esa ECOSPHERE

Monitoring marine mammals using unmanned aerial vehicles: quantifying detection certainty

Ana S. Aniceto,[1,2,]† Martin Biuw,[3] Ulf Lindstrøm,[3] Stian A. Solbø,[4] Fredrik Broms,[5] and JoLynn Carroll[1,2]

[1]*Akvaplan-niva, Fram Centre – High North Research Centre for Climate and the Environment, 9296 Tromsø, Norway*
[2] *Department of Geosciences, ARCEx – Research Centre of Arctic Petroleum Exploration, UiT The Arctic University of Norway, Dramsveien 201, P.O. Box 6050 Langnes, 9037 Tromsø, Norway*
[3]*Institute of Marine Research – Sykehusveien 23, P.O. Box 6404, 9294 Tromsø, Norway*
[4]*Northern Research Institute (NORUT) – SIVA Innovation Centre, Sykehusveien 21, P.O. Box 6434, 9294 Tromsø, Norway*
[5]*Fram Centre – The High North Research Centre for Climate and the Environment, P.O. Box 6606, Langnes, N-9296 Tromsø, Norway*

Citation: Aniceto, A. S., M. Biuw, U. Lindstrøm, S. A. Solbø, F. Broms, and J. Carroll. 2018. Monitoring marine mammals using unmanned aerial vehicles: quantifying detection certainty. Ecosphere 9(3):e02122. 10.1002/ecs2.2122

Abstract. Unmanned aerial vehicles (UAVs) are increasingly being recognized as potentially useful for detection of marine mammals in their natural habitats, but an important consideration is the associated uncertainties in animal detection. We present a study based on field trials using UAVs to carry out image-based monitoring of cetaceans in two fjords in northern Norway. We conducted 12 missions to assess the effects of both environmental- and aircraft-related variables on detection certainty. Images were inspected for animal presence and its associated detection certainty. Images were also assessed for potentially important covariates such as wave turbulence (sea state), luminance, and glare. Aircraft variables such as altitude, pitch, and roll were combined into a single variable—pixel size. We recorded a total of 50 humpback whales, 63 killer whales (KW), and 118 unidentified sightings. We also recorded 57 harbor porpoise sightings. None of the environmental conditions (sea state, glare, and luminance) affected the detection certainty of harbor porpoises. In contrast, increasing sea state and luminance had negative and positive effects, respectively, on the detection certainty of humpback and KW. The detection certainty was not significantly affected by pixel size for both harbor porpoises, and humpback and KW. Our results indicate that at lower altitudes, variations in aircraft position (pitch and roll) do not have a variable effect on detection certainty. Overall, this study shows the importance of measuring variability in both environmental and flight-related variables, in order to attain unbiased estimates of detectability for UAV-based marine mammal surveys, particularly in Arctic and sub-Arctic regions.

Key words: certainty; detection; marine mammals; survey; unmanned aerial vehicles.

Received 21 November 2017; revised 9 January 2018; accepted 16 January 2018. Corresponding Editor: Robert R. Parmenter.

† **E-mail:** asa@akvaplan.niva.no

Introduction

A major challenge for society is the management of Earth's biodiversity with the aim of protecting the integrity and health of species, their habitats, and ecosystems. Animals in their natural habitats have often proved challenging to study and manage due to their complex movement patterns and individual behaviors. This is particularly challenging for those species occupying habitats in remote regions where the use of traditional monitoring methods can be difficult, time-consuming, and expensive (Chabot and Bird 2015).

Visual monitoring for marine mammals is usually carried out from aerial, ship-based, or land-based survey platforms, each associated with its own advantages and drawbacks. Ship-based surveys are by far the most common for open-ocean abundance surveys, but are very time-consuming and costly. Aerial observations have the potential of covering large areas within relatively limited time periods, but are also costly and can be limited by flight regulations and safety considerations. The high velocity of aerial compared to ship-based surveys may also affect detection probability, depending on the diving/surfacing intervals of the animals under study in relation to the speed at which the platform passes overhead.

Irrespective of the type of platform, monitoring can be done either by direct human observation or through the collection of images or videos for post-processing. Sometimes, these are used in parallel, to combine the advantages of human observations for scanning larger regions with the advantages of later re-analysis and reassessment of images and videos. One main reason why Unmanned aerial vehicles (UAVs) are considered promising for animal surveys (Chabot and Bird 2015) is the potential for carrying out relatively large-scale aerial image-based surveys at often a fraction of the cost of manned aerial surveys, and without many of the safety issues associated with manned aircraft (Koski et al. 2009, Hodgson et al. 2017). Additionally, the low cost of UAV systems compared to manned aircraft may also allow greater flexibility in survey design, for instance by flying two or more platforms at specific time lags rather than employing the double-back maneuver (Hiby 1999). While the regulatory framework for UAVs is often similar to that for manned aerial flights, and their flight ranges have previously been limited by technology, these issues are gradually being resolved to facilitate wider implementation of this technology.

In the case of marine mammal monitoring, UAVs provide the ability to repeatedly collect high-resolution aerial imagery in a manner that is unobtrusive to animals. In addition, compared to manned surveys, UAVs may be used in areas where manned aerial operations are difficult and dangerous, such as in narrow fjords or in remote polar regions. These capabilities are particularly advantageous when performing studies focused on animal abundance and distribution. While UAVs may eliminate observer bias in the data collection phase (ability to detect an animal given it is available for detection), the behavior of the animals and the probability of detection in post-survey analyses still pose some challenges.

Environmental and survey-related variables, such as light conditions and wind, can affect detectability and must be considered during survey planning. In Arctic and sub-Arctic regions, the effectiveness of traditional survey methods can be severely limited by cold temperatures, strong winds, and seasonally low light levels. Such environmental conditions also place limitations on UAV surveys. For a given set of camera specifications, reduced visibility, mainly during the polar night, will greatly influence detection performance.

In this study, we take advantage of highly predictable and high abundances of several representative species in easily accessible fjord systems in northern Norway, which provide a unique opportunity to test the capabilities of UAVs for marine mammal monitoring under Arctic conditions, considering large-, medium-, and small-sized species. In recent years, large numbers of humpback whales (HW) have been visiting fjords in north Norway during winter; they make a stopover during their breeding migration (Broms et al. 2015, Ryan et al. 2015) to feed on high densities of over-wintering Norwegian spring-spawning herring (*Clupea harengus*). While no recent abundance estimates have been published for the area, studies involving the north Norwegian Humpback Whale catalogue have identified a minimum of 820 different whales during the winter months (November–January) in the regions of Troms and Nordland County between 2010 and 2017 (Broms 2017). These numbers correspond to approximately 82% of the estimated Barents and Norwegian Seas' population and indicate that a major proportion of the Barents Sea population utilizes the Norwegian fjords before migrating toward their breeding areas. In addition to HW, large numbers of killer whales (KW) visit the fjords in north Norway to exploit the abundant herring.

Harbor porpoises inhabit most of the Norwegian coastline. They are the smallest cetacean and are generally known for their inconspicuous behavior, making them hard to observe. They prefer coastal habitats such as fjords, inlets, reefs, straits, and gullies (Watts and Gaskin 1985, Johnston et al. 2005, Goodwin 2008, Pierpoint 2008). Harbor porpoises in open coastal regions and oceans (e.g., North Sea) are normally monitored using manned aerial surveys. However, in narrow fjord systems along the coast, where harbor porpoise numbers are expected to be high, it is not recommended to use manned airplanes due to safety reasons (e.g., unpredictable winds). In such areas, autonomous or unmanned aircrafts may be a suitable alternative, since they are capable of operating at lower altitudes and at larger range from airports (a current safety limitation of manned aircraft). We present survey results obtained through a series of field trials designed to test the applicability of UAVs in marine mammal surveys. We performed a series of flight trials in locations with known animal presence in order to (1) identify key factors affecting the reliability in detecting whales in UAV images and to (2) understand the effects of UAV movement on detection certainty in Arctic waters. Based on our results for our selection of marine mammal species, we further discuss how survey and environmental parameters can be improved in future studies of cetacean conservation and management, particularly for the estimation of marine mammal abundances. We also suggest an approach for including modeled detection certainty in future abundance estimates from UAV-based surveys.

Materials and Methods

We selected two locations based on the practicality of UAV deployment and recovery and the regular and predictable presence of marine mammals in these areas. Surveys were performed in two northern Norwegian fjords: Kaldfjord and Rystraumen (Fig. 1). These survey sites were located around Kvaløya, an island in the Tromsø municipality. The Norwegian name of this island means "whale island," which may reflect historical whale abundance around this island.

Kaldfjord (69.7° N, 18.7° E) is a 16 km long fjord, with a maximum depth of 237 m (Fig. 1).

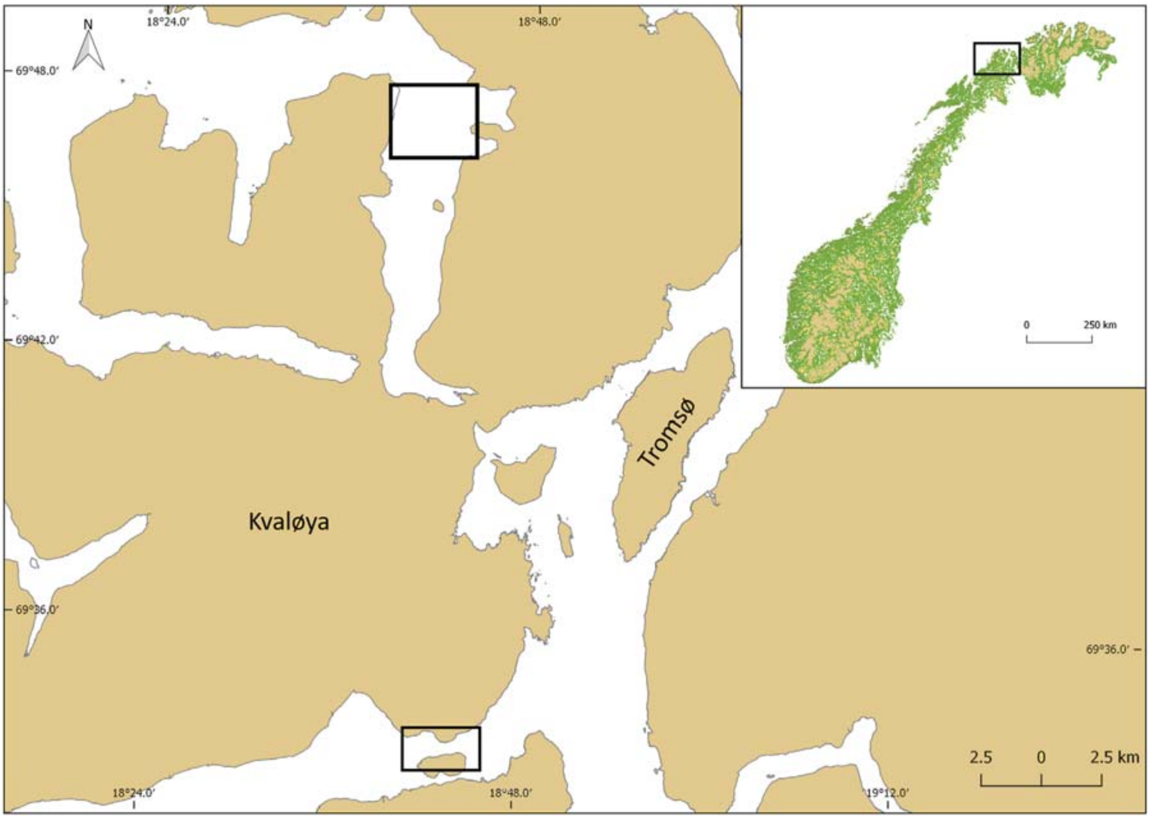

Fig. 1. Study areas in northern Norway. Humpback and killer whales were surveyed in Kaldfjord (top), and harbor porpoises were surveyed in Rystraumen (bottom).

This location was selected as it has been a region of high whale abundance during the winter months in recent years, and it allows for easy access by car to suitable launch sites and for good survey coverage over a large area. Most of Kaldfjord is observable from Ytterkollen (elevation = 95 m above sea level, an observation point located next to the Bellvik harbor (69°47′8.02″ N, 18°44′18.64″ E). This location allowed observers to confirm the presence of animals prior to the deployment of the UAVs, a requirement for these surveys. During the winter months, northern Norway experiences the polar night, where the sun does not rise over the horizon between 27 November and 15 January. During this period, light conditions are severely restricted and highly dependent on cloud cover. While clear weather days provide good light conditions for efficient drone flights for 4–5 h, during overcast days, workable light conditions are restricted. Sampling took place during winter 2014/2015 in three sections of the fjord. These sections were selected based on the distribution and density of humpback and KW in the fjord during the winter season.

Rystraumen (69.5° N, 18.7° E) is a strong tidal current channel in the Straumsfjorden strait that separates the island of Kvaløya from the mainland (Fig. 1). This strait has a minimum width of 500 m and is relatively shallow (50 m). From the Kvaløya side, it is possible to observe the entire strait. This area is well known as having a regular presence of harbor porpoises, presumably driven by prey availability due to the strong tidal currents and associated water mixing.

Unmanned aerial vehicle data

We developed and applied a protocol based on the collection of still-images with an overlap in coverage of the survey area. We included this overlap in the survey design in order to avoid gaps in the search strip and to take into account double sightings during data analyses. We deployed different aircraft systems depending on the flight distance required to achieve sufficient coverage of the two sites. The flight system of each aircraft (APM 2.6 or Pixhawk hardware running APM plane software) was programmed on the ground control station (APM mission planner) to follow a specific route based on GPS waypoints to form a full-coverage survey grid. This allowed for the aircraft to operate independently throughout the survey, except for take-off and landing. A catapult launcher was used to deploy aircraft, while landing was performed by manual flying to a suitable flat location. The UAV was equipped with a camera (Canon EOS M 22 mm, 18 megapixels) that took consecutive images every two, three, or four seconds, depending on the survey region and area covered. The camera was configured with auto-exposure and shutter speed priority to avoid motion blur. Images also recorded flight information such as GPS coordinates and altitude. The onboard flight log provided geo-referenced navigation data.

Surveys in Kaldfjord during December 2014 were performed with the Cryowing Micro RPAS (Remotely Piloted Airborne System; Thuestad et al. 2015). This aircraft has a 2.1-m wingspan made from expanded polypropylene (EPP) foam material, operating with an onboard Linux computer of the type Odroid X23 that controls the payload instruments and stores data. This equipment is powered by lithium polymer batteries (four cell [14.8 V] 10 Ah), which supplies the engine, the avionics, and the payload. In this configuration, the overall mass is 4.2 kg, which is able to operate for up to one hour at 18–20 m/s. However, with low-temperature conditions during the Arctic winter, we conducted flights with an optimal duration of 20 min to ensure recovery of the equipment. The UAV flew at an average altitude of 120 m (variation dependent on wind conditions), with a limited range of 1.2 km in 2014. This range was defined by local regulations for maintaining visual contact with the equipment and to ensure that the flight time would provide enough survey area for launch and recovery. We designed all surveys to include 50% image overlap to achieve full area coverage (similar to an aerial census).

From November to the end of December 2015, surveys were performed with the Cryowing Scout RPAS. This type of aircraft is of similar size as the Cryowing Micro RPAS, but with improved flight characteristics. It is designed as a motor glider with sleek aerodynamics and twin engines in the wings. The fuselage and wings are manufactured from composite reinforced plastic. With a wingspan of 2.7 m, this UAV is capable of carrying a payload of up to 5 kg. The flight time can exceed 90 min under optimal conditions, using 14.4 V batteries of 18 Ah. The avionics and payload control systems are equal to those used for the Cryowing Micro. During the Arctic winter, we conducted 40-min flights to ensure recovery of the equipment. As in 2014, this UAV flew at an average altitude of 120 m. Here, we also applied the full-coverage design, with 50% overlap and 2-s framerate (i.e., images taken every 2 s) with the exception of one flight, where the overlap was reduced to a minimum (5%, framerate every 4 s) in order to maximize areal coverage. In Kaldfjord, we performed five flights, covering three sections of the fjord. In survey section A (2014, Fig. 2), we conducted one flight, which covered around 1.2 km^2. In 2015, flights in sections B and C covered approximately 2.3 and 1.7 km^2, respectively (Fig. 2). Three flights took place in section B.

We performed six additional aerial surveys in Rystraumen using the Cryowing Scout RPAS (see aircraft characteristics above). The UAV flew at an average altitude of 150 m (variation dependent on wind conditions), with a limited range of 1.2 km. Similar to the flights in Kaldfjord, this range was within the local regulations. We designed a transect pattern to achieve full coverage with 72% image overlap (framerate of 3 s), taking into consideration the short surface time and small body size of porpoises. This allowed us to perform flights with a distance between transects of about 160 m. Communication at the start and end of surveys was performed as in Kaldfjord. We performed flights that covered the same survey region, resulting in a total of 2.4 km^2 of area surveyed in each flight (Fig. 3). All flights were conducted with a sensor size of 5184 × 3456 pixels and focal length of 22 mm.

Data analyses

All images collected during flights were manually examined for animal presence. For each image, we visually determined the following environmental variables: sea state, glare, luminance,

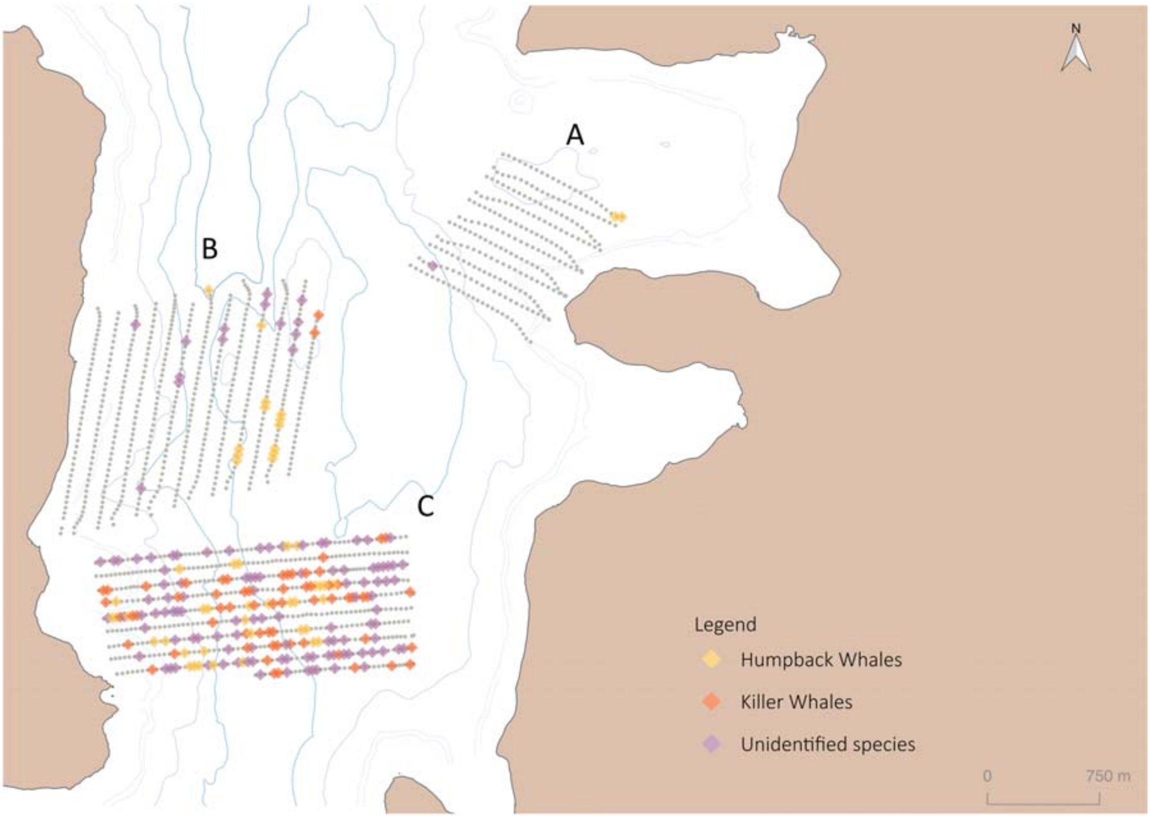

Fig. 2. Sightings during surveys in Kaldfjord in 2014 (transects on the top-right corner) and 2015 (transects on left corner) and survey sections (labeled A, B, or C). Map baseline source: Kartverket.

Fig. 3. Harbor porpoise detections (green) in Rystraumen (top). Map baseline source: Kartverket.

and tidal state. Sea state was defined as the amount of disturbance to the water's surface visible in the images as a categorical variable, where each level represented a certain degree of white caps or waves in the images (Figs. 4, 5). We also recorded light conditions (luminance) at the time of the survey by extracting information on brightness and contrast from the images. Animal sightings were categorized for presence/absence, certainty of detection, species, and number of animals. Our focus in the present study is toward understanding how the environmental and flight-related variables (altitude, pitch, roll, and image resolution) affect detection certainty (0, uncertain; 1, certain). Detectability certainty measured in UAV images was included here as a proxy for actual detectability.

We applied a generalized logistic model (GLM) to assess the detection certainty (Y) as a function of the environmental and flight-related variables (X):

$$Y_i = \mathrm{Bin}(n_i, \pi_i)$$

$$\mathrm{logit}(\pi_i) = \beta_0 + \beta_i x_{ij} + \varepsilon_i$$

where n is the number of observations for a given level (certain and uncertain), β are parameters to be estimated, and ε are the errors (assumed independent). We separated our analysis into two models, given the contrasting environmental conditions and species size, and thus accounting for analyst detectability related to both locations.

Environmental covariates

For summer surveys in Rystraumen, variables related to glare and tide were included in our analyses. These variables were excluded during winter surveys (Kaldfjord) given the low light conditions, survey frequency (in relation to tidal cycle), and the fact that tides did not affect water surface turbulence as dramatically in this relatively open fjord compared to the restricted tidal

Fig. 4. Example certain (top) and uncertain (bottom) images under different sea states, of whale detections during the winter in Kaldfjord. Sea state 1 is represented at the top image, and sea state 4 is represented at the bottom image.

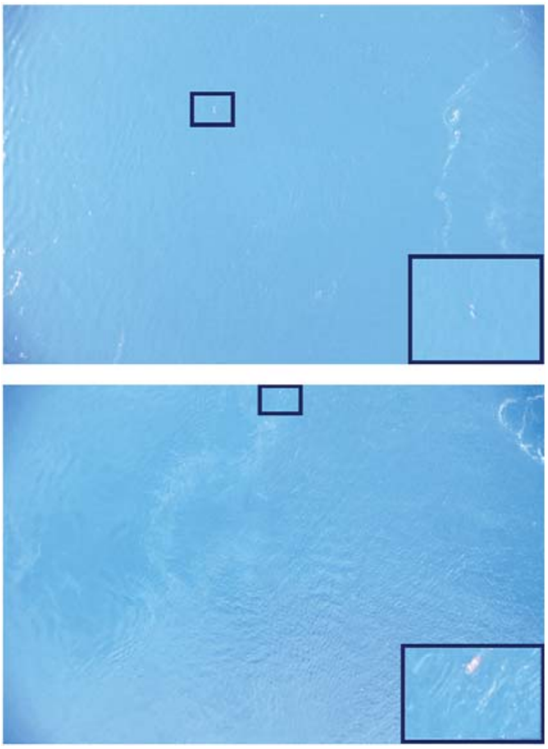

Fig. 5. Example certain (top) and uncertain (bottom) images under different sea states, of harbor porpoise detections during the winter in Rystraumen. Sea state 1 is represented at the top image, and sea state 2 is represented at the bottom image.

channel of Rystraumen. Luminance was calculated by measuring the average red, green, and blue (RGB) values for each image and then calculating a value of perceived luminance per image based. Luminance was given by Finley (2006) as:

$$\text{Luminance} = \sqrt{0.299\,R^2 + 0.587\,G^2 + 0.114\,B^2}$$

Aircraft and flight covariates

Flight information (GPS coordinates, altitude, pitch, and roll) was acquired for each image using either data stored directly in the EXIF header of the image file, or from the data record in the flight log that was nearest in time to when the image was acquired. These data were used to determine ground resolution, image footprint (area covered in each image), and ultimately to statistically examine how these variables influence detection certainty. We considered periods when the aircraft was on a straight-line transect and not intervals when the aircraft was performing sharp turns between transect lines. Similarly, we excluded transits between the launch/landing site and survey area.

We incorporated pixel ground resolution as our main predictor variable in models assessing the effects of aircraft-related parameters on detectability. Unmanned flights with re-occurring changes in pitch and roll may lead to changes in pixel ground resolution, particularly at the edges of the images (Yang et al. 2016). We took this into account by considering the actual resolution at the pixel coordinates within the image where an animal or group of animals were observed. We quantified pixel resolution for each sighting using a combination of pitch, roll, and altitude, in order to obtain a characterization of the effects of resolution in digital surveys. We used the following equation from Yang et al. (2016) to determine the difference in ground pixel resolution (pixel size) from when pitch and roll are zero:

$$X_{m,n} = \frac{H(d\cos\varphi Z + (m\,d\cos\varphi - f\sin\varphi) \times (d\sin\varphi\cos\omega))}{Z^2}$$

$$Y_{m,n} = \frac{H(d\cos\omega Z + (m\,d\sin\varphi\sin\omega + n\,d\cos\omega + f\cos\varphi\sin\omega) \times (d\sin\omega))}{Z^2}$$

$$Z = n\,d\sin\omega - m\,d\sin\varphi\cos\omega - f\cos\omega\cos\omega$$

Here, the aircraft rotation angles representing the effects of pitch and roll are defined by ω (roll) and φ (pitch). These, together with the pixel coordinates (*m* and *n*) of a sighting within the image, focal length *f* (mm), sensor pixel size *d* (mm), and altitude *H* (m), produce a change in resolution for the specific region of the image where animals were identified. The output of the equation ($X_{m,n}$ and $Y_{m,n}$) is in millimeters. Pitch and roll values provided by the aircraft flight logs come in both positive and negative values, representing the direction of movement. Negative values of pitch indicate a dive (nose points down), and negative values of roll indicate a lift of the right wing (aircraft turns left). The remaining images without animals included the theoretical resolution estimated for every pixel. Given animal orientation relative to the images is random and the pixels relatively square, *X* and *Y* axes' sizes are almost interchangeable in terms of their effect on certainty. Therefore, to avoid constraining pixel size to simply *X* or *Y* axis, we included the pixel area as representative of changes in pixel resolution.

Results

We conducted six flights in each field location. The average image size taken in Kaldfjord and Rystraumen was 9886 m^2 (121.6 × 81.3 m) and 15,423 m^2 (152.1 × 101.4 m), respectively. A total of 4660 transect images were taken during 12 flights totaling 4.2 h of flight (Table 1).

Animal detections

Of the 4398 UAV transect images, 288 images contained animals (hereafter referred to as sightings). Of these, 59 images included animals with certainty (Figs. 2, 6). In Kaldfjord, a total of 231 sightings were recorded: 50 HW, 63 KW, and 118 unidentified species. Field conditions varied considerably between winter flights (Kaldfjord). Images taken in 2015 were considerably darker in comparison with images taken in 2014. In Rystraumen, a total of 57 sightings were recorded: 5 certain and 52 uncertain (Fig. 7). In Kaldfjord, we recorded with certainty 37 HW, 50 KW, and six unidentified species. The latter numbers exclude consecutive detections of the same animals resulting from image overlap. The proportion of sightings identified as certain (as opposed to uncertain) was 23.4% (23.4, confidence interval [CI 95] = 18.2–29.5). These proportions differed significantly among the three survey sections (χ^2 = 12.175, P = 0.0023). Despite diminishing light conditions, the number of certain sightings

Table 1. Flight details for each survey: time and number of images captured in transect operations (excluding time spent in launch and recovery).

Site	Flight no.	Date	No. photos	Sea state range	Light conditions
Kaldfjord	1	12 December 2014	234	1–2	Partially cloudy
Kaldfjord	2	19 November 2015	208	1–3	Cloudy
Kaldfjord	3	10 December 2015	497	2–3	Partially cloudy
Kaldfjord	4	10 December 2015	528	2–4	Partially cloudy
Kaldfjord	5	11 December 2015	638	2–5	Partially cloudy
Rystraumen	1	26 August 2015	415	0–1	Clear
Rystraumen	2	26 August 2015	441	0–1	Clear
Rystraumen	3	02 September 2015	350	0–1	Clear
Rystraumen	4	02 September 2015	308	0–3	Clear
Rystraumen	5	02 September 2015	393	0–1	Clear
Rystraumen	6	02 September 2015	386	0–1	Clear

Note: Light conditions are defined by cloud coverage in the survey region, where partially cloudy <50% coverage; cloudy >50% coverage.

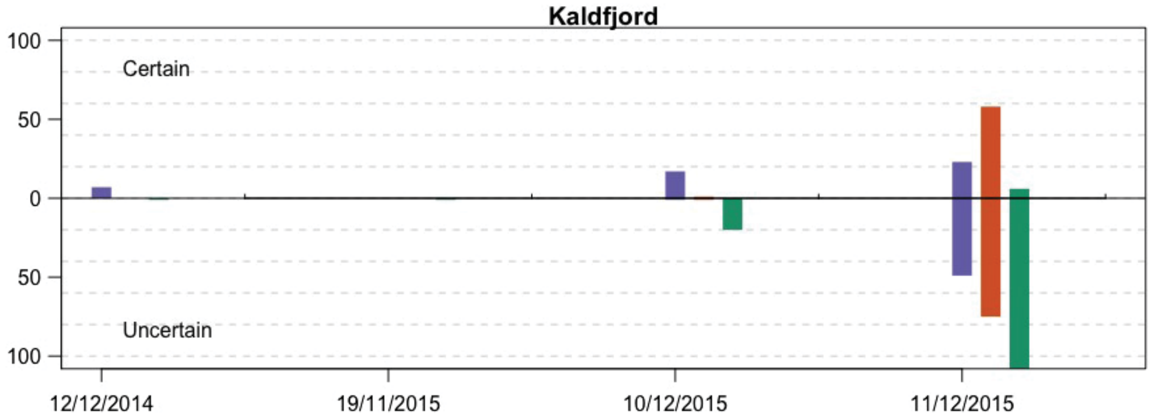

Fig. 6. Sightings of the three groups of animals recorded in unmanned aerial vehicle surveys during winter in Kaldfjord.

improved among surveys (difference in sightings in Fig. 4 between sections A, B, and C). There were more sightings with certainty in section C than in the other two sections (n_{HW} = 35, n_{KW} = 61).

In Rystraumen, the total proportion of animals identified as certain was 8.7% (8.8, CI_{95} = 3.3–2.0). This proportion, as in Kaldfjord, differed significantly among flights (χ^2 test = 13.89, P = 0.016). We obtained data for all four tide levels. In relation to tide, 88% of certain sightings were obtained at mid-tide (88.0, CI_{95} = 78.7–93.8). Most sightings (certain and uncertain) were recorded during flight 2 (n = 14), which was conducted during mid-low tidal period.

Effects of environmental conditions and aircraft/camera configuration

For Kaldfjord, we tested for a correlation between sea state and luminance, since the presence of waves may affect the measurement of brightness. We obtained a significant correlation between these two variables (r = −0.25, P < 2.2e–16; CI_{95} = −0.28 to −0.21). We therefore tested for the interaction effect between sea state and luminance. Given that this correlation is likely due to the autoexposure configuration of the camera, we compared the interaction model with an additive model for the same predictor variables (Table 2). The final model chosen was the additive model due to its goodness-of-fit

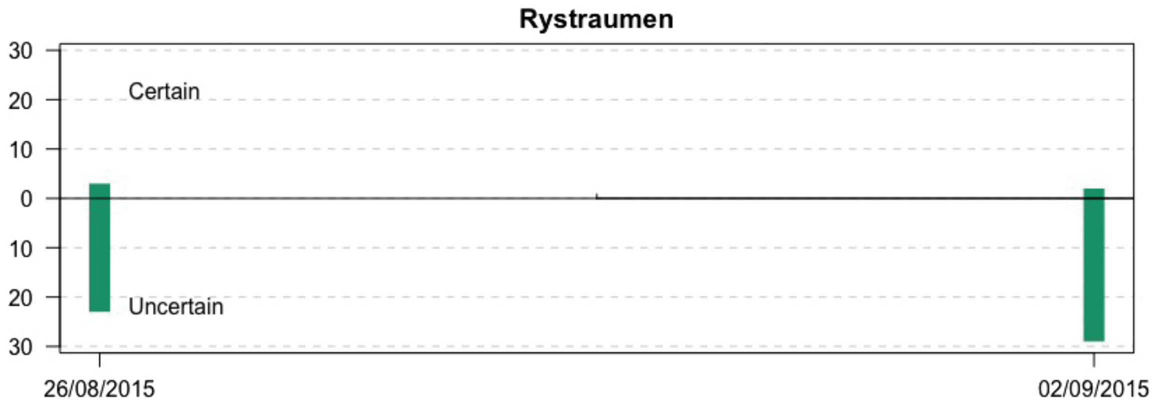

Fig. 7. Sightings of harbor porpoises (green) in Rystraumen.

Table 2. Results of model evaluation and comparison for Kaldfjord.

Model	Covariates	Coefficient	SE	*P*-value	Akaike's information criterion
Model 1	Sea state	−1.487	0.544	0.006†	989.97
Model 1	Luminance	0.012	0.133	0.930	
Model 1	Interaction (Sea state × Luminance)	0.008	0.022	0.713	
Model 1	Pixel area	0.015	0.036	0.681	
Model 2	Sea state	−1.268	0.104	<2e−16†	988.13
Model 2	Luminance	0.066	0.025	0.009†	
Model 2	Pixel area	0.081	0.222	0.716	

Note: The last three rows indicate the model selected.
† Significance at 95% confidence level.

values and resulting Akaike's information criterion (Model 2). Both sea state and luminance were found to be statistically significant (Table 2 and Fig. 8). Survey section was not included in the model given that the different sections were surveyed under different sea states. Therefore, the effect on certainty could not be separated from the effect on sea state. For the effects of pitch, roll, and altitude on the pixel size at the image location of sightings, the model tested showed no indication of an effect of pixel area on analyst detection certainty (Table 2 and Fig. 8).

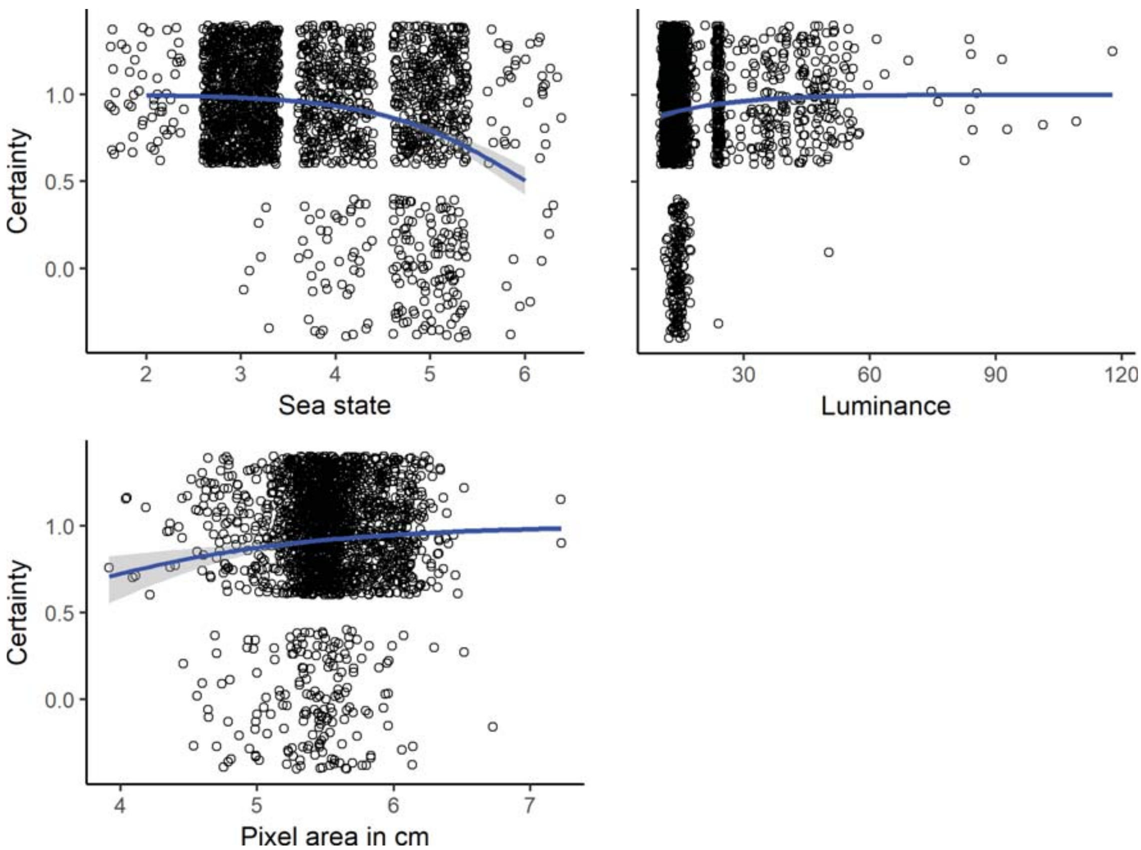

Fig. 8. Plots of the fitted jittered animal detections (1, certain; 0, uncertain) and regression lines (blue) resulting from the logistic model, as a function of sea state (left panel), luminance (right panel), and pixel size (bottom panel) in Kaldfjord in 2014 and 2015. Shaded areas indicate 95% confidence intervals.

Table 3. Results of model evaluation and comparison for Rystraumen.

Model	Covariates	Coefficient	SE	*P*-value	Akaike's information criterion
Model 1	Sea state	−0.531	0.350	0.129	496.7
	Glare	0.262	0.158	0.172	
	Pixel area	−0.526	0.387	0.174	

In Rystraumen, the data showed similar patterns as in Kaldfjord. Our measure of sea state was related to water turbulence, which in the case of Rystraumen was collinear with tide. Therefore, tide was removed from further data analyses. The variability caused by tide in our model could indeed influence the behavior of the animals (e.g., increase in prey availability during particular tidal states may influence the feeding behavior and therefore surfacing rates, presence in the channel, etc.), though this was outside of the scope of this study. We replaced luminance by glare due to the number of images that partially covered land and, as a result, influenced luminance values. From the covariates tested, surprisingly, the model showed that sea state and glare have no statistically significant effect on detection certainty (Table 3 and Fig. 9). Similar to Kaldfjord, increased pixel size/area also had no significant effect on certainty (Table 3, Fig. 9).

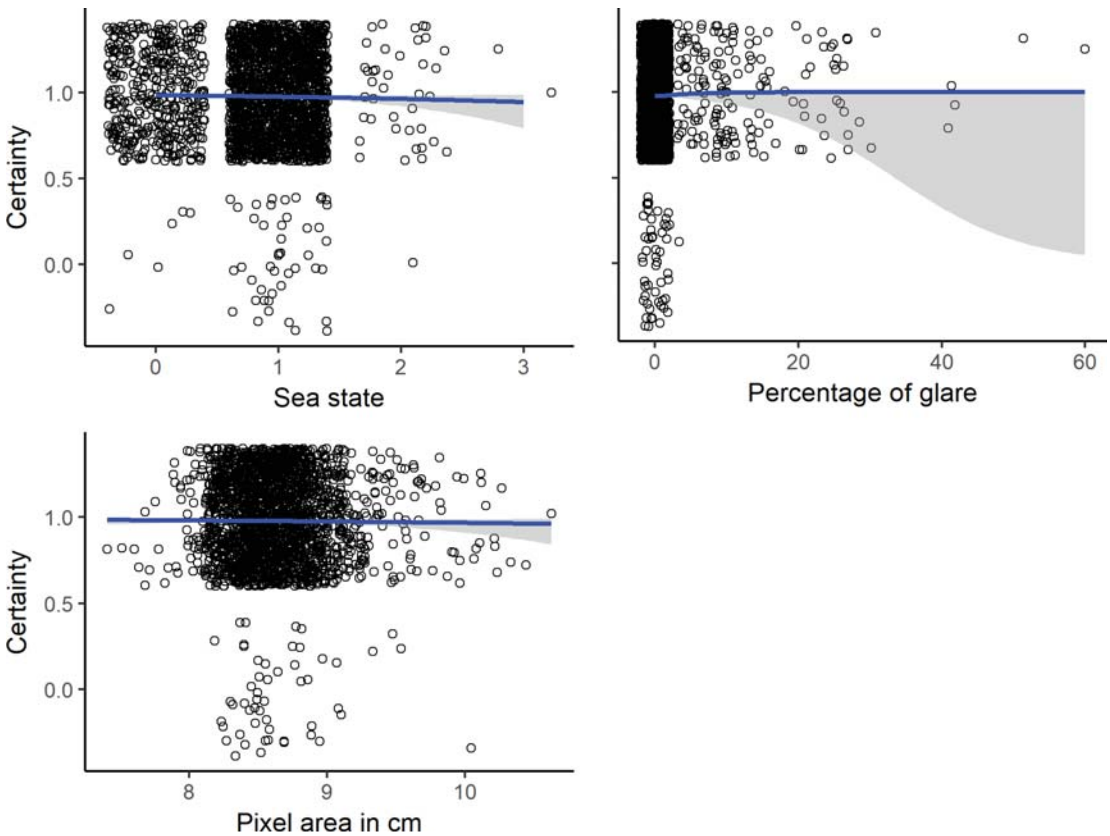

Fig. 9. Jittered detection values (1, certain; 0, uncertain) and resulting fitted regression lines (blue) given by the logistic model as a function of sea state (left panel), glare (right panel), and pixel size (bottom panel) of harbor porpoises in Rystraumen. Shaded areas indicate 95% confidence intervals.

Discussion

Unmanned aerial vehicles are becoming more widely recognized for their potential as effective and low-risk data collection platforms, with results that are easily replicated and have minimal impact on the surrounding environment (Smith et al. 2016). The present study contributes to the understanding of effects of selected environmental and technical factors on data acquisition for the detectability of KW, HW, and harbor porpoises by UAVs at high latitudes. Our results show that both humpback and KW in Kaldfjord can be detected during the polar night using UAV technology. Despite their small size, harbor porpoise sightings were recorded in Rystraumen during the summer trials. The numbers of sightings recorded indicate that there was a strong variability between survey flights, possibly due to the biological nature of the animals. Still, our work highlights the potential of using UAVs for also monitoring smaller mammals.

Factors affecting detection certainty

Certainty of detection is the primary metric for evaluating the utility of any animal survey method, and this can be affected by both environmental and aircraft conditions. In the present study, we address the role of sea state, luminance, and pixel size on detection certainty during winter UAV surveys of cetaceans. Sea state is known to affect observers' (or image analyst) ability to detect marine mammals in aerial surveys (Barlow 1988, Pollock et al. 2006). Here, our measure of sea state was represented as different category levels for water turbulence present in each image. The traditional Beaufort scale was not used in this study to describe sea state, since it is generally used to assess sea state from close to the sea surface, for example, from the bridge of a ship. Therefore, we did not consider the Beaufort scale appropriate when assessing sea state from images taken directly overhead. The certainty of detecting harbor porpoises was not affected by our measure of sea state possibly due to the low variance in the sea state levels encountered. Given that harbor porpoise surveys were conducted in the summer, the better visibility and brightness of the images could also influence the accuracy of the measurements in relation to the estimates of sea state for Kaldfjord. In contrast, the certainty of detecting humpback and KW in Kaldfjord was negatively related to sea state (Table 2), implying that an increase in sea state will decrease the certainty of detecting humpback and KW. Unlike Hodgson et al. (2013, 2017), we found that the reliability of UAV surveys can be limited by sea state conditions in the same way as traditional aerial surveys. Even though the work developed by Hodgson et al. (2013, 2017) was conducted in a different survey location, and considering whale numbers rather than analyst certainty of presence, observer/analyst understanding of sea state should be similar. Therefore, it is likely that their UAV and sensor system at high altitudes did not recognize wave turbulence at the same level of detail as in our study. Moreover, as can be seen in Fig. 4, the images can be quite dark and this could affect analyst categorization of the different sea states, particularly small variations in water turbulence as in sea states 1–3.

Glare is generally known to have an effect on detectability by limiting effective coverage of survey areas and produce false positives (Paiva et al. 2015, Kemper et al. 2016). Though glare was not included in Kaldfjord, image brightness (luminance) was found to have a positive effect on certainty of sightings in Kaldfjord (Table 2). This was anticipated for Kaldfjord, given that in the polar night it is extremely difficult to identify objects at or near the sea surface. We included glare on our second site, since the area covered included portions of land, which could affect our averaged RGB values and the resulting luminance. The results from Rystraumen are based on a relatively small number of certain sightings, and we believe that this was the main reason for the lack of statistical significance in our results (Table 3). Additionally, in this fjord floating debris can be linked to strong tidal currents could have had role in the accuracy of our measurements. Overall, our results show that UAVs in the Arctic, as any other marine survey methods, are dependent on adequate environmental conditions.

Although there was a large range in pixel size (Figs. 8, 9), results from both locations show that pixel size did not have a significant effect on observer certainty (Tables 2 and 3). This may be explained by the fact that most of the observations fell within a relatively narrow range, while the overall range was inflated by the presence of a relatively small number of outliers. Nevertheless, variations in pixel size should be taken into account when planning and analyzing future surveys in different conditions/altitudes, since pixel size is quantifiable within an image and can provide detailed information that cannot be obtained using other parameters in isolation. Though the use of altitude, pitch, and roll individually, has a value in itself, it does not quantitatively and accurately measure the output that will be given for a sighting within a UAV image. Thus, for estimates of resolution required for animal detection and species identification, we suggest that special care is taken for pixel size resulting from changes in aircraft stability that can be derived from environmental conditions or technological performance.

Improving UAV surveys

This study is among the first efforts in quantifying detectability and certainty in animal detections using UAV technology. Sea state and light conditions can affect certainty of marine mammal observations and should therefore remain under focus for environmental measurements of future surveys. In digital surveys, it is also important to not disregard aircraft movement throughout the survey and its effects on image resolution (pixel size). To minimize the effect of aircraft movements and ensure vertical camera view, we recognize that a gimbaled camera mount is an improvement to hard-mounting the camera to the UAV fuselage. Image overlap is a valuable tool in digital surveys as it can help to correctly estimate group size or assist in species identification (Hodgson et al. 2013, 2017). The high proportion of overlap in the images of Rystraumen and in flights in Kaldfjord during 2015 proved to be valuable in our measurements, though it was mainly used for confirmation of a sighting in consecutive images. However, high overlap demands a high number of images to be collected, which will affect the amount of data storage space needed onboard the aircraft.

For animal abundance and distribution estimates, spatial and temporal autocorrelation can be an issue, which will affect model validity. Spatial autocorrelation can lead to low variance between close observations and higher standard errors of mean estimates, which can then affect the levels of significance detected in parametric statistical models (Ferguson and Bester 2002), and therefore, efforts should be made to include these effects into statistical analyses of such data. Few studies have shown how to integrate spatial autocorrelation in digital surveys. (Salberg et al. 2009, Paiva et al. 2015, Conn et al. 2016). However, for the purpose of this study where the focus was the certainty of detections, we did not consider autocorrelation to have any influence in our analysis.

Detectability certainty was included as a proxy for actual detectability. Though this method is not the standard for image analysis and detectability measurements, and image analyst certainty may not be a perfect proxy for detectability, there are advantages in incorporating such measurements. For instance, higher detectability certainty is an improvement to low certainty and is probably in practice correlated to detectability. These two concepts should not be confused with one another, but seen as complementary. Even though observer detectability may remain the same, under different environmental conditions certainty may be reduced and therefore should be included as a metric in future surveys.

Unmanned aerial vehicle focus on marine studies has been increasing as the technology is quickly available at a fraction of the costs of manned aircraft. However, the payloads and camera systems that can accompany this technology play a key role in detectability improvements. As aircraft continue to develop, these issues must be addressed. Similarly, to digital surveys using manned aircraft, the use of remote-sensing equipment mounted on UAVs can increase the precision and accuracy of population size estimates for various species. For example, thermal cameras can detect animals based on their body heat and have the advantage of identifying individuals that are not easily visible to the naked eye (Christie et al. 2016). It is nevertheless relevant to incorporate animal surface behavior to develop abundance or density methods using UAVs in future studies. Hodgson et al. (2017) estimated detection probability based on whale focal follows from an UAV in combination with UAV transect surveys. This should be a focus in future attempts of estimating animal abundance based on animal availability, together with the uncertainties associated with environmental and aircraft covariates. We excluded animal behavior from our analyses since our focus was to attempt to understand the detection and measurement processes and how these could affect the quality of data acquired by UAV images. However, this study can be used as a provider of correction factors in future surveys that do incorporate animal behavior under similar conditions.

It is worth noting that we did not find any instances where the animals appeared to respond to the presence of the aircraft. We placed land-based observers during the UAV deployments and responsive behavior due to the noise or visual presence of the aircraft was not observed. Though the environmental noise present during the winter in the fjords can be quite elevated due to the large amount of fishing and recreational vessels, we believe that at the altitudes used in this study, the potential effects to the animals are minimal.

Special considerations for UAV surveys at high latitudes

Unmanned aerial vehicle in Arctic and sub-Arctic conditions is a valuable tool to study animal presence, density or abundance, and behavior. Our study shows that UAVs in the Arctic, as with the large majority of marine survey methods, are dependent on adequate environmental conditions (Forney et al. 1991, deMaster et al. 2001, Paiva et al. 2015, Kemper et al. 2016). Particularly during the polar night, light is a limiting factor that can significantly affect animal detection and certainty of sightings. Our findings are therefore not surprising, and light availability appears to have a role in unmanned surveys, which corresponds with the results reported by previous work (Parente and Elisabeth de Araujo 2011).

The effect of environmental attributes on an animal's preference will largely depend on its current biological requirements. Given the example of the winter migration of whales into the fjords of northern Norway, it would be relevant to include the effects of prey distribution, accessibility, inter-species, and human interactions. Despite lacking some behavioral conditionings to estimate animal abundance and distribution, the method presented here is an important step to further develop sound and realistic field protocols for aerial digital surveys.

Even though sea state conditions appear to have improved in the summer season, the number sightings in Rystraumen was not sufficient to provide an indication of its effects on the certainty of detections. However, the balance between the amount of observations recorded in the two survey regions, and the contrasting environmental conditions, that is, large numbers of humpback and KW in Kaldfjord surveyed in winter polar night conditions vs. few numbers of harbor porpoises in summer conditions, may have had an impact on the final results obtained in this study.

To conclude, this study shows that the UAV technology can be used to monitor sea mammals of a variety of sizes. Particularly at high latitudes, whether during summer or winter, environmental conditions can be limiting and should be incorporated into statistical models as well as UAV aircraft features that could change pixel resolution, and therefore affect detection probability.

Acknowledgments

The authors thank The Northern Research Institute's team (NORUT) for providing us with UAV equipment and data sharing. We would also like to thank Akvaplan-niva, which is a scientific partner within the Research Centre for Arctic Exploration (ARCEx). This research was partially funded by Akvaplan-niva and by ARCEx, the Research Council of Norway (Project #228107) with 10 academic and 8 industry partners. We thank the ARCTOS network for providing an environment for discussion with Arctic scientists.

Literature Cited

Barlow, J. 1988. Harbor porpoise, *Phocoena phocoena*, abundance estimation for California, Oregon and Washington: I. Ship surveys. Fishery Bulletin 86: 417–432.

Broms, F. 2017. The North Norwegian Humpback Whale catalogue (NNHWC), 2010–2017. http://hvalid.no/

Broms, F., F. Wenzel, P. L. Suarez, P. Stevick, M. Biuw, B. Jann, L. Bouveret, A. Rikardsen, C. Ryan, and S. Berrow. 2015. Recent research on the migratory destinations of humpback whales (*Megaptera novaeangliae*) from a mid-winter feeding stop-over area in Northern Norway. *In* F. W. Wenzel, editor. Recent research on the migratory destinations of humpback whales (*Megaptera novaeangliae*) from a mid-winter feeding stop-over area in Northern Norway. European Cetacean Society Special Publication Series, St. Julian's, Malta.

Chabot, D., and D. M. Bird. 2015. Wildlife research and management methods in the 21st century: Where do unmanned aircraft fit in? Journal of Unmanned Vehicle Systems 3:137–155.

Christie, K. S., S. L. Gilbert, C. L. Brown, M. Hatfield, and L. Hanson. 2016. Unmanned aircraft systems in wildlife research: current and future applications of a transformative technology. Frontiers in Ecology and the Environment 14:241–251.

Conn, P. B., E. E. Moreland, E. V. Regehr, E. L. Richmond, M. F. Cameron, and P. L. Boveng. 2016. Using simulation to evaluate wildlife survey designs: polar bears and seals in the Chukchi Sea. Royal Society Open Science 3:150561.

Ferguson, J. W. H., and M. N. Bester. 2002. The treatment of spatial autocorrelation in biological surveys: the case of the line transect surveys. Antarctic Science 14:115–122.

Finley, D. R. 2006. HSP color model—alternative to hsv (HSB) and HSL. http://alienryderflex.com/hsp.html

Forney, K. A., D. A. Hanan, and J. Barlow. 1991. Detecting trends in harbor porpoise abundance from aerial surveys using analysis of covariance. Fishery Bulletin 89:367–377.

Goodwin, L. 2008. Diurnal and tidal variations in habitat use of the harbour porpoise (*Phocoena phocoena*) in Southwest Britain. Aquatic Mammals 34:44–53.

Hiby, L. 1999. The objective identification of duplicate sightings in aerial survey for porpoise. Pages 179–190 *in* G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson, editors. Marine mammal survey and assessment methods. A. A. Balkema, Rotterdam, The Netherlands.

Hodgson, A., N. Kelly, and D. Peel. 2013. Unmanned aerial vehicles (UAVs) for surveying marine fauna: a dugong case study. PLoS ONE 8:e79556.

Hodgson, A., D. Peel, and N. Kelly. 2017. Unmanned aerial vehicles for surveying marine fauna: assessing detection probability. Ecological Applications 27:1253–1267.

Johnston, D. W., A. J. Westgate, and A. J. Read. 2005. Effects of fine-scale oceanographic features on the distribution and movements of harbour porpoises *Phocoena phocoena* in the Bay of Fundy. Marine Ecology Progress Series 295:279–293.

Kemper, G., A. Weidauer, and T. Coppack. 2016. Monitoring seabirds and marine mammals by georeferenced aerial photography. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B8, 2016 XXIII ISPRS Congress. Prague, Czech Republic.

Koski, W. R., T. Allen, D. Ireland, G. Buck, P. R. Smith, A. M. Macrander, M. A. Halick, C. Rushing, D. J. Sliwa, and T. L. McDonald. 2009. Evaluation of an unmanned airborne system for monitoring marine mammals. Aquatic Mammals 35:347–357.

deMaster, l. F., D. P. Lowry, K. J. Frost, and R. A. Bengston. 2001. The effect of sea state on estimates of abundance for beluga whales (*Delphinapterus leucas*) in Norton Sound, Alaska. Fishery Bulletin 99:197–201.

Paiva, E. G., C. Salgado-Kent, M. M. Gagnon, I. Parnum, and R. McCauley. 2015. An assessment of the effectiveness of high definition cameras as remote monitoring tools for dolphin ecology studies. PLoS ONE 10:e0126165.

Parente, C. L., and M. Elisabeth de Araujo. 2011. Effectiveness of monitoring marine mammals during marine seismic surveys off Northeast Brazil. Journal of Integrated Coastal Zone Management 11:409–419.

Pierpoint, C. 2008. Harbour porpoise (*Phocoena phocoena*) foraging strategy at a high energy, nearshore site in south-west Wales, UK. Journal of the Marine Biological Association of the United Kingdom 88:1167–1173.

Pollock, K., H. Marsh, I. R. Lawler, and M. W. Aldredge. 2006. Estimating animal abundance in heterogeneous environments: an application to aerial surveys for dugongs. Journal of Wildlife Management 70:255–262.

Ryan, C., M. W. Whooley, S. D. Berrow, C. Barnes, N. Massett, W. J. Strietman, F. Broms, P. T. Stevick, T. W. Fernald, and C. Schmidt. 2015. A longitudinal study of humpback whales in Irish waters. Journal of the Marine Biological Association of the United Kingdom 96:877–883.

Salberg, A.-B., T. A. Øigård, G. B. Stenson, T. Haug, and K. T. Nilssen. 2009. Estimation of seal pup production from aerial surveys using generalized additive models. Canadian Journal of Fisheries and Aquatic Science 66:847–858.

Smith, C. E., S. T. Sykora-Bodie, B. Bloodworth, S. M. Pack, T. R Spradlin, and N. R. LeBoeuf. 2016. Assessment of known impacts of unmanned aerial systems (UAS) on marine mammals: data gaps and recommendations for researchers in the United States. Journal of Unmanned Vehicle Systems 4:31–44.

Thuestad, A. E., H. Tømmervik, S. A. Solbø, S. Barlindhaug, A. C. Flyen, E. R. Myrvoll, and B. Johansen. 2015. Monitoring cultural heritage environments in Svalbard: Smeerenburg, a whaling station on Amsterdam Island. EARSeL eProceedings 14:37–50.

Watts, P., and D. Gaskin. 1985. Habitat index analysis of the harbor porpoise (*Phocoena phocoena*) in the southern coastal Bay of Fundy, Canada. Journal of Mammalogy 66:733–744.

Yang, Y., Z. Lin, and F. Liu. 2016. Stable imaging and accuracy issues of low-altitude unmanned aerial vehicle photogrammetry systems. Remote Sensing 8:316.