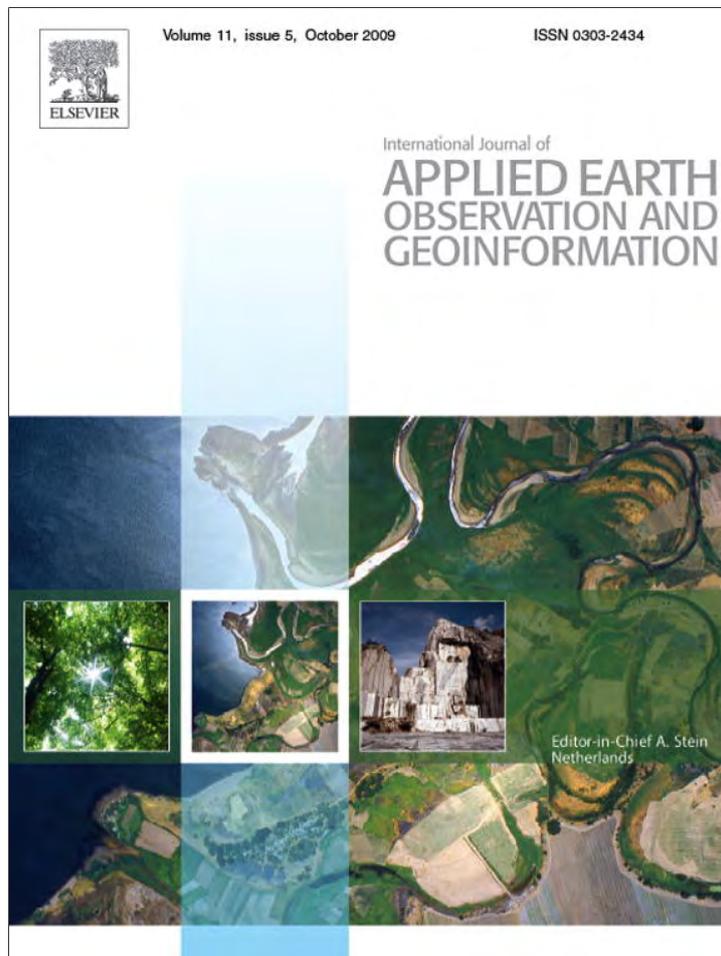


Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

An improved procedure for detection and enumeration of walrus signatures in airborne thermal imagery

Douglas M. Burn^{a,*}, Mark S. Udevitz^b, Suzann G. Speckman^a, R. Bradley Benter^a^aU.S. Fish and Wildlife Service, Marine Mammals Management Office, 1011 East Tudor Road, Anchorage, AK 99503, United States^bU.S. Geological Survey, Alaska Science Center, 4210 University Drive, Anchorage, AK 99508, United States

ARTICLE INFO

Article history:

Received 3 November 2008

Accepted 22 May 2009

Keywords:

Airborne thermal imagery

Pacific walrus

Aerial survey

Alaska

Bering Sea

ABSTRACT

In recent years, application of remote sensing to marine mammal surveys has been a promising area of investigation for wildlife managers and researchers. In April 2006, the United States and Russia conducted an aerial survey of Pacific walrus (*Odobenus rosmarus divergens*) using thermal infrared sensors to detect groups of animals resting on pack ice in the Bering Sea. The goal of this survey was to estimate the size of the Pacific walrus population. An initial analysis of the U.S. data using previously-established methods resulted in lower detectability of walrus groups in the imagery and higher variability in calibration models than was expected based on pilot studies. This paper describes an improved procedure for detection and enumeration of walrus groups in airborne thermal imagery.

Thermal images were first subdivided into smaller 200 × 200 pixel “tiles.” We calculated three statistics to represent characteristics of walrus signatures from the temperature histogram for each tile. Tiles that exhibited one or more of these characteristics were examined further to determine if walrus signatures were present. We used cluster analysis on tiles that contained walrus signatures to determine which pixels belonged to each group. We then calculated a thermal index value for each walrus group in the imagery and used generalized linear models to estimate detection functions (the probability of a group having a positive index value) and calibration functions (the size of a group as a function of its index value) based on counts from matched digital aerial photographs.

The new method described here improved our ability to detect walrus groups at both 2 m and 4 m spatial resolution. In addition, the resulting calibration models have lower variance than the original method. We anticipate that the use of this new procedure will greatly improve the quality of the population estimate derived from these data. This procedure may also have broader applicability to thermal infrared surveys of other wildlife species.

Published by Elsevier B.V.

1. Introduction

The current size and trend of the Pacific walrus (*Odobenus rosmarus divergens*) population is unknown, and recent changes in the Bering and Chukchi sea ecosystems (e.g., Hunt and Stabeno, 2000; Maslanik et al., 2007) increase the need for a reliable technique to monitor the status of this population. Between 1975 and 1990, visual aerial surveys were carried out by the United States and the former Soviet Union at 5 years intervals, producing population estimates for Pacific walruses ranging from about 200,000 to 230,000 animals. Observers counted or estimated numbers of walruses hauled out on pack ice and land, but could not accurately detect or enumerate walruses that were swimming in the water. Surveyed areas were limited to less than 5% of available

habitat. The population estimates generated from these surveys are considered minimum values that cannot be used for detecting trends in population size (Hills and Gilbert, 1994; Gilbert et al., 1992). Efforts to estimate the size of the Pacific walrus population were suspended after 1990 due to these and other unresolved problems with survey methods, which produced population estimates with unknown biases and unknown, but presumably low, precision (Gilbert et al., 1992; Gilbert, 1999).

A workshop on walrus survey methods, hosted by the U.S. Fish and Wildlife Service (USFWS) and U.S. Geological Survey (USGS), concluded that it would not be possible to obtain a population estimate with adequate precision for tracking trends using the existing visual methodology and any feasible amount of survey effort (Garlich-Miller and Jay, 2000). Workshop participants recommended exploring new survey tools, including remote sensing systems, prior to conducting another aerial survey. Remote sensing systems have the potential to address many of the shortcomings of visual aerial surveys by sampling larger areas per

* Corresponding author.

E-mail address: Douglas_Burn@fws.gov (D.M. Burn).

unit of time (Garlich-Miller and Jay, 2000), objectively detecting and quantifying walrus, and reducing observer error (Burn et al., 2006).

One of the first applications of remote sensing technology to marine mammal surveys was the Bering Sea Marine Mammal Experiment (BESMEX) conducted in the mid-1970s (Wartzok and Ray, 1980), which demonstrated the utility of thermal imagery for detecting walrus groups on ice floes. More than a decade later, Barber et al. (1991) used a forward-looking infrared radiometer (FLIR) system to determine the relation between the number of walrus in a group and the size (in pixels) of the group in a corresponding thermal image. More recently, thermal imagery has also been used to detect and count harp seals (*Pagophilus groenlandicus*) in the White Sea, Russia (V. Chernook, GiproRybFlot, unpublished data).

Burn et al. (2006) further demonstrated the feasibility of using airborne thermal imagery to detect and enumerate walrus groups as they rest on the pack ice in the Bering Sea. In order to determine the limits of this technique, Burn et al. (2006) collected thermal imagery at spatial resolutions ranging from 1 m to 4 m per pixel. Digital photographs of a subset of walrus groups were taken concurrently. Analysis of matched photographs and thermal images indicated there was a linear relation between the number of walrus in a group and the amount of heat they produced (Burn et al., 2006). This relation existed across all spatial resolutions tested, indicating that the number of walrus in a group on sea ice could be estimated using their thermal signatures.

Based on the successful results of the Burn et al. (2006) study, a pilot survey was conducted in the area of St. Lawrence Island in the Bering Sea in 2003 (Udevitz et al., 2008). The pilot survey obtained infrared images of nearly 30,000 km² of sea ice habitat, an area larger than that covered in any previous visual aerial survey of Pacific walrus, and provided the first size estimate of a walrus population based on an infrared survey (Udevitz et al., 2008).

In spring of 2006, the USFWS, in collaboration with the USGS and the Russian institutes GiproRybFlot and ChukotTINRO, conducted a range-wide survey to estimate the size of the Pacific walrus population. The survey used thermal imagery and the methods of Burn et al. (2006) and Udevitz et al. (2008) to estimate the number of walrus hauled out on sea ice. The study also used satellite-linked tags to record the haul-out status of individual walrus and estimate the proportion of the population in the water (Jay et al., 2006).

Initial analysis of the data collected in the U.S. portion of the survey area revealed that a large number of photographed walrus groups were not detected in the thermal infrared imagery. In addition, many of the groups that were detected appeared to have spatial footprints that were much smaller than their corresponding aerial photographs, which likely underestimated the magnitude of their thermal signatures. The initial calibration models for these data had large variances that would result in population estimates with unacceptably low precision. These unexpected results forced us to re-examine our image processing methodology and develop a more robust procedure for detecting walrus groups on sea ice in airborne thermal imagery. Here, we describe the new methodology and compare it to the original image processing techniques developed by Burn et al. (2006).

2. Methods

2.1. Study area and data collection

In late winter and early spring, Pacific walrus are found in the Bering Sea pack ice where open leads, polynyas, or thin ice occur (Fay et al., 1984). Walrus use floating pack ice as a substrate for

birthing, nursing, resting, and for passive transport to new feeding areas. Although capable of diving to deeper depths, walrus usually feed in shallow waters of 100 m or less (Fay, 1982; Fay and Burns, 1988). The survey targeted the extent of Bering Sea pack ice where the sea floor depth is less than 200 m. Under these criteria, all potential spring walrus habitat was included in the survey. Airborne thermal infrared surveys of Pacific walrus were conducted in the Bering Sea beginning on April 4, 2006, and ending on April 22, 2006. With the exception of a single reconnaissance flight on April 15, 2006, all flight operations were conducted south of the Bering Strait, and north of Nunivak Island, Alaska (Fig. 1). All surveys were conducted over U.S. territorial waters where pack ice conditions ranged from 50 to 100% total concentration.

Survey operations were conducted with two aircraft. The first was an Aero Commander 690B turbine engine aircraft. This aircraft contained the thermal infrared (8.5–12.5 μm) scanner with a 0.625 mrad instantaneous field of view. The system, built by Argon ST (Ann Arbor, Michigan) was equipped with a 3000 pixel detector array and had 12-bit radiometric resolution and absolute sensitivity of 0.12 °C. The system also included a position and orientation system (POS) to georeference the thermal imagery (Applanix Corp., Richmond Hill, Ontario, Canada).

Survey transects were oriented north-south and ranged in length from 60 to 225 km. Initial survey operations for the scanner aircraft were conducted at 6400 m above ground level (AGL), producing imagery with 4 m pixel size. On April 19, 2006, we conducted surveys at both 6400 m and 3200 m AGL to collect imagery with both 4 m and 2 m pixel sizes, respectively. Survey operations on April 21, 2006, and April 22, 2006, were conducted at 3200 m AGL.

The second aircraft was an Aero Commander 680 piston engine aircraft equipped with a vertical camera port. We photographed walrus groups with a high-resolution digital SLR 12.4 megapixel camera that produced images with dimensions of 4288 × 2848 pixels. Photographs were taken from a nominal altitude of 700 m AGL using an image-stabilized 200 mm f2.8 camera lens and 1.4× teleconverter. Each photograph included latitude and longitude coordinates determined by a Global Positioning System (GPS). The objective was to photograph as many walrus groups as possible within the strip surveyed by the thermal scanner within one hour of scanning, to minimize the effect of changes in group size over time. Depending on the time differential between collection of thermal imagery and aerial photography, wind and ocean currents created a slight spatial offset between the two data sources. This offset ranged from a few hundred meters up to 2 km, but was relatively constant in direction and magnitude for each thermal image and therefore did not affect our ability to match photographs with thermal images.

2.2. Data processing

We imported the thermal infrared imagery using a custom software application (Rapid Mapper) developed by Argon ST. This program integrates the thermal data and POS information to create georeferenced thermal images in Universal Transverse Mercator (UTM) projection. We used ERDAS Imagine (Leica Geosystems, Atlanta, GA) software for initial data visualization and export to ASCII format. Sensor artifacts (i.e., temperature values that were impossibly high or low) were re-coded to missing values before the data were processed with the original methods presented by Burn et al. (2006) and with the new methods presented in this paper. The following Sections describe the new methods, with the differences between the new methods and those of Burn et al. (2006) outlined in Section 2.2.4. The same procedures were used for processing both the 2 m and 4 m resolution thermal images.

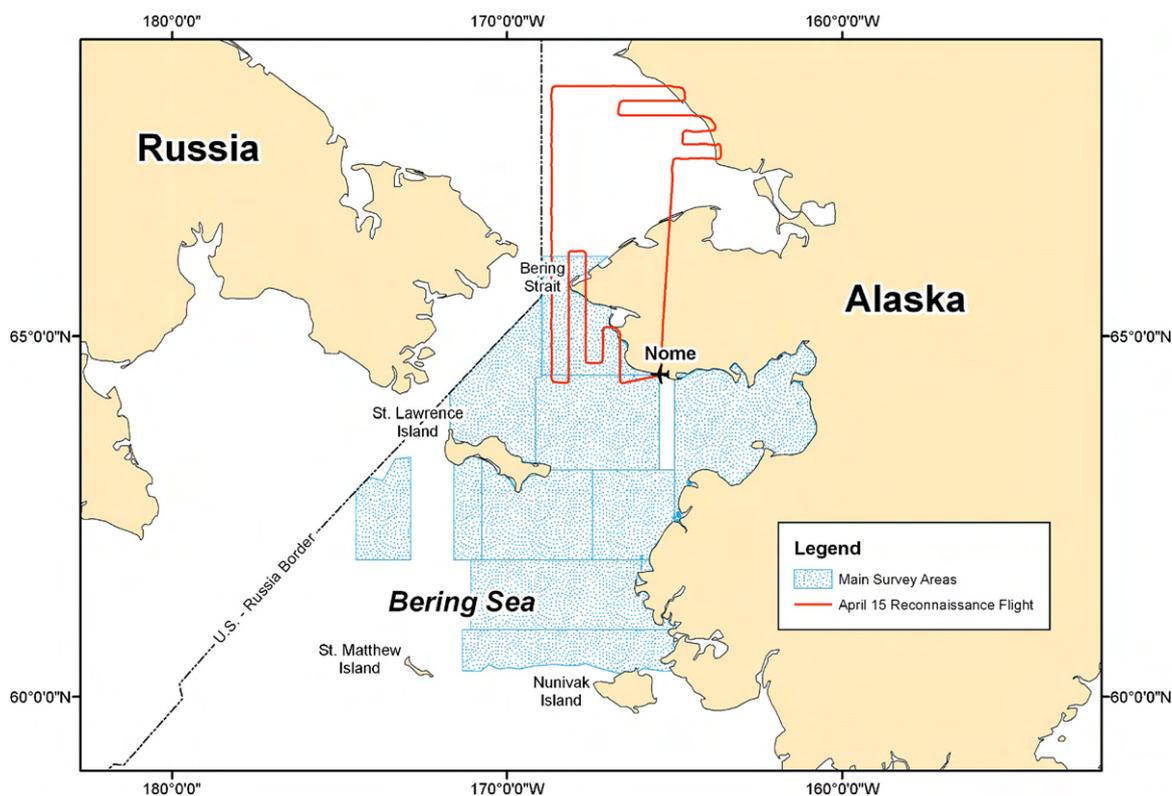


Fig. 1. Survey area for airborne thermal imagery data collection.

2.2.1. Counting walrus in photographed groups

Each photograph was overlaid on its corresponding thermal image to match each walrus group with its thermal signature (Fig. 2a and c; Burn et al., 2006). A walrus group was considered distinct from other groups if their corresponding thermal signatures were separated by one or more pixels (2–4 m, depending on resolution). The number of walrus in each photographed group was counted using ERDAS Imagine. Each photographed walrus group was counted three times by the same analyst, who marked each walrus with a uniquely colored symbol. If the three counts were not identical, we simultaneously displayed the symbols for all three counts and made a fourth count to rectify differences. To ensure that groups of walrus in photographs reflected the same groups that were recorded by the thermal scanner, only walrus hauled out completely on an ice floe were counted.

2.2.2. Detecting walrus groups

Each thermal image was subdivided into a series of 200×200 pixel “tiles” (Fig. 2a). The tiles covered an area 800 m on a side in 4 m imagery, and 400 m on a side in 2 m imagery. Depending on the orientation of each transect in UTM projection, tiles at the edge of each thermal image were of irregular size. Any edge tile with less than half the number of pixels of a full tile (40,000 pixels) was merged with an adjacent full tile. The temperature value for each pixel was rounded to the nearest tenth of a degree Celsius to create a temperature histogram for the pixels in each tile.

We derived three statistics from the temperature histogram for each tile in each thermal image: (1) maximum temperature; (2) length of right-hand tail, calculated as the difference between the maximum temperature and the warmest histogram bin with a frequency of 10 or more pixels; (3) maximum gap between histogram values (Fig. 3). Temperatures near maximum for a tile are characteristic of thermal signatures of walrus groups because walrus are typically the warmest objects in their immediate

environment. Long right-hand tails and large gaps are also characteristics of walrus thermal signatures because walrus are relatively rare features, typically present in less than 0.1% of the pixels in a tile.

We determined lower threshold values for each of these three parameters based on tiles that contained photographed walrus groups. Tiles were then assigned a set of three scores based on the values of these parameters relative to their threshold values. Tiles with a maximum temperature that exceeded the threshold value were given a score of 4. Tiles with a right-hand tail value that exceeded the threshold value were given a score of 2, and those that had a maximum gap value that exceeded the threshold were given a score of 1. Scores were set to 0 for each parameter that did not exceed its threshold value. The three scores were then summed to give a total score, which could range from 0 to 7 for each tile.

We eliminated from further consideration any tiles with total scores of 0. We then examined the data for each of the remaining tiles in greater detail, focusing on the spatial arrangement of the warmest pixels and their degree of contrast with adjacent pixels. Features consisting of warm pixels that corresponded to such features as open leads and rock faces along the shoreline could be easily eliminated based on visual inspection of the images. Walrus groups are typically located on thicker ice floes that register colder temperatures and therefore tend to be represented by pixels that have a high degree of contrast with adjacent pixels. We used these characteristics to identify which of the remaining tiles contained pixels that corresponded to walrus groups.

2.2.3. Calculating walrus thermal index values

Next, we determined which pixels belonged to each detected walrus group. We used a disjoint cluster analysis to assign every pixel in each tile to one of 10 clusters, based on their locations within the tile (row and column coordinates) and temperatures. The cluster algorithm made assignments by minimizing Euclidian

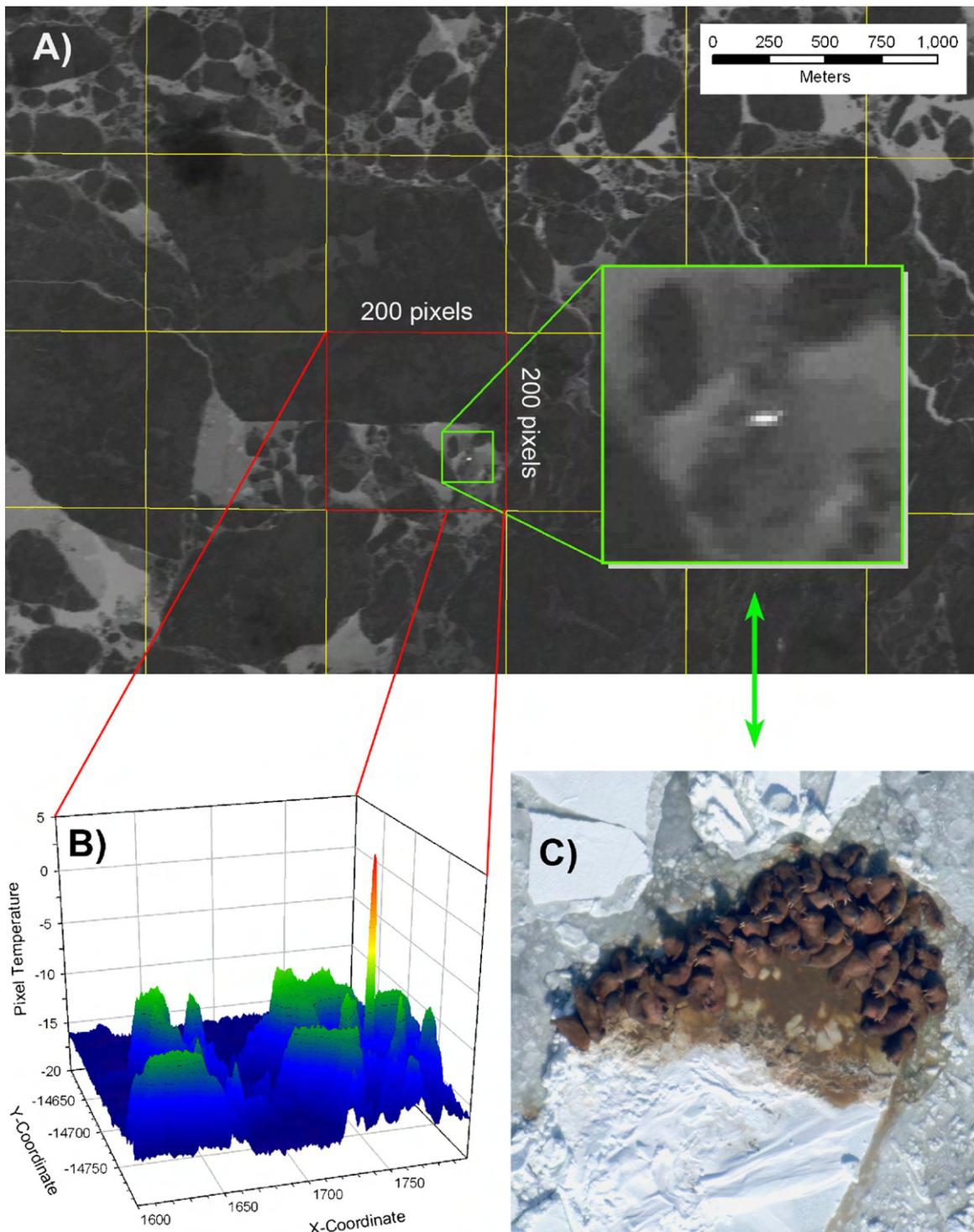


Fig. 2. (A) Airborne thermal image indicating tile structure and location of photographed walrus group (green square); (B) three-dimensional plot of data for selected (red) tile; (C) photograph of walrus group located in selected tile.

distances, relative to these three normalized variables, among pixels in the same cluster (Anderberg, 1973). By definition, the warmest cluster in a tile was always included as part of a group. However, in some cases, walrus groups consisted of more than one contiguous cluster. Clusters were ranked in order of their mean temperatures. We then added additional clusters to walrus groups until a cluster was more similar to the next coldest cluster (background) than it was to the next warmest cluster (previously designated as belonging to a walrus group).

We defined two scalar thermal indices for each walrus group. The first index (h_1) was similar to the index used by Burn et al. (2006) and was calculated by summing the temperatures of all the pixels in the walrus group and subtracting the temperature of the warmest non-walrus pixel in the tile. However, depending on whether there was thin ice or open water present within a tile, the threshold temperatures of the warmest non-walrus pixels were highly variable from tile to tile within a thermal image. We therefore calculated a second thermal index (h_2) for each walrus

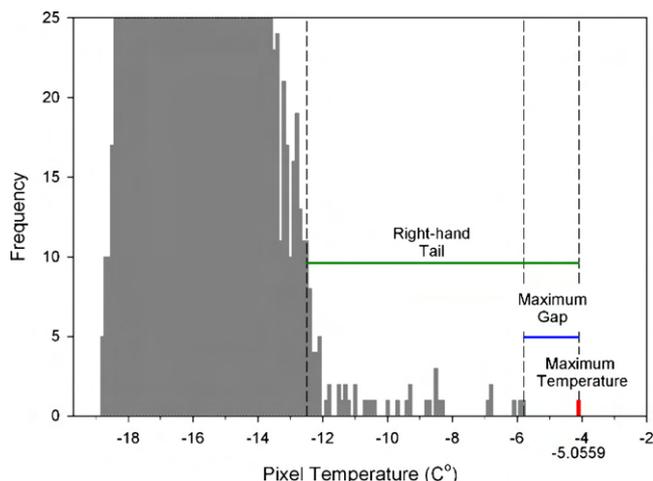


Fig. 3. Temperature histogram indicating features characteristic of walrus signatures: (1) maximum temperature (red); (2) maximum histogram gap (blue); (3) right-hand tail (green).

group as the sum of the temperatures for all pixels in the group minus the modal temperature for all non-walrus pixels in the tile. Using the modal temperature of each tile had the effect of standardizing the index relative to the local ambient temperature, thereby reducing overall variability of the relation between the index and the number of walruses in a group.

2.2.4. Comparison with original methods

We also processed both the 2 m and 4 m data using the original methods described in Burn et al. (2006) to detect walrus groups, match them with their corresponding aerial photographs, and calculate their thermal index values. In that study, walrus groups were first visually located by the pilots or observers, and digitally photographed from an altitude of 457 m. After several groups had been photographed and their locations recorded by GPS, they collected thermal imagery of those groups at 1 m, 2 m, 3 m, and 4 m resolution. Walrus group size was determined using the counting procedures outlined in Section 2.2.1. Thermal imagery was processed by developing a temperature histogram for each image, and level-slicing it at the temperature value where the histogram frequency decreased sharply from thousands of pixels to fewer than 10 pixels (Burn et al., 2006). The detected walrus groups were differentiated in the thermal imagery if they were separated by a distance of 20 m or more. Burn et al. (2006) calculated their thermal index (h_0) by summing the temperatures of all the pixels in a walrus group and subtracting the temperature of the warmest non-walrus pixel in the entire image.

The major differences between the two methods were: (1) the criterion used to differentiate walrus groups in the thermal imagery; (2) treatment of each image as a single unit (original method) vs. subdividing the image into tiles (new method); (3) detection of walrus groups by level-slicing the temperature histogram (original method) vs. cluster analysis (new method); (4) calculating the thermal index relative to the threshold temperature of the entire image (h_0) or tile (h_1), or relative to the modal temperature for the tile (h_2).

2.3. Estimating detection probabilities

We used logistic regression (Hosmer and Lemeshow, 2000) to estimate probabilities of detecting walrus groups in the thermal images, based on the data from all photographed groups. A separate analysis was conducted at each resolution, for the Burn

et al. (2006) detection algorithm and for the new detection algorithm. For both detection algorithms, we considered group size and the log of group size as possible predictors. Using the log of group size allowed the probability of detection to increase more slowly with increasing group size than with the untransformed variable. For the new algorithm, we also considered the modal ice temperature for the tile (ice temperature) and the log of ice temperature as possible predictors. Equivalent ice temperature variables were not available for the Burn et al. (2006) algorithm because images were not partitioned into tiles in that algorithm. For each algorithm, at each resolution, we fit models containing all combinations of the predictors, except that both transformed and untransformed versions of a predictor were not used in any single model. Models were fit with maximum likelihood. We used Akaike's Information Criterion (AIC, Burnham and Anderson, 2002) to select the final detection model for each algorithm, at each resolution. We evaluated the fit of the final models with the Hosmer–Lemeshow goodness-of-fit test (Hosmer and Lemeshow, 2000).

2.4. Calibrating thermal indices

We developed calibration models to estimate the number of walruses in each group based on its thermal index, given that it was detected in the thermal image (i.e., it had a positive thermal index). For calibration, we only used observations of photographed groups that were detected by the thermal scanner, because the calibration models are conditional on the group being detected. We again conducted separate analyses for each of the two detection algorithms at both 2 m and 4 m resolutions.

For the Burn et al. (2006) algorithm, there were no ice temperature variables and only one thermal index (h_0) to consider. Therefore, for this algorithm, we considered only calibration models that were functions of h_0 .

For the new algorithm, we had ice temperature values and two thermal indices (h_1 and h_2). We considered calibration models that included functions of h_1 with and without ice temperature, and models that were functions of h_2 . Because the h_2 index was based on the modal ice temperature, we did not consider models that included functions of ice temperature along with h_2 .

We estimated calibration models using generalized linear models (McCullagh and Nelder, 1999) with identity links. For each calibration model, we considered Normal, Poisson, negative binomial, and gamma distributions for fitting the error structure. Models were fit with maximum likelihood. We used AIC (Burnham and Anderson, 2002) to select the final calibration model for each algorithm, at each resolution. We used deviance and deviance residuals to assess the fit of the final models (McCullagh and Nelder, 1999).

3. Results

During the April 2006 aerial survey of Pacific walrus, we collected a total of 63 thermal images at 4 m resolution, and an additional 21 images at 2 m resolution. We also collected photographs of 124 unique walrus groups from the areas that were scanned at 4 m resolution, and photographs of 85 unique walrus groups from the areas that were scanned at 2 m resolution. Photographed walrus groups ranged from 1 to 446 in size (mean = 27).

At 4 m resolution, 25% of the detected walrus groups had summary tile scores of 7, meaning that the tile exceeded the threshold values for maximum temperature, right-hand tail, and maximum gap (Table 1). We also detected eight groups ranging in size from 9 to 70 walruses by the maximum gap value alone, and four groups ranging from 26 to 93 walruses by the right-hand tail

Table 1
Summary tile scores for walrus groups that were photographed and present in thermal imagery.

Resolution (m)	Summary tile score	Maximum temperature	Right-hand tail	Maximum gap	N groups	Mean group size	Minimum group size	Maximum group size
4	7	+	+	+	31	58	9	446
	5	+	–	+	12	22	10	34
	3	–	+	+	12	22	7	49
	2	–	+	–	4	48	26	93
	1	–	–	+	8	31	9	70
	0	–	–	–	57	9	2	24
2	7	+	+	+	70	26	2	168
	6	+	+	–	1	10	10	10
	0	–	–	–	14	3	1	6

Summary tile scores of zero indicate groups that were not detected. Plus symbol (+) indicates threshold value exceeded for the parameter; minus symbol (–) indicates threshold value not exceeded for the parameter. All group sizes were determined by photographic counts.

Table 2
Summary of detected walrus groups in airborne thermal imagery at 2 m and 4 m resolution.

Algorithm	Resolution (m)	Number of photographed groups	Number of photographed groups detected	Size of smallest detected group	Size of largest undetected group
Burn et al. (2006)	2	80	61	2	26
	4	112	35	12	93
New	2	85	71	2	6
	4	124	67	7	24

The number of photographed groups was slightly larger for the new algorithm, as this method uses a different criterion for distinguishing walrus groups. All group sizes were determined by photographic counts.

value alone. All of these groups with a summary tile score of 1 or 2 had gone undetected using the original method of Burn et al. (2006). In the 2 m imagery, nearly all detected walrus groups had summary tile scores of 7.

The new procedure detected more photographed walrus groups than the original method in both the 4 m and 2 m thermal imagery (Table 2). In addition, this method was capable of detecting smaller walrus groups in 4 m thermal imagery than the methods of Burn et al. (2006). The sizes of the largest undetected walrus groups were also smaller using the new method.

3.1. Estimating detection probabilities

Group size was strongly related to detection probability for both algorithms at both resolutions (Table 3). For the Burn et al. (2006) algorithm, models based on the log of group size fit substantially better ($\Delta AIC > 2.2$) than those based on untransformed group size. For the new algorithm, there was essentially no difference ($\Delta AIC < 1.4$) between fits of models with transformed and untransformed group size variables.

Although ice temperature (or its log) was included in some of the best models for the new algorithm, it did not substantially

Table 3
AIC values for considered detection models.

Algorithm	Resolution (m)	Model structure ^a	Number of parameters	AIC	ΔAIC^b
Burn et al. (2006)	2	log(size)	2	51.20	0.00
		Size	2	53.51	2.30
		Null	1	89.71	38.51
	4	log(size)	2	88.86	0.00
		Size	2	101.24	12.39
		Null	1	141.12	52.27
New	2	Size	2	32.63	0.00
		log(size)	2	33.99	1.36
		Size, temperature	3	34.62	1.99
		Size, log(temperature)	3	34.62	1.99
		log(size), temperature	3	35.97	3.35
		log(size), log(temperature)	3	35.98	3.36
		Temperature	2	76.27	43.65
		log(temperature)	2	76.77	44.15
		Null	1	78.06	45.43
	4	Size	2	81.23	0.00
		Size, log(temperature)	3	81.31	0.08
		log(size)	2	81.84	0.61
		log(size), log(temperature)	3	82.21	0.97
		Size, temperature	3	82.96	1.73
		log(size), temperature	3	83.71	2.48
		log(temperature)	2	171.01	89.78
		Null	1	173.09	91.86
		Temperature	2	174.26	93.03

^a Null model include only an intercept. All other models include an intercept and the listed variables.

^b ΔAIC is the difference between the AIC value for the specified model and the model with the lowest AIC value.

Table 4
Parameter estimates for final selected detection models.

Algorithm	Resolution (m)	Coefficient (SE)	
		Intercept	Group size
Burn et al. (2006)	2	-3.63 (1.09)	2.18 ^a (0.51)
	4	-8.52 (1.70)	2.56 ^a (0.53)
New	2	-3.65 (1.32)	0.85 (0.27)
	4	-4.14 (0.77)	0.27 (0.05)

^a Coefficient of log(group size).

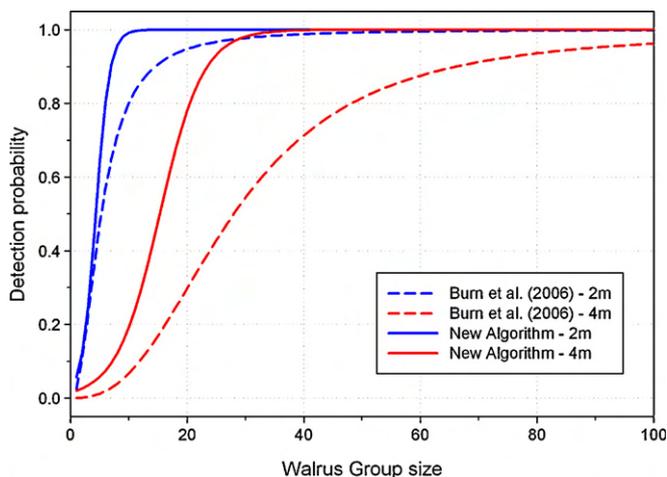


Fig. 4. Detection models for walrus groups in infrared imagery using the new algorithm and the Burn et al. (2006) algorithm at 2 m and 4 m resolutions.

improve models that already included group size (Table 3). The estimated linear effect of ice temperature was negative in all of the 2 m models that included it, but positive in the corresponding 4 m models, suggesting that there was insufficient data to reliably estimate this effect.

The final selected models for estimating detection probabilities were functions of only group size or log of group size (Table 4; Fig. 4). These models had the form:

$$Y_i \sim \text{Bernoulli}(p_i),$$

where

$$\text{logit}(p_i) = \beta_0 + \beta_1 X_i,$$

Y_i is a binary variable indicating whether group i was detected or not, and X_i is the size (or log of size) of group i . Hosmer–Lemeshow tests did not indicate lack of fit for any of the final models ($P \geq 0.45$).

Table 5
AIC values for considered negative binomial calibration models.

Algorithm	Resolution (m)	Model structure ^a	Number of parameters	AIC	ΔAIC^b
Burn et al. (2006)	2	h0	3	474.13	0.00
		Null	2	543.14	69.01
	4	h0	3	317.02	0.00
		Null	2	348.99	31.97
New	2	h2	3	481.59	0.00
		h1 temperature	4	514.00	32.41
		h1	3	515.44	33.85
		Null	2	606.38	124.79
	4	h2	3	513.26	0.00
		h1	3	545.09	31.83
		h1 temperature	4	545.49	32.23
		null	2	630.08	116.82

^a Null model includes only an intercept. All other models include an intercept and the listed variables.

^b ΔAIC is the difference between the AIC value for the specified model and the model with the lowest AIC value for the same algorithm and resolution.

3.2. Calibrating thermal indices

For both algorithms at both resolutions, calibration model variances increased with the means. Negative binomial and gamma models fit this variance structure substantially better than Normal ($\Delta\text{AIC} > 38.0$) or Poisson ($\Delta\text{AIC} > 240.0$) models. There was very little difference in AIC values for negative binomial and gamma models, with negative binomial models having slightly lower AIC values for the 2 m models (best models for the Burn et al. (2006) algorithm $\Delta\text{AIC} > 2.2$, new algorithm $\Delta\text{AIC} > 1.2$) and essentially no difference for the 4 m models (best models $\Delta\text{AIC} < 0.9$). Based on this, we confined further consideration to negative binomial models.

At both resolutions, the h0 and h1 indices had linear relations to group size (Table 5, Fig. 5). For the new algorithm, the h2 index also had a linear relation to group size and models based on h2 were substantially better than those based on h1 (Table 5). Adding ice temperature did not substantially improve any of the models based on h1 (Table 5). Based on this, we selected the linear models in h0 for the Burn et al. (2006) algorithm and the linear models in h2 for new algorithm calibrations. Examination of deviance and deviance residuals did not indicate any lack of fit for any of the final models.

Final calibration models had the form:

$$Y_i \sim \text{Negative Binomial}(\eta_i, k),$$

where

$$\eta_i = \beta_0 + \beta_1 X_i,$$

Y_i is the size of group i , X_i is the thermal index value for group i , and $\text{var}(Y_i) = E(Y_i) + kE(Y_i)^2$,

where k is the dispersion parameter. Parameter estimates for the final calibration models are presented in Table 6, along with estimates for models based on the h1 index for the new algorithm for comparison. Comparison of estimated dispersion parameters (Table 6) and variance functions (Fig. 6) indicates that variances were larger for models based on the h0 and h1 indices than those based on the h2 index. This is consistent with the lower AIC values for the h2-based models. Variances at the 2 m resolution were not much different than those at the 4 m resolution for the new algorithm, but variances were substantially lower at the 2 m resolution than those at the 4 m resolution for the Burn et al. (2006) algorithm. For models based on the h0 and h1 indices, variances were substantially lower for the new algorithm than for the Burn et al. (2006) algorithm.

4. Discussion

The new procedure for detection and classification of walrus groups in thermal imagery described in this paper marks a

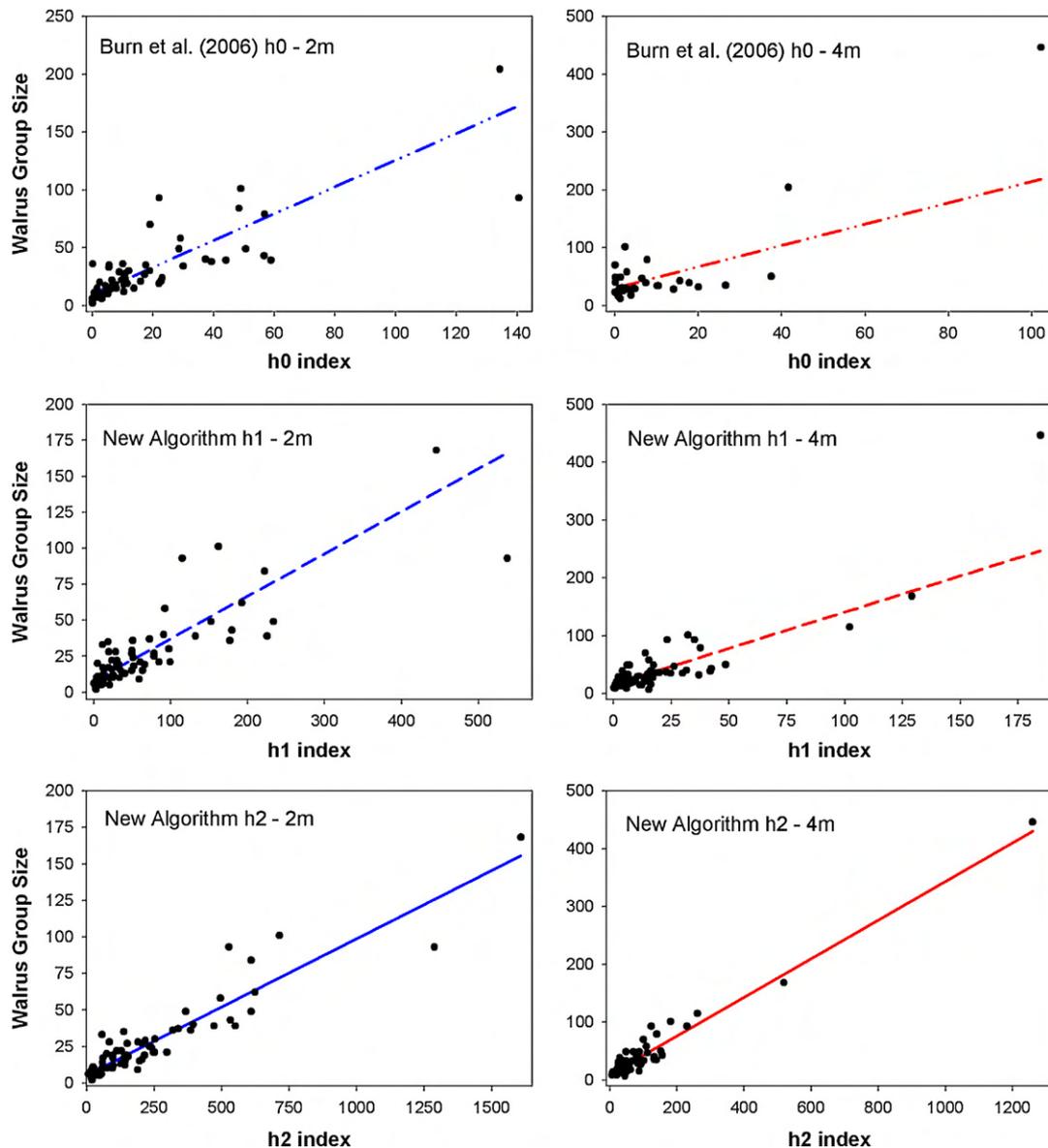


Fig. 5. Calibration models for estimating walrus group size as a function of thermal indices h_0 , h_1 and h_2 in infrared images at 2 m and 4 m resolutions.

significant improvement over the methods previously developed by Burn et al. (2006). Smaller groups were detectable in both 4 m and 2 m imagery using this procedure, and the resulting calibration models based on the h_2 index were substantially more precise than the model for the original h_0 index. Use of the new image processing procedure should result in a substantial improvement in both accuracy and precision of the population estimate from the 2006 survey (Speckman et al., 2009).

We believe the main reason for the poor performance of the Burn et al. (2006) image processing procedure was the colder ambient temperatures encountered during the 2006 survey effort compared to temperatures encountered during the 2003 pilot study. Air temperatures recorded at the base of operations in Nome, Alaska, averaged 5–8 °C colder on survey days in 2006 compared to the conditions encountered in 2003 (Burn et al., 2006). Walrus skin temperature is known to vary with ambient temperature (Ray and Fay, 1968). At extremely cold temperatures, walrus vasoconstrict to reduce blood flow to the skin as a mechanism for conserving body heat. In addition, colder air temperatures are accompanied by colder ice temperatures. For “mixed” pixels that contained both ice and walrus, the sensor

would therefore average colder walrus skin with even colder ice, which shifted those pixels to the left in the temperature histogram. This effect is demonstrated by the greater improvement in the detection function at 4 m resolution. An adult Pacific walrus can reach lengths of up to 3 m, so there are likely fewer mixed pixels in 2 m resolution thermal imagery. In fact, groups of 7 or more walrus were always detected at this finer scale.

The differences in physical size of the images also may have limited the effectiveness of the Burn et al. (2006) method applied to the 2006 survey data. During the earlier (Burn et al., 2006) study, walrus were first located visually and photographed before thermal imagery was collected. Those images were therefore considerably smaller than the ones collected in 2006, which typically had over 100 million pixels. By tiling these large images and looking for characteristic walrus signatures, we were able to identify features that were distinct on a local level, but may have been masked when examining the image as a single unit. For example, the original method did not detect any walrus groups during survey operations on April 10, 2006, despite visual observation of numerous small walrus groups made from the photography aircraft. Using the new method, we were able to

Table 6
Parameter estimates for final selected calibration models.

Algorithm	Resolution (m)	Thermal index	Coefficient (SE)		Dispersion parameter (SE)
			Intercept	Thermal index	
Burn et al. (2006)	2	h0	10.08 (1.47)	1.16 (0.15)	0.19 (0.04)
	4	h0	30.29 (3.64)	1.84 (0.47)	0.25 (0.06)
New	2	h1 ^a	7.79 (1.06)	0.29 (0.03)	0.16 (0.04)
		h2	5.34 (0.83)	0.09 (0.01)	0.08 (0.02)
	4	h1 ^a	15.90 (1.93)	1.25 (0.17)	0.16 (0.03)
		h2	9.91 (1.70)	0.33 (0.03)	0.09 (0.02)

^a These were not selected models, but are included here for comparison.

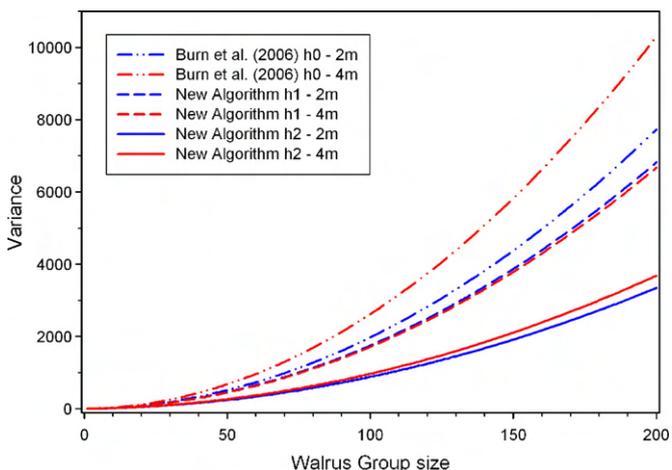


Fig. 6. Variance functions for thermal index calibration models using h0, h1 and h2 indices at 2 m and 4 m resolutions.

detect 12 walrus groups that had previously been missed on that survey date. Although we did not have photographs of these groups, the area where they were detected corresponded to the area where many small groups were seen from the photography aircraft.

The most common non-walrus features that had high summary tile scores (i.e., “false positives”) were open leads that extended only slightly into a tile that otherwise contained only ice. These tiles typically did not exceed the maximum temperature threshold, and were easily identified with visual inspection of the neighboring tiles to verify that the feature was an open lead. The process of visually inspecting all tiles with summary scores greater than zero could be improved by adding an additional step that centers the warmest pixel within a new tile and then re-examines the temperature histogram. Walrus groups would still be relatively rare features, whereas open leads would not, likely resulting in low values for the right-hand tail and maximum gap parameters and a reduction in the number of false positives.

An additional feature that created false positives that exceeded the maximum temperature threshold occurred when survey lines intersected land. Bare rock faces, especially those facing south, typically had relatively warm temperatures. These features were also easy to identify by overlaying GIS data layers for land onto the thermal images.

During the process of developing this new method, we also examined a variety of other techniques that were unsuccessful. Regardless of the ambient temperature, walrus always have a high degree of thermal contrast in comparison to the ice floes upon which they rest. We initially investigated a suite of kernel operations (Russ, 2002) including edge detection, standard deviation, variance, and minimum-maximum difference using both 3 × 3 and 5 × 5 kernel sizes. Results with these techniques

were equivocal, and in many cases they did not detect photographed walrus groups that were detected using the original methods of Burn et al. (2006). Performance of the kernel operations may have been affected by the high degree of variability in walrus group sizes, which may range from 1 to more than 100 pixels in 4 m imagery. In very large groups that contain multiple “pure” walrus pixels, there was often relatively little contrast in the center of the group.

The use of airborne thermal imagery to survey marine mammal populations has the potential to sample considerably larger areas per unit time than visual or photographic surveys. The procedure described in this paper would be directly applicable to the Atlantic walrus (*O. rosmarus rosmarus*), and may also have broader applicability to surveys of other ice-associated pinnipeds, including harp (*P. groenlandicus*), ribbon (*Phoca fasciata*), spotted (*Phoca largha*), ringed (*Phoca hispida*), and bearded (*Erignathus barbatus*). These species do not form dense concentrations similar to walrus, and their detection would probably require smaller pixel sizes, which would reduce the amount of area that could be surveyed per unit time. A pilot study similar to that of Burn et al. (2006) using the image processing procedures we describe here may help identify the sampling resolution that optimizes the compromise between the ability to detect animals in the imagery and the amount of area that can be surveyed.

Acknowledgements

This study was conducted under USFWS permit No. MA039386-0 and Department Management Authority (DMA) Letter of Confirmation No. MA-039582.

Aircraft support was provided by Commander Northwest (Wenatchee, Washington). We thank Andy Harcombe, Jeremy Weintraub, and John Sumero for their piloting skills and camaraderie. This paper is dedicated to the memory of Ralph Aiken, chief pilot of the scanner aircraft, computer wizard, and friend—he is deeply missed by all who knew him.

Tom Ory and Fred Osterwisch from Argon ST worked closely with the authors in development and operation of the scanner system to meet our survey needs. William Baker operated the thermal imaging system and drums for all survey operations.

The U.S. Marine Mammal Commission provided funding for the Applanix Position and Orientation System. The Municipality of Anchorage donated high-resolution satellite imagery to aid in calibration of the thermal sensor.

Tom Heinlein, Bering Land Bridge National Preserve, provided logistical support for the aerial survey team while in Nome, Alaska. The National Weather Service in Nome provided weather information and logistical support. Jan Bennett and Lark Wuerth of the Department of Interior, Office of Aircraft Services, provided flight-following of both aircraft during survey operations.

The authors also thank Steve Fleischman, Anupma Prakash, and Mike Simpkins for reviewing an earlier version of this manuscript. Some of the concepts in this paper were the result of a conversation

between the first author and his brother, David Burn, at our parents' kitchen table.

Reference to brand name products does not imply endorsement by the Federal Government.

References

- Anderberg, M.R., 1973. Cluster Analysis for Applications. Academic Press, New York.
- Barber, D.G., Richard, P.R., Hochheim, R.P., Orr, J., 1991. Calibration of aerial thermal infrared imagery for walrus population assessment. *Arctic* 44 (Suppl. 1), 58–65.
- Burn, D.M., Webber, M.A., Udevitz, M.S., 2006. Application of airborne thermal imagery to surveys of Pacific walrus. *Wildlife Society Bulletin* 34, 51–58.
- Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference. New York, Springer.
- Fay, F.H., 1982. Ecology and biology of the Pacific walrus, *Odobenus rosmarus divergens* Illiger. *North American Fauna* 74, 1–279.
- Fay, F.H., Burns, J.J., 1988. Maximal feeding depth of walruses. *Arctic* 41 (3), 239–240.
- Fay, F.H., Kelly, B.P., Gehrnich, J.L., Sease, J.L., Hoover, A.A., 1984. Modern populations, migrations, demography, trophics, and historical status of the Pacific walrus. Final Report R.U. #611. NOAA Outer Continental Shelf Environmental Assessment Program, Anchorage, AK.
- Garlich-Miller, J., Jay, C.V., 2000. Proceedings of a Workshop Concerning Walrus Survey Methods. USFWS R7/MMM Technical Report 00-2. 92 pp.
- Gilbert, J.R., Fedoseev, G.A., Seagars, D., Razlivalov, E., LaChugin, A., 1992. Aerial census of Pacific walrus, 1990. USFWS R7/MMM Technical Report 92-1. 33 pp.
- Gilbert, J.R., 1999. Review of previous Pacific walrus surveys to develop improved survey designs. In: Garner, G.W., Amstrup, S.C., Laake, J.L., Manly, B.F.J., McDonald, L.L., Robertson, D.G. (Eds.), *Marine Mammal Survey and Assessment Methods*. A.A. Balkema, Rotterdam, pp. 75–84.
- Hills, S., Gilbert, J.R., 1994. Detecting Pacific walrus population trends with aerial surveys—a review. In: *Transactions of the Fifty-ninth North American Wildlife and Natural Resources Conference*. pp. 201–210.
- Hosmer, D.W., Lemeshow, S., 2000. *Applied Logistic Regression*, second ed. John Wiley and Sons, New York.
- Hunt Jr., G.L., Stabeno, P.J., 2000. Ecosystem responses to weather patterns in the Southeastern Bering Sea. *Arctic Research of the United States* 14, 17–24.
- Jay, C.V., Heide-Jørgensen, M.P., Fischbach, A.S., Jensen, M.V., Tessler, D.F., Jensen, A.V., 2006. Comparison of remotely deployed satellite radio transmitters on walruses. *Marine Mammal Science* 22, 226–236.
- Maslanik, J.A., Fowler, C., Stroeve, J., Drobot, S., Zwally, J., Yi, D., Emery, W., 2007. A younger, thinner Arctic ice cover: increased potential for rapid, extensive sea-ice loss. *Geophysical Research Letters* 34, L24501.
- McCullagh, P., Nelder, J.A., 1999. *Generalized Linear Models*, second ed. Chapman and Hall, Boca Raton.
- Ray, C., Fay, F.H., 1968. Influence of climate on the distribution of walruses, *Odobenus rosmarus* (Linnaeus). II. Evidence from physiological characteristics. *Zoologica* 53, 19–32.
- Russ, J.C., 2002. *The Image Processing Handbook*, fourth ed. CRC Press, Boca Raton.
- Speckman, S.G., Chernook, V., Burn, D.M., Udevitz, M.S., Kochnev, A., Vasilev, A., Benter, R.B., Lisovsky, A., 2009. Range-wide survey of Pacific walruses in 2006: estimated number of walruses on sea ice. Progress Report. U.S. Fish and Wildlife Service and U.S. Geological Survey, Anchorage, AK, 55 pp.
- Udevitz, M.S., Burn, D.M., Webber, M.A., 2008. Estimation of walrus populations on sea ice with infrared imagery and aerial photography. *Marine Mammal Science* 24, 57–70.
- Wartzok, D., Ray, G.C., 1980. The hauling-out behavior of the Pacific walrus. U.S. Marine Mammal Commission Report MMC-75/15, Bethesda, MD.