during an aborted attempt to complete transects in Lancaster Sound in mid-August. Based on previous studies, we expected low densities in these areas (Innes et al., 2002) and they were to be covered with relatively low effort intensity (Figure 2). We would therefore expect the resulting negative bias on the estimate for this large summer stock to be inconsequential. Similarly, we were unable to cover the outer part of Cumberland Sound, but again low densities were expected in this area during the summer months, as it is more likely a migration corridor in the spring and fall (Heide-Jørgensen, Richard et al., 2013).

Table 7. Corrected abundance estimates by summer stock. C_a : Correction factor for availability bias (from Watts et al., 2015). CV_{dup} : CV of duplicate identification uncertainty. The 95% CI assume a log-normal distribution.

Stock	Surface abundance	CV	C_a	CV_{ca}	CV_{dup}	Abundance (corrected)	CV	95% CI	
Jones Sound	4,316	0.32	2.94	0.03	0.03	12,694	0.33	6,324	– 25,481
Smith Sound	5,563	0.65	2.94	0.03	0.00	16,360	0.65	3,833	– 69,836
Somerset Island	16,921	0.20	2.94	0.03	0.01	49,768	0.20	32,945	– 75,182
Admiralty Inlet	11,915	0.42	2.94	0.03	0.02	35,043	0.42	14,188	– 86,553
Eclipse Sound	3,566	0.24	2.94	0.03	0.03	10,489	0.24	6,342	– 17,347
East Baffin Island	3,921	0.34	2.94/4.53*	0.03/0.04*	0.01	17,555	0.35	8,473	– 36,373
Total						141,908	0.16	102,464	– 196,536

* For East Baffin Island, the first correction applies to the design-based stratum (EB) and the second one applies to the fjord stratum (EBF).

Based on information received from residents of Grise Fjord and a reconnaissance survey conducted in 2012, we expected to find narwhals in the main part of Jones Sound. However, few narwhals were seen in the Jones Sound stratum itself, with Norwegian Bay accounting for most of the estimate. Grise Fjord residents reported that narwhals arrived late to the area in 2013, but we were unable to repeat the coverage in acceptable weather conditions later in the survey period. Future tagging projects will likely increase our knowledge of narwhal movements and stock structure in this area.

We have no information on movements between fjords and open-water strata that may have occurred during the survey, for any of the stocks. Most fjords, however, were surveyed in the same day as the neighbouring open-water areas, and we have no reason to believe that directional movements into or away from fjords may have biased our estimates. Telemetry studies have confirmed that narwhals are largely sedentary within fjords during the first 3 weeks of August (Watt et al., 2012).

Perception bias correction

We estimated the proportion of narwhals that were missed by observers using MRDS methods applied to the double platform data from our independent observer configuration. MRDS requires the certain identification of sightings seen by both observers (duplicate sightings), however, there is no means to independently and unequivocally determine whether or not a given pair of sightings is in fact a duplicate pair, or to select the most likely duplicate among a set of candidate sightings observed in close proximity. The identification of duplicate sightings becomes much less certain when sightings are aggregated, as a sighting within an aggregation may have several candidate duplicate sightings from the other platform, requiring a robust method to select the best candidate.

We used a data-driven approach to inform the choice of covariate thresholds and to select the most likely duplicate pairs among different candidates. We attempted to identify covariate thresholds using the graphical methods described by Southwell et al. (2002). Unfortunately, inflection points in our data were not nearly as clear-cut as in the dataset used by Southwell et al. (2002). This is probably because our sightings were more aggregated, which leads to greater uncertainty in measures of beam time, declination, group size and species identification. To alleviate some of the resulting uncertainty, we have deviated Southwell et al. (2002) in two ways. First, the Southwell approach places a “hard” threshold on each covariate. If a sighting pair exceeds this threshold for any of the 4 covariates, it is ruled out as a valid duplicate candidate. Our logistic approach, however, used a single threshold level obtained by injecting the 4 covariate thresholds in the logistic equation. Therefore, if a sighting pair slightly exceeds a threshold for 1 covariate, it can still be considered a valid candidate provided it scores low enough on the other covariates. This allows for additional flexibility and considerably lessens the impact of choosing arbitrary thresholds. Table 4 shows that the resulting number of duplicate and unique sightings under the logistic method is relatively robust to the choice of thresholds.

Second, we used a comparison of opposite vs same-side datasets to identify the relative contribution of each covariate to best discriminate between the 2 datasets, assuming that this

would also yield the best weighted combination of covariates to identify true duplicates. However, our approach is not “fully” probabilistic (i.e., it does not provide an actual probability of a pair of sightings being a duplicate). A further methodological development has been proposed by Hamilton et al. (2018), who also used a comparison of opposite and same-side datasets to fit a mixture distribution fitted by maximum likelihood to estimate a probability that a same-side pair is a duplicate. These probabilities were then used in a bootstrap of the entire MRDS analysis to ensure that the uncertainty around duplicates is fully propagated into the abundance estimates.

The complex stratification structure of our HACS study design, and the inclusion of spatial models in addition to design-based analyses, did not allow us to implement a fully probabilistic resampling approach. Instead, we quantified the uncertainty due to duplicate identification by using a sensitivity analysis in which the entire abundance estimation process (except for fjord strata) was repeated over a range of threshold values for duplicate assignment. The results show that our method to assign duplicates was robust to reasonable variations in the threshold values (difference in beam time and in declination angle), and that the effect on the abundance estimates of each stratum was small (an additional CV component of 1–3%). We included this additional uncertainty in our final estimates.

In the design-based strata, our realized probability of detection on the track line ($p(0)$) of 0.82 for the combined platforms is similar to those estimated for previous surveys of narwhals. Asselin & Richard (2011) realized $p(0)$ values of 0.90 and 0.77 for 2 surveys conducted in Admiralty Inlet in 2010, while Richard et al. (2010) found that $p(0)$ ranged between 0.84 and 0.92 in Canadian Arctic surveys conducted between 2002 and 2004. The magnitude of perception bias depends on several factors, including the aptitude of observers, environmental conditions, and the degree of aggregation of sightings, which can vary substantially between surveys, so some degree of variation in this correction is to be expected.

Duplicate identification in the fjord strata suggested that only 23% of groups were seen by both observers, which resulted in a combined $p(0)$ of 0.55. While this low value is plausible and within the range of other narwhal surveys, it is considerably lower than $p(0)$ values for the design-based strata, which were found to vary little among aircraft and strata. It is also noteworthy that the effective strip width was lower in fjords than in design-based strata. Since there were no major differences in detection probabilities among the 3 aircraft, and considering the significant effect of the sighting rate covariate on the detection function (with higher densities having the effect of reducing detection distances), it is likely that high aggregation rates have contributed to lowering the detection distances in fjords.

Examination of sighting data in fjords showed that duplicates were restricted to group sizes of 1–5, while any group size above 5 never resulted in being considered a duplicate (except for 1 group of 332), suggesting that larger groups were more likely to be missed by at least 1 observer (including numerous groups of 25, 50 and even 200 individuals). This negative relationship with group size is counter-intuitive and likely stems from the issue that large groups were recorded differently by different observers, some of which tend to group sightings into larger clusters when faced with large aggregations while others split them into smaller groups. This level of discrepancy between

platforms makes duplicate identification difficult in high density areas. Another problem is that in areas of high aggregation, the observers missed many declination measurements, leaving only beam time and group size to identify duplicates, whereas declination was by far the strongest discriminator of duplicates. Because the extreme clustering observed in several fjords made duplicate identification uncertain, we decided against correcting for perception bias in fjord strata as there was a risk that applying the correction of 1.82 to the double-platform data (i.e., unique sightings of both observers combined) would not be precautionary. However, our estimate is affected by 2 opposing biases: a negative bias because it is not corrected for perception, and a positive bias because duplicates were likely under-identified (which would result in over-estimating the number of unique sightings that are used to calculate encounter rates). The relative magnitude of these biases can be inferred by comparing 2 alternative options. Using only primary sightings (i.e., assuming full duplication) would decrease the abundance estimate by 8%. Conversely, assuming that there were no duplicate sightings and that all detections were unique would result in an abundance estimate 20% greater. These relatively small differences are due to primary observers accounting for most of the detections in the fjords. In contrast, the correction factors for perception bias in the design-based strata (which are based on more data and showed little variability across aircraft and strata) were 1.72 for single platforms, and 1.22 for the combined observers' data. This suggests that the magnitude of the negative perception bias likely exceeds the positive bias caused by encounter rate inflation, and thus that our estimates for fjords are negatively biased overall.

In such situations of extreme clustering, the use of photographic data would likely decrease the uncertainty. High-resolution imagery collected by aircraft or unmanned aerial vehicles would allow automating the process of counting narwhals, thereby incorporating more objective and quantifiable results while alleviating perception bias. These new methods introduce other considerations and new assumptions that require testing (Brack et al., 2018), but the results of UAV-based photographic surveys of marine mammals have been shown to be comparable to those using visual sightings (Bröker, Hansen, Leonard, Koski, & Heide-Jørgensen, 2019). A further advantage of using UAVs would be to decrease risks of flying in remote areas for the pilots and observers. At the time that HACS was conducted, the existing technology (fuel endurance, weather limitations) did not make it logistically feasible to fly such a survey in the High Arctic.

Availability

Our correction for narwhals that were not seen because they were submerged during the passage of the aircraft (availability bias) is based on the diving behaviour during daylight hours of 24 narwhals tagged with satellite-linked time/depth recorders between 2009 and 2012 in Admiralty Inlet and Eclipse Sound (Watt et al., 2015). As we lack comparable data for other strata, particularly areas with sea ice, we assumed that narwhal diving behaviour does not vary significantly across the survey areas. We assumed that narwhals were visible to a depth of 2 m in areas with clear water, based on research conducted by Richard et al. (1994) using narwhal-shaped and coloured targets suspended at various depths. East Baffin fjords had highly-coloured water due to glacial inflow, and for these we assumed that narwhals were visible to a maximum depth of 1 m.

Our correction implicitly assumes an instantaneous change in availability with depth; that is, narwhals are always visible to a depth of 2 m, and never visible at greater depths. In all likelihood the relationship between availability and depth is more complex, with a gradual decrease in sightability as the animal is deeper in the water column. However, we lack data with which to quantify this relationship. Given the magnitude of the correction (multiplying by 3 to 4.5), future research is required to assess the limitations of these assumptions and to better quantify the uncertainty around these multipliers.

Narwhal diving behaviour may vary with bathymetry and slope, however bathymetry data in the survey area are too coarse to incorporate in our modelling (Watt et al., 2015). Narwhals spent slightly more time near the surface in late August compared to mid-August, however only 10 narwhal sightings were made in late August and therefore we decided to apply the mid-August correction to all data.

Movement between strata

While narwhals appear to be largely sedentary within their stock areas during the summer (Dietz et al., 2001; Dietz et al., 2008), some movement between areas has been observed, particularly late in the season (Kenyon, 2017). In addition, while narwhals usually return to the same summering area year after year, some limited inter-annual exchange between summering areas has been detected (Watt et al., 2012; Heide-Jørgensen, Richard et al., 2013; Kenyon, 2017). The latter issue was addressed by surveying the entire Baffin Bay stock area within 1 season, rather than surveying individual components in different years as has been done previously. Within-season movements between summering areas, which also serve as our survey strata, could cause bias if they are correlated with survey timing. For example, a directed movement between strata could cause positive bias if it coincided with the temporal order in which the strata were surveyed, or negative bias in the opposite case. For this reason, we attempted to survey individual strata within a short period of time, and to survey neighbouring strata within as short a period as feasible, in order to limit the possibility for inter-stratum movement. For example, the Admiralty Inlet and Eclipse Sound strata, which tagging studies have shown not to be completely isolated, were surveyed in quick succession over a period of 7 days. The magnitude of any bias resulting from inter-stratum movement is therefore likely to be small.

Spatial model uncertainty

Fjords could not be surveyed using a systematic design because they are narrow and high-sided, restricting the flight path of a low-flying survey aircraft. The non-systematic placement of transect lines in those areas could result in 2 violations of the assumptions of distance sampling: animals may not be distributed uniformly within the searched strip-width (e.g., if they are influenced by the distance to shore), and the spatial coverage within fjord strata is unequal and non-random.

To investigate the issue of non-uniform distribution of perpendicular distances within strips, we used a distinct detection curve for the fjord strata than the design-based strata. We did not observe any obvious signs of such an issue in the histogram of perpendicular distances within fjords, for which the best model was a function with monotonic decrease rather than a gamma curve (unlike the detection curve in the design-based strata). We hypothesize that there was enough

variety in fjord width, in the distance between the aircraft and the shore, and enough opportunities to run improvised zigzags to alleviate any potentially non-uniform distribution of sightings with respect to the shore.

To address the issue of non-random placement of lines, we used density surface modelling, which allows abundance to be estimated for any subset of a survey region, by numerically integrating under the relevant section of the fitted density surface. A spatial model does not require track lines to be designed according to a formal survey sampling scheme, and is robust to violations of assumptions of random and equal coverage (Williams, Hedley, & Hammond, 2006). Moreover, the resulting variance of the abundance estimate incorporates both the variance from the detection function and that of the spatial model (Hedley & Buckland, 2004). This approach allowed us to take into account the uneven, non-random transect design of our fjord surveys, and also to estimate a CV for each abundance estimate.

Spatial models, however, have their own sources of uncertainty. In fjords (PSUs) with few observations, it was often difficult to select the right set of smoothers and their level of flexibility, and the resulting abundance estimates were therefore imprecise and sensitive to modelling decisions. Even in some fjords with larger narwhal numbers, the goodness-of-fit of spatial models was not always satisfactory, suggesting there was unmodelled spatial heterogeneity due to covariates that were not included (e.g., AIF01 where 143 sightings still yielded a CV of 0.85).

We assumed that larger fjords were more likely to contain narwhals based simply on their area. An alternative approach would have been to stratify or consider each fjord as a unit of equal value (i.e., probability of containing narwhals), regardless of its size. Ultimately, we do not know what characteristics of fjords are attractive to narwhals (area, length, coastline complexity, water temperature at the surface or at depth, etc.). There could also be latitudinal or longitudinal gradients in narwhal density among fjords but we did not have enough data to model narwhal distribution at a larger (among-fjord) scale. These assumptions could introduce biases in the estimates. It is possible that detailed analyses of narwhal movement among fjords, using satellite tracking, could yield insights on this issue, but such data are missing at present for most fjord strata (Smith Sound, Jones Sound, East Baffin Island).

As in design-based approaches, the main source of uncertainty (and the reason for the relatively large CVs that characterize the total stratum estimates) are the high aggregation rates at 2 scales: within, and more importantly, among fjords, as well as the relatively small effort in terms of the number of fjords surveyed, even though the total area covered was in most cases comparable to that of regular strata (e.g., in EBF, 9 out of 54 fjords were surveyed, but their total area was equal to 48% of the stratum area). These relatively large CVs reflect the imprecision in our estimates and their sensitivity to the random selection of PSUs: for instance, if the first PSU of AIF (or ESF) had not been surveyed, the abundance estimate for the entire stratum would be zero instead of 143 (or 1,135 for ESF).

Comparison to previous estimates

Our corrected abundance estimate of ca. 50,000 for the Somerset Island summer stock is higher than the most recent previous estimate of 35,000 based on surveys done in 1996, 2002, and 2004 (Table 1), however the confidence intervals of

both estimates overlap. The estimates become more similar if the higher estimate for Prince Regent Inlet from the 1996 survey (Innes et al., 2002) is used, in which case the previous estimate sums to ca. 52,000. Taken stratum-by-stratum as well as overall, our 2013 estimate is more precise (CV=20%), likely due to increasing coverage intensity in PS and PRI, and based on more complete coverage than any previous survey of this stock.

Our 2013 results for Admiralty Inlet (ca. 35,000) and Eclipse Sound (ca. 10,000) differ substantially from the most recent previous survey estimates of the same stocks (18,000 for Admiralty Inlet in 2010 and 20,000 for Eclipse Sound in 2004) but again the confidence intervals overlap. In this survey, narwhal sightings in these 2 areas were characterized by extremely high clustering and were encountered almost entirely in the low-intensity strata, while very few sightings were made in the high-intensity strata of both regions. These factors decreased the precision of our estimate for these strata. The 2010 Admiralty Inlet result was obtained by averaging 2 complete surveys that yielded abundance estimates (24,398 and 13,729), which, while not significantly different, illustrate the magnitude of difference that can occur even in the same year. This could be due partially to movement in and out of the surveyed area, but also to differences in aggregation patterns between surveys (Asselin & Richard, 2011). We note that the total estimate for these 2 areas (~45,000) is similar to the total of previous abundance estimates, although previous surveys were never conducted in the same year. Further research is required to assess connectivity between these 2 stocks and is particularly relevant given the industrial activity and increased shipping occurring in the region.

The surface estimate of ca. 3,800 narwhals for East Baffin Island fjords is close to the 2003 estimate of 3,487 (Richard et al., 2010), despite different statistical approaches. The HACS results confirm that a large number of narwhals use the East Baffin fjords during summer, especially in the northern part. Once corrected for availability bias, our 2013 estimate of ca. 17,500 is larger than the corrected estimate from 2003 (10,000) primarily because of our use of an availability correction assuming sightability to a depth of 1 m, rather than to 2 m as used by Richard et al. (2010).

The Jones Sound and Smith Sound estimates are new and cannot be compared to previous surveys. The HACS results confirm that relatively large numbers of narwhals inhabit Jones Sound and the adjacent Norwegian Bay. However, few narwhals were seen in the Jones Sound stratum itself, with Norwegian Bay making up for most of the stock's estimate. Our coverage of Smith Sound was incomplete but resulted in a high encounter rate and a large abundance estimate. The CV in Smith Sound is quite high (0.65) despite the large number of sightings because a large proportion of the sightings were made on a single transect (the line leading into SSF-03, where large numbers of narwhals were sighted). The separation between fjord and design-based strata was mostly arbitrary and thus if the sightings at the mouth of fjord SSF-03 were wrongfully attributed to the SS estimate, it could have increased the estimate due to the large area of this strata. To mitigate this possible effect, we post-stratified SS based on the realized effort (Figure 4). Future telemetry projects will likely increase our knowledge of narwhal movements within and between these areas, as well as stock structure.

Our total estimate of 142,000 narwhals for the Canadian portion of the High Arctic region, is higher than any previous estimate. Estimates from surveys conducted between 2002 and 2004 totalled 66,000; however these surveys did not include Peel, Smith, or Jones Sounds (Richard et al., 2010). Restricting our 2013 estimates to those strata covered in 2002–2004, reduces the total to ca. 95,000, still higher than the earlier estimate. Summing all most recent estimates produced prior to 2013 results in a total of ca. 84,000 (Table 1), but this does not include Smith or Jones Sound. Our 2013 survey was more extensive and intensive than any conducted previously, and the resultant estimates do not suffer from the potential biases resulting from surveying a non-stationary stock piecemeal over several years. We therefore consider the HACS 2013 estimate to be the best available information for Canadian Baffin Bay narwhals. Estimates for Inglefield Bredning (8,368 in 2007, Heide-Jørgensen et al., 2010) and Melville Bay (3,091; Heide-Jørgensen, Hansen, & Strandler Sinding, 2014) can be added to approximate the total number in the Baffin Bay Stock of about 153,000 narwhals. Estimates from other stocks of narwhals include: East Greenland: <1,000 (Heide-Jørgensen et al., 2010) and Northern Hudson Bay: 4,452 in 2011 (Asselin et al., 2012). Narwhal are also found around Svalbard and the northwest Russian Arctic, but numbers here appear to be small (Gjertz, 1991; Lydersen, Martin, Gjertz, & Kovacs, 2007). The 6 High Arctic Canadian stocks may therefore represent up to 95% of the world's narwhal population.

Overall, our objectives of updating abundance estimates of narwhal stocks in the Canadian High Arctic and improving their precision and accuracy were met. Concurrent, long-term telemetry studies of diving behaviour were critical to obtaining estimates of availability bias. Abundance estimates also were improved by implementing new analysis techniques to address specific challenges associated with narwhal use of fjords. The success of HACS was also due to involvement of the Inuit communities and co-management partners, including their participation in a 2012 reconnaissance survey of previously unsurveyed areas.

ADHERENCE TO ANIMAL WELFARE PROTOCOLS

The research presented in this article has been done in accordance with the institutional and national animal welfare laws and protocols applicable in the jurisdictions in which the work was conducted.

ACKNOWLEDGEMENTS

We thank the communities and Hunters' and Trappers' Organizations of Arctic Bay, Clyde River, Grise Fjord, Hall Beach, Kugaaruk, Pangnirtung, Pond Inlet, Qikiqtarjuaq, Resolute Bay, and Taloyoak. Christine Abraham and Mike Hammill assisted with project management, Pierre Richard with survey design, and Blair Dunn, Jason Hamilton, and Bernard LeBlanc with logistics. We thank DFO observers Catherine Bajzak, Alejandro Buren, Pierre Carter, Blair Dunn, Patt Hall, Lee Sheppard, Samuel Turgeon, and Brent Young, as well as Inuit observers Leo Ihakik, Robert Karetak, Ivan Koonoo, Karlene Napayok, and Bonnie Tapati. Pierre Rivard and Jean-François Lussier read and georeferenced sightings on aerial photographs. We thank Mike Kristjanson (Chief Infrastructure and Logistics) and Jodi MacGregor (Arctic Logistics Operations Manager) from the

Polar Continental Shelf Program, Earth Sciences Sector, as well as Ken Borek pilots and engineers. This project benefitted considerably from previous DFO satellite tracking projects, and we thank in particular Jack Orr and Bernard LeBlanc for their efforts and input. Funding came from the DFO Ecosystem Science Directorate, the Nunavut Implementation Fund, the Nunavut Wildlife Management Board, the Polar Continental Shelf Program, the Government of Nunavut Fisheries and Sealing, and CEMAM. We also thank 3 anonymous reviewers and Rikke Guldborg Hansen for their thorough revisions and comments on the manuscript.

REFERENCES

- Asselin, N.C., Ferguson S.H., Richard P.R., & Barber, D.G. (2012). Results of narwhal (*Monodon monoceros*) aerial surveys in northern Hudson Bay, August 2011. *DFO Canadian Science Advisory Research Secretariat Research Document, 2012/037*. iii + 23 p.
- Asselin, N.C., & Richard, P.R. (2011). Results of narwhal (*Monodon monoceros*) aerial surveys in Admiralty Inlet, August 2010. *DFO Canadian Science Advisory Research Secretariat Research Document, 2011/065*. iv + 26 p.
- Brack, I.V., Kindel, A., & Oliveira, L.F.B. (2018). Detection errors in wildlife abundance estimates from Unmanned Aerial Systems (UAS) surveys: synthesis, solutions, and challenges. *Methods in Ecology and Evolution*, 9(8), 1864–1873. <https://doi.org/10.1111/2041-210X.13026>
- Bröker, K.C.A., Hansen, R.G., Leonard, K.E., Koski, W.R., & Heide-Jørgensen, M.P. (2019). A comparison of image and observer based aerial surveys of narwhal. *Marine Mammal Science*, 35, 1253–1279. <https://doi.org/10.1111/mms.12586>
- Buckland, S.T., Anderson, D., Burnham, K., Laake, J., Borchers, D., & Thomas, L. (2001). *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press.
- Burt, M. L., Borchers, D. L., Jenkins, K. J., & Marques, T. A. (2014). Using mark-recapture distance sampling methods on line transect surveys. *Methods in Ecology and Evolution*, 5(11), 1180–1191. <https://doi.org/10.1111/2041-210X.12294>
- Cochran, W. (1977). *Sampling techniques*. New York: John Wiley & Sons.
- Cope, J.M., & Punt, A.E. (2009). Drawing the lines: resolving fishery management units with simple fisheries data. *Canadian Journal of Fisheries and Aquatic Sciences*, 66(8), 1256–1273. <https://doi.org/10.1139/F09-084>
- Dietz, R., Richard, P.R., & Acquarone, M. (2001). Summer and fall movements of narwhals (*Monodon monoceros*) from northeastern Baffin Island towards northern Davis Strait. *Arctic*, 54(3), 244–261. <https://doi.org/10.14430/arctic785>
- Dietz, R., Heide-Jørgensen, M. P., Richard, P., Orr, J., Laidre, K., & Schmidt, H. C. (2008). Movements of narwhals (*Monodon monoceros*) from Admiralty Inlet monitored by satellite telemetry. *Polar Biology*, 31(11), 1295–1306. <https://doi.org/10.1007/s00300-008-0466-4>
- Dunn, P. K., & Smyth, G. K. (1996). Randomized Quantile Residuals. *Journal of Computational and Graphical Statistics*, 5(3), 236–244. <https://doi.org/10.2307/1390802>
- Forney, K. A. (2000). Environmental models of cetacean abundance: Reducing uncertainty in population trends. *Conservation Biology*, 14(5), 1271–1286. <https://doi.org/10.1046/j.1523-1739.2000.99412.x>
- Gjertz, I. (1991). The Narwhal, *Monodon monoceros*, in the Norwegian High Arctic. *Marine Mammal Science*, 7(4), 402–408. <https://doi.org/10.1111/j.1748-7692.1991.tb00115.x>
- Hamilton, O.N.P., Kincaid, S.E., Constantine, R., Kozmian-Ledward, L., Walker, C.G., & Fewster, R.M. (2018). Accounting for uncertainty in duplicate identification and group size judgements in mark-recapture distance sampling. *Methods in Ecology and Evolution*, 9, 354–362. <https://doi.org/10.1111/2041-210X.12895>
- Hedley, S. L., & Buckland, S. T. (2004). Spatial models for line transect sampling. *Journal of Agricultural, Biological and Environmental Statistics*, 9, 181–199. <https://doi.org/10.1198/1085711043578>

- Heide-Jørgensen, M. P., Dietz, R., Laidre, K., & Richard, P. (2002). Autumn movements, home ranges, and winter density of narwhals (*Monodon monoceros*) tagged in Tremblay Sound, Baffin Island. *Polar Biology*, 25, 331–341. <https://doi.org/10.1007/s00300-001-0347-6>
- Heide-Jørgensen, M. P., Dietz, R., Laidre, K. L., Richard, P., Orr, J., & Schmidt, H. C. (2003). The migratory behaviour of narwhals (*Monodon monoceros*). *Canadian Journal of Zoology*, 81(8), 1298–1305. <https://doi.org/10.1139/z03-117>
- Heide-Jørgensen, M. P., Laidre, K. L., Burt, M. L., Borchers, D. L., Marques, T. A., Hansen, R. G., Rasmussen, M., & Fossette, S. (2010). Abundance of narwhals (*Monodon monoceros*) on the hunting grounds in Greenland. *Journal of Mammalogy*, 91, 1135–1151. <https://doi.org/10.1644/09-MAMM-A-198.1>
- Heide-Jørgensen, M. P., Hansen, R. G., Westdal, K., Reeves, R. R., & Mosbech, A. (2013). Narwhals and seismic exploration: Is seismic noise increasing the risk of ice entrapments? *Biological Conservation*, 158, 50–54. <https://doi.org/10.1016/j.biocon.2012.08.005>
- Heide-Jørgensen, M. P., Richard, P. R., Dietz, R., & Laidre, K. L. (2013). A metapopulation model for Canadian and West Greenland narwhals: Narwhal metapopulation. *Animal Conservation*, 16(3), 331–343. <https://doi.org/10.1111/acv.12000>
- Heide-Jørgensen, M. P., Hansen, R. G., & Strandler Sinding, M. H. (2014). Abundance and distribution of narwhals in Melville Bay in 2014. Report by the Greenland Institute of Natural Resources. Retrieved from http://www.natur.gl/fileadmin/user_files/Doku
- Hiby, L. & Lovell, P. (1998). Using aircraft in tandem formation to estimate abundance of harbour porpoise. *Biometrics*, 54, 1280–1289. <https://doi.org/10.2307/2533658>
- Higdon, J. W., & Ferguson, S. H. (2017). Database of aerial surveys and abundance estimates for beluga whales (*Delphinapterus leucas*) and narwhals (*Monodon monoceros*) in the Canadian Arctic. *Canadian Technical Report of Fisheries and Aquatic Sciences*, 3211, 48.
- Innes, S., Heide-Jørgensen, M. P., Laake, J. L., Laidre, K. L., Cleator, H., Richard, P., & Stewart, R. E. A. (2002). Surveys of belugas and narwhals in the Canadian High Arctic in 1996. *NAMMCO Scientific Publications*, 4, 169–190. <https://doi.org/10.7557/3.2843>
- Kenyon, K. (2017). Movement patterns and winter habitat use of narwhal (*Monodon monoceros*) in the Eastern Canadian Arctic. (unpublished doctoral dissertation) University of Manitoba, Winnipeg, Manitoba, Canada. Retrieved from <https://mspace.lib.umanitoba.ca/xmlui/handle/1993/32587>
- Laake, J. L., & Borchers, D. L. (2004). Methods for incomplete detection at distance zero. In *Advanced Distance Sampling* (Eds. S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers and L. Thomas.) pp. 108–189. Oxford University Press: Oxford.
- Laake, J.L., Borchers, D., Thomas, L., Miller, D., & Bishop, J. (2019). mrds: Mark-Recapture Distance Sampling. R package version 2.2.1.
- Lydersen, C., Martin, A. R., Gjertz, I., & Kovacs, K. M. (2007). Satellite tracking and diving behaviour of sub-adult narwhals (*Monodon monoceros*) in Svalbard, Norway. *Polar Biology*, 30(4), 437–442. <https://doi.org/10.1007/s00300-006-0200-z>
- Maling, D.H. (1992). *Coordinate systems and map projections*, 2nd edition. Pergamon Press. Oxford. 476 p.
- Marcoux, M., Auger-Méthé, M., & Humphries, M.M. (2009). Encounter frequencies and grouping patterns of narwhals in Koluktoo Bay, Baffin Island. *Polar Biology*, 32(12), 1705–1716. <https://doi.org/10.1007/s00300-009-0670-x>
- McLaren, I. A. (1961). Methods of determining the numbers and availability of ringed seals in the eastern Canadian Arctic. *Arctic*, 14, 162–175. <https://doi.org/10.14430/arctic3671>
- Miller, D.L. (2012). *On smooth models for complex domains and distances*. (unpublished doctoral dissertation). University of Bath, Bath, UK.
- Miller, D. L., & Wood, S. N. (2014). Finite area smoothing with generalized distance splines. *Environmental and Ecological Statistics*, 21(4), 715–731. <https://doi.org/10.1007/s10651-014-0277-4>
- Miller, D.L., Rexstad, E. A., Burt, M. L., Bravington, M. V., & Hedley, S. L. (2013). dsm: Density Surface Modelling of Distance Sampling Data version 2.2.17 from CRAN. Retrieved from <https://rdrr.io/cran/dsm/>
- North Atlantic Marine Mammal Commission (NAMMCO). (2015). Report of the Joint Meeting of the NAMMCO Scientific Working Group on the Population Status of Narwhal and Beluga in the North Atlantic and the Canada/Greenland Joint Commission on Conservation and Management of Narwhal and Beluga Scientific Working Group In: NAMMCO Annual Report 2015; pp. 248–304. Retrieved from <https://nammco.no/wp-content/uploads/2017/08/annual-report-2015.pdf>
- North Atlantic Marine Mammal Commission (NAMMCO). (2018). Report of the NAMMCO Global Review of Monodontids. 13-16 March 2017, Hillerød, Denmark. Retrieved from <https://nammco.no/topics/sc-working-group-reports/>
- Pike, D., & Doniol-Valcroze, T. (2015). Identification of duplicate sightings from the 2013 double-platform High Arctic Cetacean Survey. *DFO Canadian Science Advisory Research Secretariat Research Document*, 2015/034. 27.
- Priest, H., & Usher, P. (2004). *The Nunavut Wildlife Harvest Study*. Retrieved from <https://www.nwmb.com/iku/publications/harvest-study/1824-156-nwhs-report-2004-156-0003/file>
- Quang, P.X., & Lanctot, R.B. (1991). A line transect model for aerial surveys. *Biometrics*, 47, 1098–1102. <https://doi.org/10.2307/2532661>
- Richard, P. R., Hall, P. A., Dueck, L., & Barber, D. G. (1994). Distribution and numbers of Canadian High Arctic narwhals (*Monodon monoceros*) in August 1984. *Meddelelser Om Grønland, Bioscience*, 39.
- Richard, P. R., Laake, J. L., Hobbs, R. C., Heide-Jørgensen, M. P., Asselin, N. C., & Cleator, H. (2010). Baffin Bay narwhal population distribution and numbers: Aerial surveys in the Canadian High Arctic, 2002-04. *Arctic*, 63(1), 85–99. <https://doi.org/10.14430/arctic649>
- Richard, P. R. (2010). Stock definition of belugas and narwhals in Nunavut. *DFO Canadian Science Advisory Research Secretariat Research Document*, 2010/022. iv + 14.
- Richard, P. R., Weaver, P., Dueck, L., & Barber, D. G. (1994). Distribution and numbers of Canadian High Arctic narwhals (*Monodon monoceros*) in August 1984. *Meddelelser Om Grønland, Bioscience*, 39, 41–50.
- Smith, T., Hammill, M. O., Burrage, D., & Sleno, G. A. (1985). Distribution and abundance of belugas, *Delphinapterus leucas*, and narwhals, *Monodon monoceros*, in the Canadian High Arctic. *Canadian Journal of Zoology*, 42(4), 676–684. <https://doi.org/10.1139/f85-087>
- Southwell, C., de la Mare, B., Underwood, M., Quartararo, F., & Cope, K. (2002). An automated system to log and process distance sight-resight aerial survey data. *Wildlife Society Bulletin (1973–2006)*, 30(2), 394–404. <https://www.istor.org/stable/3784496>
- Stevenson, B. C., Borchers, D.L., & Fewster, R.M. (2018). Cluster capture-recapture to account for identification uncertainty on aerial surveys of animal populations. *Biometrics*, 1–11. <https://doi.org/10.1111/biom.12983>
- Thomas, L., Williams, R., & Sandilands, D. (2007). Designing line transect surveys for complex survey regions. *Journal of Cetacean Research and Management*, 9(1), 1–13.
- Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A., & Burnham, K.P. (2010). Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, 47(1), 5–14. <https://doi.org/10.1111/j.1365-2664.2009.01737.x>
- Watt, C. A., Marcoux, M., Asselin, N. C., Orr, J. R., & Ferguson, S. H. (2015). Instantaneous availability bias correction for calculating aerial survey abundance estimates for narwhal (*Monodon monoceros*) in the Canadian High Arctic. *DFO Canadian Science Advisory Research Secretariat Research Document*, 2015/044, 18.
- Watt, C. A., Orr, J., Leblanc, B., Richard, P., & Ferguson, S. H. (2012). Satellite tracking of narwhals (*Monodon monoceros*) from

- Admiralty Inlet (2009) and Eclipse Sound (2010-2011). *DFO Canadian Science Advisory Research Secretariat Research Document*, 2012/046.
- Williams, R., Hedley, S. L., & Hammond, P. S. (2006). Modeling distribution and abundance of Antarctic baleen whales using ships of opportunity. *Ecology and Society*, 11(1), art1. <https://doi.org/10.5751/ES-01534-110101>
- Wood, S. N. (2001). *mgcv: GAMs and generalized ridge regression for R*. R News, 1.
- Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1), 95–114. <https://doi.org/10.1111/1467-9868.00374>
- Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. Second Edition. Boca Raton: Chapman and Hall/CRC. <https://doi.org/10.1201/9781420010404>
- Wood, S. N., Bravington, M. V., & Hedley, S. L. (2008). Soap Film Smoothing. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(5), 931–955. <https://doi.org/10.1111/j.1467-9868.2008.00665.x>