

Beluga, *Delphinapterus leucas*, Group Sizes in Cook Inlet, Alaska, Based on Observer Counts and Aerial Video

RODERICK C. HOBBS, JANICE M. WAITE, and DAVID J. RUGH

Introduction

It is extremely difficult to get accurate counts of large groups of cetaceans from aerial surveys, even though aircraft are usually better observation platforms than vessels or shore-based sites. To overcome this problem, methods and analyses for using video records were developed to obtain counts and estimate group sizes to correct observers' aerial counts. These methods were applied to a study of beluga whales, *Delphinapterus leucas*, in Cook Inlet that involved annual aerial surveys conduct-

ed by NOAA's National Marine Fisheries Service (NMFS) from 1994 to 2000 (Rugh et al., 2000).

The status of the Cook Inlet stock has been of special concern because of its small size (Hobbs et al., 2000), isolation (Laidre et al., 2000), and annual hunting pressures by Native Alaskans (Mahoney and Shelden, 2000). Studies of abundance of this stock and related parameters for correcting aerial counts and assessing variance have been conducted by NMFS annually since 1993.

During late spring and early summer, large aggregations of beluga whales are found near river mouths of upper Cook Inlet (Rugh et al., 2000). In recent years, very few whales have been observed anywhere other than in these nearshore areas, and most of the observed whales were within dense concentrations. Con-

sequently, the accuracy of an abundance estimate for this population depends on accurate estimates of the sizes of these few large groups.

Since 1994, a consistent method has been used by the NMFS for counting groups of beluga whales during annual aerial surveys of Cook Inlet. This method adapted available survey designs (flying consistently at an altitude of 244 m (800 ft) and 185 km/h (100 knots) along coastal tracklines or straight offshore tracklines) to make systematic, thorough searches for whale groups (Rugh et al., 2000). Paired, independent observers provided sighting information used in an analysis of the likelihood that a group of whales was not seen (Hobbs et al., 2000). When a group was found, repeated counts were made using an extended racetrack flight pattern, such that observers always counted on one side of the aircraft while passing a group on a straight line. During each pass, two observers counted the group independently, and a third recorded the group on videotape. The timing of the start and stop of each count was recorded precisely along with the apparent quality of the counts—whether or not the whole group was in view or visibility interfered with the count.

These counting methods worked well for obtaining counts for each group and for comparing the performance of different observers. However, counts made during each pass varied widely among observers, and counts by each observer varied for successive counts of the same group, even when visibility did not change. Large, dense groups remained especially difficult to count accurately despite the standardized approach. Observers responded differently to increases

The authors are with the National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way N.E., Seattle, WA 98115-0070 (e-mail address: rod.hobbs@noaa.gov).

ABSTRACT—Beluga, *Delphinapterus leucas*, groups were videotaped concurrent to observer counts during annual NMFS aerial surveys of Cook Inlet, Alaska, from 1994 to 2000. The videotapes provided permanent records of whale groups that could be examined and compared to group size estimates made by aerial observers. Examination of the video recordings resulted in 275 counts of 79 whale groups. The McLaren formula was used to account for whales missed while they were underwater (average correction factor 2.03; $SD=0.64$). A correction for whales missed due to video resolution was developed by using a second, paired video camera that magnified images relative to the standard video. This analysis showed that some whales were missed either because their image size fell below the resolution of the standard video recording or because two whales surfaced so close to each other that their images appeared to be one large whale. The correction method that resulted depended on

knowing the average whale image size in the videotapes. Image sizes were measured for 2,775 whales from 275 different passes over whale groups. Corrected group sizes were calculated as the product of the original count from video, the correction factor for whales missed underwater, and the correction factor for whales missed due to video resolution (averaged 1.17; $SD=0.06$). A regression formula was developed to estimate group sizes from aerial observer counts; independent variables were the aerial counts and an interaction term relative to encounter rate (whales per second during the counting of a group), which were regressed against the respective group sizes as calculated from the videotapes. Significant effects of encounter rate, either positive or negative, were found for several observers. This formula was used to estimate group size when video was not available. The estimated group sizes were used in the annual abundance estimates.

in beluga density, with some observers continuing to count individual whales but narrowing their field of view, while other observers switched to counting in 10's or 20's, maintaining a larger field of view but reducing the accuracy of their counts. In addition, animals were missed due to visibility biases (availability and perception biases) (Marsh and Sinclair, 1989), which were not easily measured.

Availability bias occurs when whales are not visible to observers because the whales are completely below the surface—especially in turbid water—and so are not “available” to be counted (Marsh and Sinclair, 1989). In upper Cook Inlet, the waters are extremely turbid to the point of being nearly opaque. Secchi disk depth readings range from 1 cm to 37 cm with an average of 14 cm in areas where beluga whales are typically found (Shelden and Angliss¹), such that any part of a beluga that is below the surface is out of sight.

Water disturbance patterns created when whales swim near the surface are often visible and can serve as sighting cues; however, observers were asked to count only those belugas that they actually saw at the surface. It is unknown to what extent their counts are influenced by the water disturbance cues. Researchers have corrected for availability bias using the formula of McLaren (1961) which uses dive interval information to estimate the inverse of the probability that a typical animal will be at the surface during the period of observation (Frost et al., 1985; Barlow et al., 1988; Laake et al., 1997). The McLaren formula also requires knowing the time that an observer spends counting whales in each patch of water which, as noted above, can vary by observer and group size.

Perception bias occurs when whales are on the surface of the water (avail-

able to be seen) but are not detected by an observer due to various possibilities, including adverse conditions (such as glare or rough seas), camouflage coloration, or observer fatigue or inexperience (Marsh and Sinclair, 1989). Perception bias is further compounded for beluga whales due to their age-specific coloration. Belugas are dark gray when born and gradually lighten as they get older, until they become completely white as adults (9–11 years) (Sergeant, 1973). When at the surface, belugas appear as white or gray ovals against a dark background of water. The white adult whales strongly contrast with the water, which makes them easy to see even at a distance. The smaller, darker whales are more likely missed in aerial counts.

To circumvent these inherent problems with observers' counts from an aircraft, a method using video recordings was developed as an alternative. Video recordings have precise timing, have a well defined field of view, and can be examined closely frame-by-frame to ensure that all of the recorded beluga images are counted. These properties allowed for precise, repeatable counts, and accurate measurements of the time available for counting, including the time that whales were at the surface. The objectives of the video recordings during these surveys and subsequent analysis were to collect the following types of data:

- 1) Accurate counts of all visible surfacings in a group over a measured period of time.
- 2) Accurate measurement of the time whales were visible at the surface (start to end of a surfacing).
- 3) Apparent size of whale images in the video recordings.
- 4) Gray-scale classification for whale images in video recordings.
- 5) And, matching of whale images between simultaneous video recordings at two different magnifications.

Using these data and the average dive interval from a radio-tagging study (Lerczak et al., 2000), the counts of surfacing whales from the video could be corrected for availability bias and perception bias. Availability bias could be

estimated using the McLaren (1961) formula, and perception bias could be estimated by matching images between simultaneous high and low magnification video recordings of the same group. The calculations of group size resulting from the video analysis could then be used to develop a correction method for the observer counts independent of assumptions about observer behavior and counting techniques.

Methods

Aerial Counts of Beluga Groups

Aerial counts were made by pairs of observers tallying beluga whale sightings independently during counting passes (Rugh et al., 2000). A racetrack flight pattern, typically 2–4 km from end-to-end and 1–2 km across, depending on the size of a group, usually allowed two counting passes per circuit around the group, if glare was not a problem (Fig. 1). Start and stop times were precisely noted for each counting pass, and observers graded the conditions during each count (from A to D for excellent to poor). The protocol was to continue this process until each of the observers had made at least four counts under acceptable conditions (A or B). Then observers traded positions, and the process was repeated.

Typically, then, each group of whales was counted 16 times (4 times by 4 different observers). This method allowed for repeatable, independent counts by four observers with essentially the same presentation of the group and a measure of time spent counting. The number of counts was reduced when groups were small (<5 whales) or dense air traffic prevented staying in the area. Each observer recorded counts independently, and counts were not discussed among observers during the remainder of the project. To further maintain independence, the counts were entered into the database by a colleague who did not participate in the counts, or they were entered only after the survey was completed.

Standardized Counting Video

Videotape recordings (hereafter referred to as counting video), concurrent to observers' counts of beluga whale

¹ Shelden, K. E. W., and R. P. Angliss. 1995. Characterization of beluga whale (*Delphinapterus leucas*) habitat through oceanographic sampling of the Susitna River delta in Cook Inlet, Alaska, 11–18 June 1994. In D. P. DeMaster, H. W. Braham, and P. S. Hill (Editors), Marine Mammal Assessment Program status of stocks and impacts of incidental take, 1994, p. 77–90. Annu. Rep. submitted to Off. Protect. Resour., Natl. Mar. Fish. Serv., NOAA, 1335 East-West Highway, Silver Spring, MD 20910.

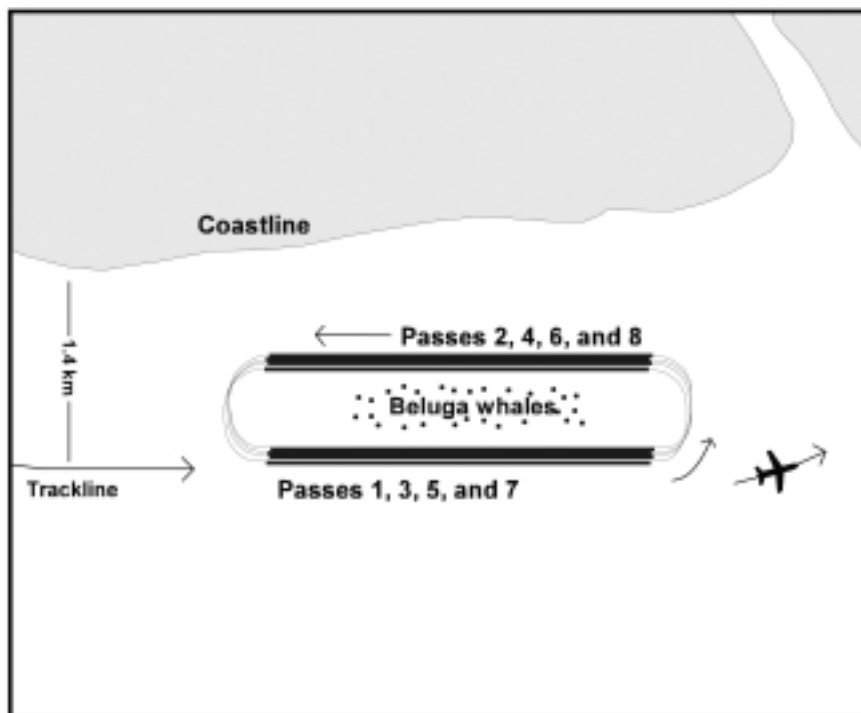


Figure 1.—Schematic of the racetrack flight pattern used to count beluga whale groups in Cook Inlet. After observers saw a group of whales, the aircraft circled in a series of straight-line passes such that each pair of observers had four opportunities to count the group from the same side of the aircraft. The group was videotaped concurrent to each counting pass.

groups, were collected with a Canon² Hi-8 814 XL-S video camera during 1994–98 and a Sony Digital 8 DCR-TRV103 video camera during 1999–2000. These tapes were used to count whales in each group independent of the observers. In the aircraft, the videographer was positioned at an open window facing the same direction as the two observers so that all three would have essentially the same view of each group of whales (Rugh et al., 2000 provides details).

Videotape, with a time-stamp, was recorded during nearly every counting pass over a beluga group. The counting video was taken in either of two ways: 1) The “point” method was used if the whole group could be seen in the viewfinder of the video camera. In this case, the camera was pointed at the group and moved to keep the entire group

of whales in the field of view until they were out of sight (Fig. 2a); or 2) The “scanned” method was used when groups were too dispersed for all individuals to be in view at one time. With this method, the camera was held still and perpendicular to the trackline so that the video scanned the length of the group as the aircraft moved past (Fig. 2b).

Zoomed Video

Although the counting video was valuable for counting clearly visible, distinct whales in a group, small whales may not have been visible in the video at the limits of the resolution, or whales may not have been distinguishable due to close proximity with other whales. In 1996, a second video camera (Ricoh Hi-8 R800H) was used at a higher magnification (8×) during counting passes, to test the visibility of small and gray whales in the standard video.

Upon examination of the zoomed video, we found that the narrow field

of view resulted in capturing few usable whale images during these passes. An experiment was then devised in which a large group of beluga whales was circled continuously while both cameras videotaped the group. The two video cameras (standard and zoomed) were mounted side by side on a board so they were held steady and parallel to each other. The circling allowed whales to be in view for longer periods of time and, thus, they were more likely to be captured on the zoomed video. This circling experiment was repeated in 1998 and 1999.

In June 2000, a Sony Digital Camcorder DSR-PD100A replaced the Ricoh as the zoomed video camera, and it was used during counting passes instead of a circling experiment. It had higher resolution and could be zoomed to a lower magnification (~5×) and still collect useable whale images while allowing a broader field of view. Sufficient zoomed images were captured using this camera during counting passes so that a separate experiment was not necessary.

It was assumed that within the viewing range of the zoomed video, all whales at the surface could be distinguished. We believe this assumption is reasonable, because the higher magnification allowed positive identification of small dark gray whales next to large white whales and the distinction between larger gray and white animals.

Video Analysis

Each pass on the videotapes was reviewed to evaluate its quality for counting beluga whale groups and was given a rating (excellent, good, fair, poor, or unacceptable). The highest rating was given to passes in which the camera panned smoothly while remaining in focus and the margins of a group were visible. Video passes were given a lower rating if all or a part of a group was obscured by glare, confused by waves, or if focus, magnification, or panning varied rapidly. Only excellent and good passes were used for the group size estimation analysis.

The video recordings were examined using a Panasonic high resolution monitor and a Hi-8 video cassette recorder (VCR) (Sony EVO 9500A) (1994–98)

² Mention of trade names or commercial firms does not imply endorsement by the National Marine Fisheries Service, NOAA.

or the Sony digital camera in playback mode (1999–2000) capable of advancing and reversing the tape frame-by-frame and assigning a unique time and frame count number to each frame. Each frame corresponded to 1/30th of a second.

Groups were counted in two different ways according to group size: 1) for small groups, whales were counted directly from the monitor screen as the video played at regular or slow speed and 2) for large groups, whale locations were “captured” by stopping the video every 0.5 seconds (15 frames). This timing was based on the assumption that no whale would surface and dive again within a 0.5-sec period.

Transparency sheets were placed on the monitor screen and marked with dots to indicate the position of each beluga image. The sheets were then compared by placing one on top of the next; this allowed for differentiating between new sightings and whales which were resighted from one sheet to the next. New sightings were marked as new and tallied by sheet. The number of whales on the first sheet plus the number of “new” whales on each successive sheet were then summed to derive a total count for the aerial pass.

To correct for availability bias, the time spent counting was needed. This was determined differently depending on how the video was taken. For the “point” groups, the amount of time that the group was in view was used. For the “scanned” groups, timing was based on how long an object on the water was in view across the screen.

This method of counting beluga whales from video recordings was developed in several steps. Initially, two people viewed an aerial pass together with open discussion about which images represented actual beluga whales. Then a second series of counts were made by three people independently (including the two who had made the first count together). These counts were then examined sheet-by-sheet (every 0.5-sec) by all three reviewers. Discrepancies between reviewers were discussed, and a consensus was reached on whether a dot on the screen could be considered a whale or not.

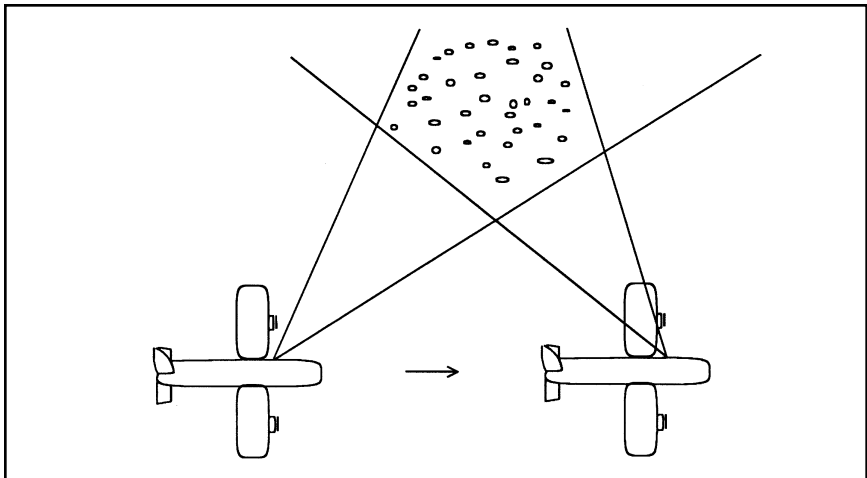


Figure 2a.—Video camera position during a counting pass. In a “point” view of a whale group, the camera is centered on the group, and the videographer changes the camera angle to keep the group centered.

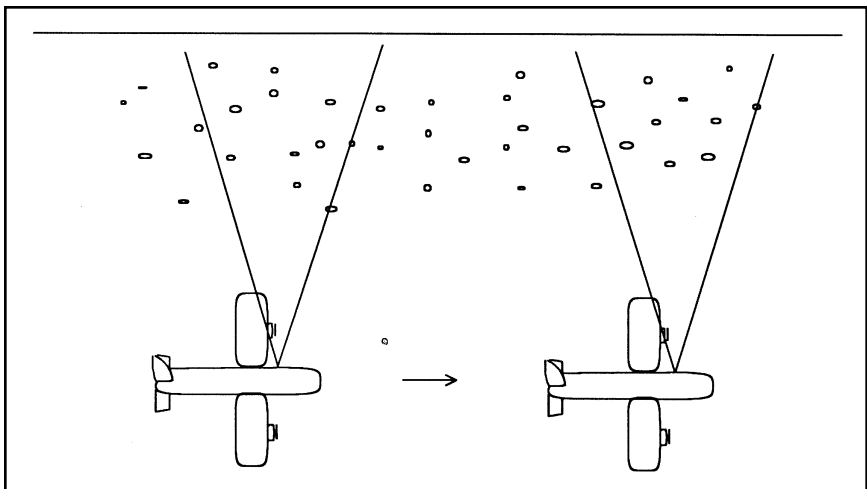


Figure 2b.—In a “scan” view, the video camera is held at a fixed angle (usually centered on the beam line), and the group is videotaped as the plane flies past.

For example, we found that a white bird in the video could be mistaken for a beluga whale, but we learned through this process how to recognize which dots were birds. After the method was developed by the initial group of three analysts, new analysts were trained by following a progression from open discussion, checks on their counts by experienced analysts, to independent analysis.

Time-at-the-surface, Image Size, and Gray-scale

To determine the distribution of whale image sizes visible from the aerial video

and to estimate the time that whales were visible at the surface, a sample of whales was examined from each of the passes that were counted on transparency sheets. For each group, whales from every tenth sheet (every 5 sec) were identified with a number. Since the average dive interval is typically longer than 20 sec (Lerczak et al., 2000), this method essentially grabbed a random subsample of 10–30 whales per group. Each numbered whale was then followed using the frame-by-frame mode on the video cassette recorder; essentially following a whale from the time it

first appeared to the time it disappeared. Time and frame numbers were used to determine the length of time each whale was visible at the surface, so that the error in timing was at most 0.07 sec.

Images of whales that were measured for time-at-the-surface were also measured for image size and given a gray-scale rating. Measurements were taken at the halfway point in each whale's time-at-the-surface. A plastic metric ellipse template was used as a scale for size. The template was copied, doubled, and reduced by half (using a photocopier) to create a set of standard sizes of white ellipses on a gray background to match the typical image and range of sizes of the whale images. Each whale ellipse was classified with the length of the major axis (in millimeters) relative to the standardized scaled chart. Measurements by two independent analysts were averaged. A gray-scale rating for each whale (bright white, dull white, and gray) was determined by consensus.

Comparison of Counting Video and Zoomed Video

The zoomed video was evaluated for usable sections. Using two video players and monitors side-by-side, segments from the zoomed video were matched to segments taken from the counting video by comparing the spatial and movement patterns of whales. The two videotapes were synchronized to the nearest frame possible. Once segments were synchronized, the relative magnification was measured by finding a pair of whales on the zoomed video that could be matched on the standard video.

Positions of the whales were marked on transparencies sheets placed on the monitor screens and distances in millimeters measured between the pair of whales on each screen. Using the position of the matched whales and the relative magnification, the boundary of the zoomed image on the counting video could be calculated. A rectangle was then drawn on the standard video sheet, marking the boundaries of the zoomed image. Marks were made on the transparencies indicating whether whales were visible on both videotapes or only seen on the zoomed video.

The size of whale images in the zoomed frames were then measured following the method of the counting video (using the half-way time calculated for the counting video when it was available) and assigned a gray-scale rating following a scheme with four shades (white, off-white, light gray, and dark gray). Because the images were larger, it was possible to distinguish four gray-shades rather than the three shades used in the standard video.

Correction for Perception Bias

Detecting whales in video recordings is limited by the resolution of the video system. Probability of detection was measured by comparing the whales seen in the zoomed video to those seen in the corresponding region on the counting video. The whale images in the zoomed video were each assigned to one of three categories: 1) whales that were seen in both the zoomed and counting video, 2) whales in the zoomed video that were missed in the counting video due to proximity—two whales surfacing close to each other appear as one large image on the counting video, or 3) image size—a whale seen in the zoomed is too small or gray so that it falls below a threshold and does not form a visible image on the counting video. The two mechanisms (proximity and image size) that affect whale detection in the counting video require different approaches for correction.

Proximity Correction

When two whales were close enough together to appear as a single whale on the counting video, the space between them was much narrower than the width of an average whale. Consequently, these two images would be merged throughout the typical range of magnifications used in the counting video, regardless of their relative size. Thus, a constant ratio could be used to correct for whales missed due to proximity:

$$\frac{J_z}{J_z - J_p},$$

where J_z is the number of whales seen in the zoomed video and J_p is the number of whales missed due to proximity.

Image Size Correction

The resolution of a video system is limited by the density of scan lines in the video recording system and the density of pixels on the display monitor. This process of scanning and pixelation has the effect of smearing images and edges by averaging the gray scale and hue across each pixel. If a pixel is half water and half beluga, then it will appear to have a gray scale and hue halfway between that of the water and the beluga. A large, white beluga will appear as a bright white ellipse with a fuzzy edge that fades to the gray-scale of the water. A gray beluga will appear as a gray ellipse with a less distinct fading to the water color. Small, gray belugas or distant belugas of any hue may not have a sufficiently large image to completely fill any pixels so that the image is entirely made up of these averaged pixels.

Because the edge of the image has been blurred, it is necessary to interpret by eye the margin of the image from the surrounding background. Experience has shown that with a limited amount of training, consistent and repeatable measurements of beluga images can be made. However, the measurement method is partially subjective, so it is necessary to estimate the bias in the interpreted image size. The smearing occurs only at the edges, so the bias should be independent of size. The gradient that is interpreted is dependent on the difference in hue between the object and the background. The subjectivity involves a determination of the point along this gradient that is the edge of the image.

The lengths of the images at the mid times of beluga surfacings matched between the counting and zoomed videos can be related by the following formula:

$$\frac{L_z}{m} = \frac{l_z - b}{m} = l_c - b = L_c,$$

where L_z and L_c are the unbiased sizes of the whale images on the zoomed and counting videos, respectively; l_z and l_c are the measured sizes of the whale images on the zoomed and counting videos, respectively; m , is the relative magnifica-

tion between the zoomed and counting video frames (obtained as the ratio of the distance between centers of two whale images seen on both the counting and zoomed video); and b is the bias resulting from smearing of the edge.

An average value for the bias can be estimated from several image pairs as,

$$\hat{b} = \frac{\sum_{j=1}^{J_n} (l_{zj} - m_j l_{cj})}{\sum_{j=1}^{J_n} (1 - m_j)},$$

where J_n is the number of whales seen in both the zoomed and counting video, and j is the index of the j th pair. If b was not significantly different from zero, it was not necessary to correct for bias.

The following equation was then used to estimate the image size in the counting video for the whales that were visible in the zoomed video but, because of their size, were not detected in the counting video. The estimated image size for these whales in the counting video was:

$$\hat{l}_c = \frac{l_z - \hat{b}}{m} + \hat{b}.$$

A binomial logistic regression was applied to the resulting combined distribution of measured and estimated standard image sizes to estimate the probability that a whale with a given image size would be seen in the counting video.

For a given group, g , and pass, p , m is not known. Instead, the average of image sizes, $\mu_{g,p}$, and the fractions of whales that would be detected, $F(\mu_{g,p})$, in a counting video are related. To determine this relationship, arbitrary values for magnification, m' , (e.g. magnification increasing at 0.01 intervals) are chosen to span the range of possible magnifications. The combined distribution of observed (whales seen in both the zoomed and counting video) and estimated (whales seen in the zoomed but missed in the counting because of small size) counting video sizes are then re-scaled by

$$\hat{l}_{m'j} = (l_{cj} - \hat{b})m' + \hat{b}$$

to simulate the distribution of image sizes under these arbitrary magnifications. For this re-scaled distribution, the average of image sizes, $\mu(m')$, and the fractions of whales, $F(\mu(m'))$, that would be detected in a counting video are

$$\mu(m') = \frac{\sum_{J_n+J_w} P(\hat{l}_{m'j}) \hat{l}_{m'j}}{\sum_{J_n+J_w} P(\hat{l}_{m'j})},$$

$$F(\mu(m')) = \frac{\sum_{J_n+J_w} P(\hat{l}_{m'j})}{J_n + J_w},$$

where $P(l)$ is the probability that an image of size l will be seen in the counting video.

A lookup table relating average image size for a group counted from video, $\mu_{g,p}$, to the correction for the fraction that were missed because of image size, $1/F(\mu_{g,p})$, was created from this analysis. For passes with a sample of measured images, the fraction missed was found in the table. Passes of small groups where images were not measured were given the average fraction missed from other passes of the same group, or if no other passes on the group had measured images, the pass was given the average fraction missed of all measured passes from all groups.

Combined Correction Factor

The correction for perception bias was the product of the proximity correction and the image size correction. For a video count with an estimated average image size, $\hat{\mu}_{gp}$, the correction factor, $D_{g,p}$ is then,

$$D_{gp} = \left[\frac{J_z}{J_z - J_p} \right] \frac{1}{F(\hat{\mu}_{gp})}.$$

Correction for Availability

The formula of McLaren (1961) for the correction for availability bias is the inverse of the probability that a typical beluga is at or will appear at the surface during the videotaping. The correction factor, $A_{g,p}$, for a group and pass

depending on the time spent counting, $t_{g,p}$, is calculated as,

$$A_{g,p} = \frac{T_l}{T_s + t_{g,p}},$$

where T_l is the average dive interval (24.1 sec., Lerczak et al., 2000), and T_s is the average time at the surface from the video analysis described above.

Estimation of Group Size

The group size, \hat{n}_g , was estimated by averaging the corrected video counts for a group:

$$\hat{n}_g = \frac{1}{P_g} \sum_{P_g} c_{g,p} D_{g,p} A_{g,p},$$

where, $c_{g,p}$ is the count for group g from pass p , and P_g is the number of passes for group g that were counted. When a video pass contained two or more distinctly different segments (e.g. it began using the point method, then switched to the scan method when the first portion of the group came abeam of the plane), the counts were corrected separately to create a group size estimate for each subgroup. These subgroup estimates were then summed to estimate the total group size.

The coefficient of variation (CV) for \hat{n}_g was estimated as:

$$CV(\hat{n}_g) =$$

$$\sqrt{\frac{CV^2(n)}{P_g} + CV^2(D_g) + CV^2(A_g)}.$$

An average CV for a group size estimate made from a single count was estimated by averaging the variation of the group size estimates of all groups where more than one pass from the group was counted from video,

$$\overline{CV^2}(n) = \frac{\sum_{G^2} \frac{1}{P_g - 1} \sum_{P_g} (n_{g,p} - \bar{n}_g)^2}{\sum_{G^2} \bar{n}_g^2}.$$

Where more than one count is used to estimate group size, this average CV is

scaled appropriately. The value $CV(n)$ includes an empirical measure of stochastic variation between counts that is not corrected by the two correction factors, but it does not account for the variation of the correction factors themselves which must be accounted for separately.

The component of the CV resulting from the correction for perception, $CV(D_g)$, is estimated by the delta method as,

$$\overline{CV}^2(D_g) = \frac{\sum_{P_g} \left[\frac{\partial D}{\partial \mu} \Big|_{\mu_{g,p}} \right]^2 SE^2(\mu_{g,p})}{\left[\sum_{P_g} D_{g,p} \right]^2}$$

For cases where $\mu_{g,p}$ was not estimated, the correction factor $D_{g,p}$ was derived from an average of $\mu_{g,p}$ from other passes of the same group or an average of other groups. In these cases $SE(\mu_{g,p})$ was the standard deviation of the set of the estimated average image sizes of the averaged groups.

The component of the CV resulting from the correction for availability, $CV(D_g)$, is dominated by the variation of T_I . The variation of T_I has a component related to the variability between individuals and the variation of a typical individual. Following the delta method yields,

$$\overline{CV}^2(A_g) = \frac{\sum_{P_g} \left[\frac{1}{T_s + t_{g,p}} SE_g(T_I) \right]^2}{\left[\sum_{P_g} A_{g,p} \right]^2} = \frac{\frac{\sigma_A^2 - \sigma_I^2}{\hat{n}_g} \sum_{P_g} \left[\frac{1}{T_s + t_{g,p}} \right]^2}{\left[\sum_{P_g} A_{g,p} \right]^2},$$

where σ_A^2 ($= 41 \text{ sec}^2$, $\sigma_A = 6.4 \text{ sec}$) and σ_I^2 ($= 707 \text{ sec}^2$, $\sigma_I = 26.6 \text{ sec}$) are

the variance of the average dive interval among individuals and the average variance of the dive interval of individuals, respectively (values taken from Lerczak et al., 2000). Note that in this formulation, $CV(A)$ was not independent of group size because of the assumption that the dive behavior of individuals in the group is uncorrelated so that the variation in the average of dive intervals during the counting interval decreases as group size increases.

Group Size Estimates from Observer Counts

Good quality video was not available for all groups, so a method for estimating group size from observer counts was devised. Aerial counts of beluga whales were corrected for observer differences and the effect of encounter rate (group density in whales per second). Data from observers who had participated in the equivalent of one or more complete survey seasons (three surveys of the upper inlet and one survey of the lower inlet) were included in the analysis. Only counts made during passes considered by the observers to be excellent or good in quality (A or B) were used. Group sizes, estimated from video recordings, were used to represent the true group size.

This method provided a correction for availability and perception as well as the uncertainty in the time available to observers to count individual whales. The correction formula was derived by regression of the video-derived group sizes against the observer counts for those groups and an interaction term between the counts and the observed encounter rate with the intercept fixed at zero:

$$\hat{n}_{g,p,o} = c_{g,p,o} \left(\hat{b}_{1,o} + \hat{b}_{2,o} \frac{c_{g,p,o}}{t_{g,p}} \right)$$

$$Var(\hat{n}_{g,p,o}) = c_{g,p,o}^2 \begin{pmatrix} SE^2(\hat{b}_{1,o}) \\ +2Cov(\hat{b}_{1,o}, \hat{b}_{2,o}) \frac{c_{g,p,o}}{t_{g,p}} \\ +SE^2(\hat{b}_{2,o}) \frac{c_{g,p,o}^2}{t_{g,p}^2} \end{pmatrix}$$

where $\hat{n}_{g,p,o}$ is the size estimate for group g from a count by observer o during pass p , $\hat{b}_{1,o}$, $\hat{b}_{2,o}$ are the parameters estimated for each observer by linear regression, $C_{g,p,o}$ is the count by observer o of group g during pass p , $t_{g,p}$ is the time spent counting group g during pass p , $SE^2(\hat{b})$ is the squared standard error of the regression coefficients, $\hat{b}_{1,o}$, $\hat{b}_{2,o}$, and $Cov(\hat{b}_{1,o}, \hat{b}_{2,o})$ is the estimated covariance of the regression coefficients, $\hat{b}_{1,o}$, $\hat{b}_{2,o}$.

This approach weights the correction formula to be most accurate for large groups where a bias would have the greatest impact on the abundance estimate. The first summand estimates a multiplicative correction for counts to group size; the second summand estimates an additive bias proportional to the count multiplied by the density of the group. For aerial counts without recorded time, a single multiplicative correction was also estimated. The correction formula was applied to counts from groups where no group size estimate was available through the video analysis. These corrected counts were then averaged to estimate the group size:

$$\hat{n}_g = \frac{\sum_{J_g} \hat{n}_{g,p,o}}{J_g}$$

$$Var(\hat{n}_g) = \frac{\sum_{J_g} (\hat{n}_g - \hat{n}_{g,p,o})^2}{J_g - 1} + \frac{\sum_{J_g} Var(\hat{n}_{g,p,o})}{J_g^2},$$

where \hat{n}_g is the estimated size of group g , and J_g is the set of corrected observer counts for group g .

Results

Aerial Counts

There were 144 sightings of beluga whale groups during aerial surveys of Cook Inlet in 1994–2000 (Table 1). Many of these represent multiple resightings of groups encountered on different survey days and years. Of the 144 sightings, 126 were counted and video

Table 1.—Beluga whale group size estimates made from aerial surveys of Cook Inlet, Alaska, 1994–2000. Count method refers to either group size estimates made from video counts only (Vid.) or from corrected observer counts (Obs.). Elapsed times are: video point passes (P), video scan passes (S), observer counts (O), or elapsed time not available (na). Correction factors for group size estimates made from video counts were calculated for each pass separately with group averages given here. For groups with no usable video, group size was estimated from observer counts using a formula derived by regression of counts vs. group size where video was available (see Table 2 for parameters).

Date	Group	Location	Count method	Average group count	Number of counts	Average elapsed time during count (sec)	Correction for subsurface whales (A)	CV(A)	Correction for whales missed at the surface (D)	CV(D)	CV (video count)	Estimated group size (N_g)	CV(N_g)
6/1/94	1	Big Susitna	Vid.	156	4	17.98 (P)	1.22	0.13	1.12	0.01	0.11	209	0.17
6/1/94	2	Big Susitna	Vid.	145	1	14.00 (P)	1.46	0.22	1.24	0.04	0.21	263	0.31
6/2/94	1	W of Big Susitna	Obs.	119	4	na						394	0.6
6/2/94	2	Turnagain Arm	Obs.	6	4	na						18	0.47
6/2/94	3	Chickaloon Bay	Obs.	15	4	na						47	0.43
6/3/94	1	Pt Possession	Obs.	10	5	38.00 (O)						27	0.39
6/3/94	2	Kachemak Bay	Obs.	3	5	21.00 (O)						8	0.31
6/3/94	3	Kachemak Bay	Obs.	5	3	25.00 (O)						13	0.29
6/4/94	1	Iniskin Bay	Obs.	2	4	25.67 (O)						4	0.3
6/4/94	2	W of Big Susitna	Obs.	59	4	20.50 (O)						144	0.31
6/4/94	3	Big Susitna	Vid.	155	2	14.97 (P)	1.38	0.17	1.17	0.02	0.15	252	0.22
6/4/94	4	W of Little Susitna	Vid.	123	1	11.50/4.60 (P/S)	3.39	0.11	1.14	0.02	0.21	475	0.24
6/5/94	1	Pt Possession/E Foreland	Obs.	1	3	11.00 (O)						2	0.93
6/5/94	2	Beluga R	Vid.	16	1	17.60/8.40 (P/S)	2.21	0.45	1.16	0.09	0.21	41	0.51
6/5/94	3	W of Big Susitna	Vid.	16	1	19.80 (P)	1.08	0.94	1.16	0.09	0.21	20	0.96
6/5/94	4	W of Big Susitna	Obs.	9	6	17.00 (O)						21	0.34
6/5/94	5	W of Big Susitna	Vid.	29	1	31.47/12.70 (P/S)	1.59	0.46	1.19	0.02	0.21	55	0.51
6/5/94	6	Little Susitna	Obs.	145	6	38.50 (O)						337	0.19
6/5/94	7	Chickaloon	Vid.	6	1	65.93/13.38 (P/S)	1.52	1.08	1.16	0.09	0.21	11	1.1
7/18/95	1	Chickaloon	Vid.	18	5	15.97 (P)	1.36	0.34	1.26	0.03	0.09	29	0.36
7/18/95	2	McArthur R	Obs.	1	8	25.60 (O)						1	1.15
7/18/95	3	Big Susitna	Obs.	331	7	97.71 (O)						731	0.36
7/19/95	1	Chickaloon	Vid.	14	6	17.27 (S)	1.42	0.39	1.16	0.04	0.09	20	0.41
7/19/95	2	McArthur R	Obs.	4	5	20.00 (O)						8	0.58
7/19/95	3	Shirleyville	Obs.	4	1	na						4	0.42
7/19/95	4	Big Susitna	Vid.	101	2	14.88 (S)	1.72	0.12	1.35	0.1	0.15	348	0.22
7/20/95	1	Chickaloon	Vid.	7	4	8.95 (S)	2.11	0.38	1.16	0.05	0.11	16	0.4
7/20/95	2	Big Susitna	Vid.	48	1	17.00 (P)	1	0.45	1.23	0.03	0.21	73	0.5
7/20/95	3	Big Susitna	Obs.	104	5	56.40 (O)						309	0.41
7/21/95	1	Big Susitna (E)	Vid.	18	1	13.00 (S)	1.55	0.61	1.16	0.09	0.21	32	0.65
7/21/95	2	Big Susitna (W)	Obs.	132	5	109.71 (O)						278	0.77
7/21/95	3	Knik Arm	Obs.	2	1	na						2	0.42
7/21/95	4	Chickaloon	Obs.	18	8	22.75 (O)						43	0.26
7/22/95	1	Big R	Vid.	9	4	12.29 (S)	1.75	0.41	1.16	0.05	0.11	17	0.42
7/24/95	1	Drift R	Obs.	2	5	18.50 (O)						5	1.05
7/24/95	2	McArthur R	Obs.	2	12	11.83 (O)						4	0.68
7/24/95	3	Big Susitna (W)	Vid.	86	1	32.50/7.00 (P/S)	2.54	0.16	1.25	0.03	0.21	272	0.27
7/24/95	4	Big Susitna (E)	Obs.	34	7	30.00 (O)						94	0.45
6/11/96	1	S of Beluga R	Obs.	2	1	na						2	0.42
6/11/96	2	Lewis R	Obs.	6	6	18.00 (O)						13	0.69
6/11/96	3	Ivan R	Obs.	2	1	na						2	0.42
6/11/96	4	Theodore R	Obs.	4	3	20.67 (O)						9	0.49
6/11/96	5	Lewis R	Obs.	117	13	115.33 (O)						256	0.25
6/12/96	1	Knik Arm	Obs.	5	5	17.33 (O)						12	0.39
6/12/96	2	Knik Arm	Obs.	2	1	na						4	0.42
6/12/96	3	Big Susitna (after stranding)	Vid.	43	2	19.50/8.53 (P/S)	1.64	0.3	1.16	0.07	0.15	69	0.34
6/12/96	4	Pt Possession	Obs.	17	1	na						48	0.58
6/12/96	5	Lewis R	Obs.	136	12	83.14 (O)						304	0.29
6/12/96	6	Theodore R	Obs.	17	9	28.20 (O)						12	0.39
6/12/96	7	Lewis R	Vid.	32	2	25.03 (P)	0.88	0.58	1.16	0.07	0.15	33	0.6
6/12/96	8	Big Susitna R	Obs.	73	2	52.50 (O)						199	0.186
6/12/96	9	Big Susitna R	Obs.	20	5	40.75 (O)						47	0.294
6/13/96	1	Knik Arm	Obs.	7	7	25.50 (O)						17	0.26
6/13/96	2	Knik Arm	Obs.	8	7	24.50 (O)						18	0.45
6/13/96	3	Pt Possession	Vid.	26	3	9.58 (S)	2.13	0.22	1.16	0.05	0.12	69	0.26
6/13/96	4	Ivan R	Obs.	80	8	36.25 (O)						168	0.2
6/13/96	5	Big Susitna R	Vid.	79	3	7.39 (S)	2.6	0.1	1.16	0.05	0.12	229	0.17
6/14/96	1	Pt MacKenzie	Obs.	18	6	24.00 (O)						39	0.39
6/16/96	1	Knik Arm	Obs.	17	6	17.80 (O)						37	0.15
6/16/96	2	Knik Arm	Obs.	13	10	83.60 (O)						28	0.32
6/16/96	3	Pt Possession	Vid.	15	2	6.77 (P/S)	2.32	0.32	1.16	0.07	0.15	40	0.36
6/16/96	4	Lewis/Ivan R	Vid.	107	8	5.36 (S)	3.08	0.05	1.12	0.01	0.07	365	0.09
6/16/96	5	Big Susitna	Vid.	34	4	11.24 (P/S)	1.75	0.2	1.16	0.05	0.11	69	0.23
6/16/96	6	Big/ Little Susitna	Obs.	55	8	32.50 (O)						132	0.39
6/16/96	7	Little Susitna	Vid.	21	2	21.74 (P)	1.02	0.73	1.16	0.07	0.15	22	0.74
6/17/96	1	Ivan/ Big Susitna	Vid.	125	2	5.07 (S)	3.25	0.08	1.11	0.01	0.15	446	0.17
6/17/96	3	Little Susitna	Vid.	22	3	18.28 (P)	1.19	0.42	1.15	0.02	0.12	30	0.44
6/17/96	4	Ivan/ Big Susitna	Vid.	65	1	15.50/5.38 (P/S)	3.06	0.16	1.16	0.09	0.21	230	0.28
6/8/97	1	Knik Arm	Vid.	14	3	13.27/4.70 (P/S)	2.14	0.31	1.1	0.01	0.12	30	0.34
6/8/97	2	Knik Arm	Vid.	35	5	17.50/12.64 (P/S)	1.41	0.21	1.29	0.02	0.09	63	0.23
6/8/97	3	Knik Arm	Vid.	42	5	26.50/9.42 (P/S)	1.8	0.18	1.17	0.01	0.09	76	0.2

Continued on next page.

Table 1.—Continued.

Date	Group	Location	Count method	Average group count	Number of counts	Average time elapsed during count (sec)	Correction for subsurface whales (A)	CV(A)	Correction for whales missed at the surface (D)	CV(D)	CV (video count)	Estimated group size (N_g)	CV(N_g)
6/8/97	4	Knik Arm	Obs.	1	1	na						2	0.42
6/8/97	5	Knik Arm	Vid.	35	5	8.90 (S)	2.18	0.18	1.18	0.01	0.31	96	0.4
6/8/97	7	Knik Arm	Obs.	4	1	na						4	0.42
6/8/97	8	Chickaloon	Vid.	11	6	16.20/8.28 (P/S)	2.02	0.31	1.19	0.02	0.09	20	0.32
6/8/97	9	Chickaloon	Vid.	9	4	17.12/9.47 (P/S)	1.73	0.42	1.19	0.01	0.11	19	0.43
6/8/97	10	Big Susitna	Vid.	49	3	9.30 (S)	2.1	0.17	1.21	0.03	0.12	127	0.21
6/9/97	1	Tuxedni Bay	Obs.	2	1	na						2	0.42
6/9/97	2	Big Susitna	Obs.	58	9	37.67 (O)						103	0.36
6/10/97	1	Chickaloon	Vid.	30	4	4.87 (S)	3.3	0.12	1.14	0.01	0.11	113	0.16
6/10/97	2	Big Susitna	Obs.	70	15	57.75 (O)						140	0.35
6/10/97	3	Knik Arm	Vid.	67	7	9.37 (S)	2.15	0.09	1.14	0.01	0.08	153	0.12
6/10/97	4	Knik Arm	Vid.	30	5	26.50/8.06 (P/S)	2.03	0.19	1.18	0.02	0.09	60	0.21
6/10/97	5	Knik Arm	Obs.	1	1	na						2	0.42
6/10/97	6	Knik Arm	Obs.	5	2	na						9	0.42
6/9/98	1	Little Susitna	Vid.	59	1	18.00 (S)	1.18	0.45	1.13	0.02	0.21	78	0.5
6/10/98	1	Fire I	Obs.	11	8	13.25 (O)						21	0.52
6/10/98	2	Chickaloon	Vid.	14	5	12.38 (S)	1.62	0.31	1.16	0.04	0.09	27	0.33
6/10/98	3	Susitna	Vid.	75	3	16.36 (S)	1.45	0.19	1.23	0.02	0.12	139	0.23
6/10/98	4	Knik	Obs.	4	1	na						4	0.42
6/10/98	5	Knik	Obs.	11	4	18.00 (O)						21	0.39
6/10/98	6	Knik	Obs.	4	4	15.00 (O)						7	0.48
6/10/98	7	Knik	Obs.	4	5	26.67 (O)						8	0.41
6/10/98	8	Knik	Obs.	30	3	29.00 (O)						64	0.42
6/10/98	9	Knik	Obs.	6	2	31.00 (O)						9	0.5
6/10/98	10	Knik	Obs.	24	5	24.67 (O)						49	0.33
6/12/98	1	Little Susitna	Vid.	26	4	11.74 (S)	1.8	0.26	1.21	0.02	0.11	53	0.28
6/12/98	2	Knik	Obs.	5	5	16.75 (O)						9	0.89
6/12/98	3	Knik	Vid.	14	2	12.54 (S)	1.89	0.43	1.16	0.07	0.15	26	0.46
6/12/98	4	Knik	Vid.	10	3	11.20 (S)	1.93	0.51	1.16	0.05	0.12	18	0.53
6/12/98	5	Knik	Vid.	8	3	7.49 (S)	2.54	0.34	1.16	0.05	0.12	21	0.37
6/12/98	6	Knik	Obs.	9	4	64.25 (O)						19	0.66
6/12/98	7	Knik	Vid.	19	1	9.33 (S)	2.04	0.45	1.16	0.09	0.21	45	0.51
6/12/98	8	Chickaloon	Vid.	16	3	12.19 (S)	1.74	0.38	1.16	0.05	0.12	31	0.4
6/12/98	9	Chickaloon	Vid.	14	2	18.33 (S)	1.16	0.66	1.15	0.02	0.15	19	0.68
6/15/98	1	Chickaloon	Vid.	34	3	8.35 (S)	2.34	0.18	1.18	0.02	0.12	89	0.22
6/15/98	2	Little Susitna	Obs.	2	1	na						2	0.42
6/15/98	3	Little Susitna	Vid.	132	4	11.24 (S)	1.79	0.1	1.21	0.01	0.11	285	0.14
6/15/98	4	Knik	Obs.	12	3	na						21	0.42
6/15/98	5	Knik	Vid.	16	7	9.69 (S)	2.13	0.2	1.16	0.04	0.08	40	0.21
6/15/98	6	Knik	Vid.	2	2	8.20 (S)	2.29	1.07	1.16	0.07	0.15	4	1.08
6/15/98	7	Knik	Vid.	5	5	9.51 (S)	2.05	0.65	1.16	0.04	0.09	11	0.66
6/9/99	1	Little Susitna	Vid.	123	6	8.05(S)	2.62	0.03	1.08	0.00	0.17	314	0.17
6/9/99	2	Knik Arm	Obs.	23	3	na						55	0.11
6/9/99	3	Chickaloon	Vid.	16	1	11.60(S)	1.71	0.21	1.06	0.00	0.40	29	0.46
6/11/99	1	Little Susitna	Vid.	72	1	5.00(S)	3.21	0.07	1.06	0.00	0.40	245	0.41
6/12/99	1	Chickaloon	Vid.	1	4	7.25(S)	2.62	0.36	1.19	0.14	0.20	3	0.44
6/12/99	2	Chickaloon	Vid.	16	6	16.76(S)	1.63	0.09	1.14	0.02	0.17	30	0.19
6/12/99	4	Beluga R.	Vid.	7	1	14.00(S)	1.46	0.32	1.24	0.00	0.40	13	0.52
6/12/99	5	Big Susitna	Vid.	31	4	6.94(S)	2.73	0.06	1.06	0.00	0.20	92	0.21
6/12/99	6	Little Susitna	Vid.	69	1	7.33(S)	2.45	0.08	1.06	0.00	0.40	179	0.41
6/12/99	7	Knik Arm	Vid.	19	3	13.33(S)	1.69	0.11	1.20	0.00	0.23	39	0.26
6/13/99	1	Big Susitna	Vid.	16	2	14.53(S)	1.42	0.16	1.15	0.00	0.29	25	0.33
6/13/99	2	Little Susitna	Vid.	165	4	15.41(P)	1.46	0.04	1.13	0.00	0.20	258	0.21
6/13/99	4	Knik Arm	Vid.	15	5	22.16(P)	1.35	0.14	1.15	0.00	0.18	18	0.23
6/7/00	1	Knik	Vid.	18	3	10.00 (P/S)	1.87	0.10	1.34	0.15	0.24	44	0.30
6/7/00	2	Knik	Vid.	9	5	8.40 (S)	2.28	0.10	1.18	0.05	0.19	25	0.22
6/8/00	1	Little Susitna	Vid.	109	8	09.12 (P/S)	2.37	0.02	1.25	0.07	0.15	317	0.16
6/8/00	2	Little Susitna	Vid.	22	3	18.67 (P/S)	1.17	0.12	1.24	0.05	0.24	33	0.27
6/8/00	3	Chickaloon	Vid.	4	2	5.00 (S)	3.32	0.24	1.11	0.03	0.30	11	0.38
6/11/00	1	Little Susitna	Vid.	86	9	8.00 (S)	2.37	0.03	1.15	0.04	0.14	231	0.15
6/11/00	2	Beluga River	Vid.	1	5	12.40 (S)	1.77	0.35	1.27	0.09	0.19	2	0.41
6/12/00	1	Chickaloon	Vid.	11	8	9.25 (S)	2.28	0.08	1.26	0.11	0.15	31	0.20
6/12/00	2	Chickaloon	Vid.	4	5	6.60 (S)	2.86	0.16	1.13	0.04	0.19	11	0.25
6/12/00	3	Little Susitna	Vid.	131	8	7.50 (S)	2.52	0.02	1.11	0.02	0.15	357	0.15
6/12/00	4	Little Susitna	Vid.	7	1	4.00 (S)	3.61	0.22	1.08	0.02	0.42	27	0.47
6/12/00	5	Knik	Obs.	5	2	na						13	0.79
6/12/00	6	Knik	Vid.	1	1	6.00 (S)	3.00	0.61	1.16	0.09	0.42	3	0.74
6/12/00	7	Knik	Vid.	20	5	9.80 (S)	2.00	0.07	1.26	0.07	0.19	49	0.21
6/13/00	1	Chickaloon	Obs.	4	2	52.00 (O)						9	0.51
6/13/00	2	Big Susitna	Vid.	51	6	7.67 (S)	2.40	0.04	1.31	0.06	0.17	156	0.18
6/13/00	3	Knik	Obs.	7	2	53.50 (O)						18	0.32
6/13/00	4	Knik	Obs.	3	1	na						9	0.68
6/13/00	5	Knik	Vid.	8.6	5	6.00 (S)	2.88	0.10	1.14	0.08	0.19	28	0.22

recorded following the survey protocol. The remaining were either: 1) small groups (<10) encountered in areas of high aircraft traffic, where groups could not be circled due to safety concerns, or 2) groups encountered during non-survey flights (e.g. in support of vessel operations or as a part of adjunct experiments). Video of sufficient quality for group size estimates was available for 79 groups, averaging 3.5 (SD=2.1) counts per group. Average count times ranged from 4 sec to 13 sec for scan counts and from 6 sec to 65 sec for point counts. The remaining 65 group size estimates were derived from observer counts, averaging 4.7 (SD= 3.2) counts per group with count times ranging from 11 sec to 115 sec.

Linear regression, comparing observer counts to counts from video for the years 1994–98, indicated significant effects for each of the covariates: count ($\hat{b}_{1,o}$) and count multiplied by encounter rate ($\hat{b}_{2,o}$). All but one of the individual observer parameters for the count correction were significant, and for three observers, the parameter in the adjustment for encounter rate was significantly different from zero (Table 2). The video recordings in 1999 and 2000 were so successful that sizes for all but a few small groups were obtained from video. Sample sizes for comparison between the video-derived group sizes and observer counts within each year were sufficiently large so that each year was treated separately in the linear regression analysis (Table 2). Group sizes were then estimated for groups which had no group size available from video.

Video Counts, Time-at-the-surface, and Image Size

A total of 275 passes representing 79 groups were counted using aerial video recordings and a total of 2,775 whales were measured in the counting video. The average time-at-the-surface for all of the whales that were measured was 2.34 sec (SD=0.94, SE=0.018). This value, rounded to the nearest 0.5 sec ($T_s = 2.5$ sec), was used in the correction for whales missed under the water's surface. The average image size was 1.76 mm (SD=0.46, SE=0.009).

Table 2.—Parameters used to estimate beluga whale group sizes based on the observers' counts and the time spent counting each group. When there was no record of time spent counting, the respective aerial count was multiplied by the simple correction. $\hat{b}_{1,o}$ is the parameter used to estimate each observer's counting performance. $\hat{b}_{2,o}$ is the correction of counts as a function of group size. In 2000, none of the $\hat{b}_{2,o}$ parameters were significant, so only the simple corrections were used.

Years	Observer	Simple correction	SE	$\hat{b}_{1,o}$	SE	$\hat{b}_{2,o}$	SE
1994–98	1	2.01	(0.10)	1.80	(0.13)	0.07	(0.03)
	2	2.43	(0.28)	2.88	(1.61)	-0.60	(2.11)
	3	3.77	(0.38)	2.60	(0.65)	0.17	(0.08)
	4	2.84	(0.16)	2.60	(0.39)	0.12	(0.18)
	5	1.58	(0.07)	2.25	(0.28)	-0.18	(0.07)
	6	2.82	(1.64)	2.04	(0.58)	0.41	(0.29)
	7	1.43	(0.22)	1.45	(0.70)	0.00	(0.19)
	8	2.82	(0.29)	1.59	(0.79)	0.60	(0.36)
1999	1	2.68	(0.16)	2.07	(0.52)	0.24	(0.20)
	4	2.19	(0.11)	2.24	(0.33)	-0.02	(0.10)
	5	2.41	(0.11)	1.97	(0.23)	0.14	(0.07)
	8	2.80	(0.17)	1.56	(0.54)	0.56	(0.23)
2000	1	2.45	(0.10)				
	4	2.90	(0.09)				
	5	2.87	(0.16)				
	8	2.83	(0.14)				

Zoomed vs. Counting Video and Correction for Whales Missed at the Surface

Two of the three circling experiments provided enough zoomed images suitable for matching to the counting video. A whale group videotaped 13 June 1996 resulted in 103 usable whale images in the zoomed video, of which 91 were also seen in the counting video, 9 were not seen, and 3 were missed due to proximity to another whale that was counted. A group videotaped in 15 June 1998 resulted in 231 usable whale images in the zoomed video, of which 192 were also seen in the counting video, 29 were not seen, and 10 were missed due to proximity. The average whale image size between years was not significantly different (1996: 1.92 mm, SE=0.033 mm; 1998: 1.90 mm, SE=0.023), so the two data sets were combined into one size distribution.

Logistic regression estimated the probability that a whale image of a given size was seen as:

$$P(l) = \frac{e^{-13.72+9.56l}}{1 + e^{-13.72+9.56l}}$$

Thus an image of 1.43 mm in size would have a 50% chance of being seen in the counting video (Fig. 3). There was a poor correspondence between the gray-scale codes given to the images in the zoomed and counting videos. Gray

belugas always appeared gray, but white belugas in the zoomed video would sometimes appear gray in the counting video when their image size was small. Therefore, gray-scale codes were not used to stratify this analysis, and the ratio of gray to white images in the counting video is not considered to be representative of the ratio of juveniles to adults in the population. The correction for missed whales at the surface, by average image size is fairly constant for larger values ($\mu > 3$ mm) but climbs quickly as the average image size declines below 2.5 mm (Fig. 4). This correction was not considered valid for values greater than 1.5 ($\mu < 1.63$ mm); however, there were no passes counted from videotape when average sizes were below this cutoff.

For groups counted from video using the Canon Hi-8 video camera (1994–98) this correction averaged 1.17 (SD=0.04) and ranged from 1.10 to 1.35 (Table 1). The component of this correction that corrected for whales missed due to image size averaged 1.13 (SD=0.04) and ranged from 1.06 to 1.30. The component related to whales missed due to proximity in the counting video was 1.04 (CV=0.01).

Zoomed video of groups during counting passes in 2000 was of sufficient quality for comparison to the counting video (with the Sony Digital 8 camera). A total of 24 passes on 11 groups resulted in 170 usable whale

images in the zoomed video; of these, 154 were also seen in the counting video, 13 were not seen, and 3 were missed due to proximity to another whale that was counted. Accordingly, a logistic regression estimate of the probability that a whale image of a given size was seen was:

$$P(l) = \frac{e^{-5.77+7.00l}}{1 + e^{-5.77+7.00l}}$$

Thus an image of 0.834 mm in size would have a 50% chance of being seen in the counting video (Fig. 3b). Gray scale codes did not correspond sufficiently well between the images in the zoomed

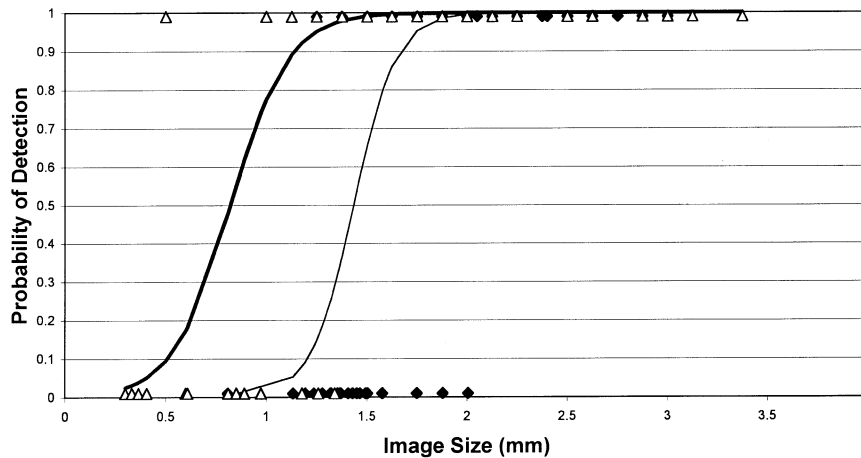


Figure 3.—Estimated probability from logistic regression of detection in the counting video by image size, with image sizes used in the analysis. Detected images are given a value of 1 and undetected images are given a value of 0. Diamonds represent 1994–98 data and triangles represent 1999–2000 data. The dark line is the probability of detection in the video from 1994–98, the light line is the probability of detection in the video from 1999–2000.

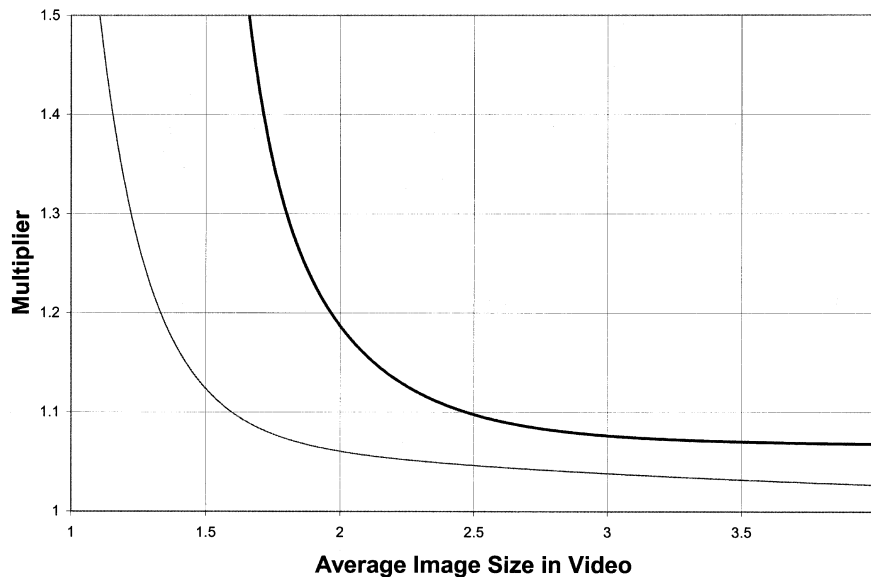


Figure 4.—Correction factors for beluga whales at the surface that were missed in the counting video, determined as a function of average image size of whales in the video counts for each pass of a group. This includes whales missed due to proximity to other whales and missed due to small image size. The dark line is the correction for the video from 1994–98, the light line is the correction for the video from 1999–2000.

and counting videos and therefore were not used to stratify this analysis.

The correction for missed whales at the surface, by average image size is fairly constant for larger values ($\mu > 2$ mm) but climbs quickly as the average image size declines below 1.5 mm (Fig. 4). This correction was not considered valid for values greater than 1.5 ($\mu < 1.10$ mm); however, there were no passes counted from videotape where average sizes fell below this cutoff. For groups counted from video using the Sony Digital 8 video camera (1999–2000), this correction averaged 1.17 (SD=0.08) and ranged from 1.06 to 1.34 (Table 1). The component of this correction that corrected for whales missed due to image size averaged 1.15 (SD=0.08) and ranged from 1.04 to 1.31. The component related to whales missed due to proximity in the counting video was 1.018 (CV=0.01).

Correction for Whales Missed Below the Surface in Video

Of the 275 passes counted from video, 30 were point counts, 189 were scan counts, and 56 were a mixture of scan and point. Point counts averaged 18.4 sec and ranged from 6 sec to 66 sec; scan counts averaged 9.7 sec and ranged from 4 sec to 30 sec. Correction factors for individual passes averaged 2.03 (SD=0.64) and ranged from 0.88 to 3.62 per whale at the surface.

Group Size Estimates

Group size estimates ranged from 1 to 731 (Table 1). CV's for group size estimates from video averaged 0.36 and ranged from 0.09 to 1.10. CV's for group size estimates from observer counts averaged 0.45 and ranged from 0.11 to 1.15. The largest component of the CV's of the group size estimates from video is CV(A), the correction for availability, averaging 0.28, which is largely determined by the uncertainty in the dive interval.

Discussion

Video-derived group size estimates proved to be a valuable tool for examining observer counts made from an aircraft. When confronted with large

groups, it was difficult for observers to quickly see and count each individual whale. Therefore, observers made quick estimates by tallying whales as best as they could.

Data from video counts indicate that for these large groups, too many whales were present for observers to mentally register each whale. Often, biased counts were made depending on the response of individual observers to this difficulty. The observer counts of relatively small groups (< 50 whales) were often larger than counts from the video but not larger than the corrected video group size estimates, suggesting that observers had more time to count individuals in small groups, and they could search a larger area than the video could record at any one time while scanning.

In making corrections for missed whales in the video analysis, it was necessary to have a measure of the portion of a group of whales that were actually visible in the video. Of the two ways that whales could be missed, it was far more likely that they were unavailable to be videotaped due to submersion than undetectable due to the resolution of the video camera.

The average correction factor for availability bias was 2.03 (SD=0.64), thus more than half of the whales in a typical group were missed because they were underwater, and in some groups more than 2/3 of the group were missed. By comparison, the average correction for whales missed at the surface was 1.17, or roughly 1/7 of the whales available to be recorded were missed. On the surface it would seem that a higher resolution video system would reduce these two corrections by providing: 1) a wider field of view, which would allow for longer scan times and more whales to surface in the view, and 2) a crisper image so that small whales would be easier to detect, and whales in close proximity would form more distinct images. Such a system was incorporated in the aerial survey program in 1999 and 2000 when it started using a digital video camera. This resulted in an improved resolution of over 40% when comparing the estimated probabilities by size in Figure 3, but it did not significantly change the average correction

factors. What did change was the success rate of the video data collection, resulting in 84% of the group size estimates being made from video in 1999 and 2000, whereas 46% of estimates were made from video during 1994–98.

The measure of time-at-the-surface was necessary as a component in the correction factor for whales missed because they were underwater during the count. We found differences in surface times between whale groups, although two passes from the same group had similar surface times. There are several possible reasons for these differences. The behavior of the animals may have been different between groups (or in the same group on different days). Whales may have faster surfacing times when they are active (feeding or traveling) than when at rest, or they may have different surfacing times while feeding on different prey types (such as herding small fish in dense schools compared to targeting large, single fish). However, we were unable to determine the behavior of a whale group from the air, other than that they were clumped or spread out.

Surface times may also be different between groups due to differences in distance from the aircraft. The farther the group was from the aircraft, the smaller the whales appeared. They would, therefore, be less distinct when measuring surface times. The brightness of the day and the sea condition may also affect how well animals were seen. These weather conditions affect how whales contrast against the background water gray-scale. There could also have been a difference in the brightness settings of the video monitor during analysis. However, variation of the average surface times of passes and groups was small in comparison to the uncertainty in the dive interval and the half-second resolution of the counting time, so it was not necessary to estimate this separately for each pass.

The variability of dive intervals of individuals is problematic for variance calculations. Individual behavior may be correlated if all or most of the members of a group are engaged in the same activity; this component of the variance could remain constant regard-

less of group size. An example would be a group that was feeding and consequently taking longer dives; each individual would be diving for a longer time within its own range of variability. The variance of the average dive interval among individuals would still decline with the inverse of group size. It is unknown to what degree the dive behavior of individuals in a group is correlated. We do not see obvious patterns during aerial surveys, so we assume that—although small subgroups may have correlated behavior—the overall effect is negligible.

The gray-scale codes that were determined for the counting video did not correspond well to the codes given to the corresponding zoomed images. Zoomed images were a better record of the true shade of the whale, so the lack of good correspondence to the gray-scale codes of the counting video indicated that the counting video was not a useful record of the ratio of gray to white in a group of beluga whales (i.e. the ratio of juveniles to adults). The most likely reason is that sizes of whale images in the counting video are close to the limits of the resolution of the video system. Higher resolution in the new digital video system made these gray-scale codes more reliable, but in the size range where whales were most likely to be missed, there was little improvement.

Despite conscientious effort by observers to count as many whales as they could, it is apparent from the magnitude of the correction for each observer's group size estimates that a significant number of beluga whales were missed during aerial counts even among well-trained observers. The magnitude of this correction is considerably larger with untrained observers. The most likely explanation is that each observer can only effectively count in an area somewhat smaller than the area covered by a large group. The time spent counting any portion of the group was then only an unknown fraction of the time spent counting the whole group.

The McLaren formula was not appropriate for use with these data because the time available to count an individual whale was not well determined

(except in the video analysis). Whale groups were often spread over several kilometers of water, so the time spent counting was sometimes as much as a minute or more, but no area of water was in view for more than several seconds. Also, observers have noticed that when whale density was high, the effective field of view for counting was reduced, decreasing the time available to count each individual whale.

Using video analysis to develop correction factors resolved this problem because there was no limit to the amount of time available to study the image. A correction specific to each observer allowed for variation in the area that each observer searched during an aerial count and the different responses of individual observers to high densities of whales.

A stranding of a group of beluga whales on a mud flat during low tide on 12 June 1996 allowed a rare opportunity to compare a precise count from aerial video recording of the group when stranded (63) against the typical, systematic aerial observer and video counts (racetrack method) made after the tide returned, and the whales were able to swim away. The resultant corrected count from aerial video of the swimming whales (69; Table 1, group 3) compares well to the precise count (63) obtained from the video of the group when stranded. The average of observer counts for this group (31) after it was waterborne again provides an anecdotal correction factor estimate of 2.2 (= 69/31), which is consistent with the correction factors for observer counts obtained from the video passes used in the abundance estimates (Table 2).

Other authors have applied correction factors to calculate beluga abundance from aerial counts. Brodie (1971) used a factor of 1.4 based on observed diving behavior (not including hard-to-see juvenile whales). Sergeant (1973) estimated that belugas in fairly turbid water would be visible for about one-third of the time (i.e. a factor of 3), based on observations of whales one day at the mouth of the Churchill River. Fraker (1980) used a factor of 2 to account for belugas invisible beneath the surface, but he considered this “largely

arbitrary.” Kingsley (1998), in a review of surveys of belugas in the St. Lawrence River, used a minimum availability correction of 1.15 for whales that were not at the surface in aerial photographs, but this was to improve the conformity with visual estimates, not a correction for absolute abundance.

The most substantial correction factor, 2.75, was developed by Frost et al. (1985) through results from VHF tags kept 2 weeks on two beluga whales in Bristol Bay and the assumption that aerial observers have 10 sec to search an area. However, their technique was designed to correct for whales missed in small groups while making a single transect pass through a sample area.

Although the Cook Inlet surveys have continued to operate on straight flightlines when counting whales, the multiple fly-overs (“racetracks”) allow ample time to determine the extent of each group. This provides a better counting situation as observers can concentrate their search on the known location of the respective group. The amount of time spent counting each group—typically 20–60 sec—is recorded and can be included in calculations of abundance.

In conclusion, use of video to count groups of beluga whales removes some of the uncertainty associated with aerial observer counts. The area of view is well defined, and the time spent counting an area can be precisely measured. Little subjectivity is involved in interpreting the images, and counts from video recordings are highly repeatable among trained video analysts. Associated variables necessary for correction factors, such as time-at-the-surface and average image size, can be easily measured. Yet, the average coefficient of variation for a single count is over 20%. A portion of this is the result of the binomial variation associated with the correction factors for missed animals; the remainder may be due to variations in group behavior that we are as yet unable to identify from the air.

Acknowledgments

We thank Robyn Angliss, Douglas DeMaster, Barbara Mahoney, Kim Shelden, and Laura Litzky for operating the video camera during aerial sur-

veys. Lisa Baraff, Robin Bush, Suzanne Harkness, Laura Litzky, Marcia Muto, Stephanie Norman, and Kim Shelden counted beluga whales from videotape. Suzanne Harkness helped develop the zoomed video analysis methods and organized and trained video analysts for the 1998 data analysis. Funding for this project was provided by the NMFS Marine Mammal Assessment Program. This research was conducted under Permit No. 791 and 782-1438 issued by the National Marine Fisheries Service. Reviews of this manuscript were provided by Robyn Angliss, Jeff Breiwick, Jeff Laake, Sally Mizroch, Sue Moore, Kim Shelden, and two anonymous reviewers.

Literature Cited

- Barlow, J., C. W. Oliver, T. D. Jackson, and B. L. Taylor. 1988. Harbor porpoise, *Phocoena phocoena*, abundance estimation for California, Oregon, and Washington: II. Aerial surveys. *Fish. Bull.* 86:433–444.
- Brodie, P. F. 1971. A reconsideration of aspects of growth, reproduction, and behavior of the white whale (*Delphinapterus leucas*), with reference to the Cumberland Sound, Baffin Island, population. *J. Fish. Res. Board Can.* 28:1309–1318.
- Fraker, M. A. 1980. Status and harvest of the Mackenzie stock of white whales (*Delphinapterus leucas*). *Rep. Int. Whal. Comm.* 30:451–458.
- Frost, K. J., L. F. Lowry, and R. R. Nelson. 1985. Radiotagging studies of belukha whales (*Delphinapterus leucas*) in Bristol Bay, Alaska. *Mar. Mammal Sci.* 1(3):191–202.
- Hobbs, R. C., D. J. Rugh, and D. P. DeMaster. 2000. Abundance of belugas, *Delphinapterus leucas*, in Cook Inlet, Alaska, 1994–2000. *Mar. Fish. Rev.* 62(3):37–45.
- Kingsley, M. C. S. 1998. Population index estimates for the St. Lawrence belugas, 1973–1995. *Mar. Mammal Sci.* 14(3):508–530.
- Laake, J. L., J. Calambokidis, S. D. Osmeck, and G. W. Garner. 1997. Probability of detecting harbor porpoise from aerial surveys: Estimating $g(0)$. *J. Wildl. Manage.* 61:63–75.
- Laidre, K. L., K. E. W. Shelden, D. J. Rugh, and B. A. Mahoney. 2000. Beluga, *Delphinapterus leucas*, distribution and survey effort in the Gulf of Alaska. *Mar. Fish. Rev.* 62(3):27–36.
- Lerczak, J. A., K. E. W. Shelden, and R. C. Hobbs. 2000. Application of suction-cup-attached VHF transmitters to the study of beluga, *Delphinapterus leucas*, surfacing behavior in Cook Inlet, Alaska. *Mar. Fish. Rev.* 62(3):99–111.
- Mahoney, B. A., and K. E. W. Shelden. 2000. Harvest history of belugas, *Delphinapterus leucas*, in Cook Inlet, Alaska. *Mar. Fish. Rev.* 62(3):124–133.
- Marsh, H., and D. F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *J. Wildl. Manage.* 53(4):1017–1024.
- McLaren, I. A. 1961. Methods of determining the numbers and availability of ringed seals in the

- eastern Canadian Arctic. *Arctic* 14(3):162–175.
- Moore, S. E., K. E. W. Shelden, L. K. Litzky, B. A. Mahoney, and D. J. Rugh. 2000. Beluga, *Delphinapterus leucas*, habitat associations in Cook Inlet, Alaska. *Mar. Fish. Rev.* 62(3):60–80.
- Rugh, D. J., K. E. W. Shelden, and B. A. Mahoney. 2000. Distribution of belugas, *Delphinapterus leucas*, in Cook Inlet, Alaska, during June/July 1993–2000. *Mar. Fish. Rev.* 62(3):6–21.
- Sergeant, D. E. 1973. Biology of white whales (*Delphinapterus leucas*) in western Hudson Bay. *J. Fish. Res. Board Can.* 30:1065–1090.