Polar Biology (2026) 49:43
https://doi.org/10.1007/s00300-026-03484-z

RESEARCH

A comparison of uncrewed aerial vehicle and on-ground surveys of penguin populations on the Antarctic Peninsula

Mairi Hilton[1,2] · Grant Humphries[2] · Camilla Nichol[1] · Clare M. Flynn[3,4] · Ron Naveen[2] · Steve Forrest[2] · Gemma V. Clucas[5] · Philip N. Trathan[6]

Received: 7 March 2025 / Revised: 18 February 2026 / Accepted: 24 March 2026

Abstract

Monitoring changes in wildlife populations over time requires regular assessments using robust methodologies. For many species of seabirds, including Antarctic penguins, population assessments have traditionally involved on-ground methods using standardised approaches. Recently, uncrewed aerial vehicles (UAVs) have emerged as a complementary tool to address challenges with traditional methods. To understand any potential impacts of this methodological shift on long-term datasets, we compared on-ground counts with two independent counts from UAV-derived imagery from 12 sites across the Antarctic Peninsula. We also investigated changes to breeding phenology at one site on the Peninsula, to help to contextualise other factors that influence population sizes across different temporal scales, as UAV surveys are usually conducted opportunistically and so may not capture peak populations. Whilst on-ground and UAV counts generally showed close comparability, only half of the counts fell within the generally accepted 5% repeatability threshold. UAV counts tended to be higher and more consistent with each other than on-ground counts, suggesting the UAV counts were more accurate, and may have captured parts of colonies missed during on-ground surveys. We observed significant variation in breeding phenology, and a negative correlation between survey timing and counts of breeding success. Through comparing our counts with historic data, our findings confirm gentoo penguin (Pygoscelis papua) population growth across the Peninsula, and showed mixed results for chinstrap (Pygoscelis antarctica) populations. We discuss potential sources of discrepancy between the methods and considerations for using UAVs in penguin surveys, particularly for informing management decisions.

Keywords Antarctica · Breeding phenology · Drone · Pygoscelis · RPAS · Seabird

Introduction

To understand and record changes to wildlife populations across time, regular assessments of population size are essential. Timely and accurate monitoring data facilitate informed decision-making, potentially leading to more effective management and improved conservation outcomes (Robinson et al. 2018). For many species of seabird, population assessments have traditionally involved on-ground methods where researchers hand-count the number of nests, often by entering the colonies (Hodgson et al. 2018). Such methods have become well established in Antarctica, including population monitoring of brushtail penguins (Genus Pygoscelis; Cobley and Shears 1999; Trathan et al. 2008; Dunn et al. 2019). Standardised approaches for conducting on-ground surveys have been developed (Agnew 1997; CCAMLR 2014) and applied across multiple sites, enabling population trends to be monitored and compared

✉ Mairi Hilton
mairihilton@gmail.com

1 United Kingdom Antarctic Heritage Trust, High Cross, Madingley Road, Cambridge CB3 0ET, UK

2 Oceanites, Inc., 3292 Arcadia Place Northwest, Washington DC 20015, USA

3 Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11794, USA

4 Institute for Advanced Computational Science, Stony Brook University, Stony Brook, NY 11794, USA

5 Cornell Lab of Ornithology, Cornell University, Ithaca, NY 14850, USA

6 British Antarctic Survey, High Cross, Madingley Road, Cambridge CB30ET, UK

Published online: 22 April 2026

across space and time (Lynch et al. 2008; Dunn et al. 2019; Herman et al. 2020).

There are, however, several drawbacks with on-ground survey methodologies, particularly in locations such as Antarctica. Penguin colonies are often remote, or in areas of complex topography, presenting access and observation challenges, especially where colonies extend over large areas of terrain where parts of a colony may be difficult to access (Trathan 2004; Dickens et al. 2021; Dunn et al. 2021). Brushtail penguin colonies can range in size from tens of birds to hundreds of thousands (Lynch et al. 2008; Borowicz et al. 2018; Herman et al. 2020), so opportunities to conduct surveys at larger colonies can be both challenging and time consuming. Colonies that are remote from research stations are generally limited by logistical opportunities and temporal constraints (Pfeifer et al. 2019; Dickens et al. 2021). Researcher proximity to birds can also cause disturbance (Chabot and Bird 2012), which may cause birds to temporarily abandon their nests (Ellenberg et al. 2015), potentially with long-term influences on population size (Bricher et al. 2008).

To help address these issues, there has been a recent increase in the use of uncrewed aerial vehicles (UAVs) to survey seabird populations (Hodgson et al. 2018; Rush et al. 2018; Dickens et al. 2021), including penguins (Ratcliffe et al. 2015; Pfeifer et al. 2019, 2025; Dunn et al. 2021; Krause et al. 2021; Dickens et al. 2021; Pina and Vieira 2022). Surveys are generally conducted by flying the UAV over colonies and capturing vertical images, which are then combined in post-processing to produce a single, mosaic image of the complete colony. Researchers can then count the number of nests from this mosaic image, including through the use of automated methods (e.g. Borowicz et al. 2018; Qian et al. 2023). Images can also be stored for later reference, or recounted as software methods evolve.

By eliminating the need to enter colonies, UAVs can help reduce disturbance (Chabot et al. 2015; Krause et al. 2021). However, as penguins have aerial predators whilst on land, they can exhibit aggressive or escape behaviours if UAVs are flown too close, as the UAV may be perceived as a threat (Rümmler et al. 2018; Harris et al. 2019). Nevertheless, it has been demonstrated that the disturbance impacts on nesting penguins are reduced compared to ground survey methods (Krause et al. 2021), if UAV surveys observe recommended minimum flight heights (see e.g. Mustafa et al. 2018b; Harris et al. 2019).

Alongside reducing disturbance, there are further benefits to surveying using UAVs. They can make previously inaccessible sites accessible to researchers (Dunn et al. 2021; Dickens et al. 2021), and the top-down vertical view can also reveal birds that may have been obstructed from sight during on-ground surveys (Chabot and Bird 2012). UAVs also facilitate surveying of larger colonies due to the increased rate of surveying (Borowicz et al. 2018; Pfeifer et al. 2019), and the images captured provide a permanent record of the colony at that time, which can be used to assess environmental information related to colony distribution or habitat variables (Hodgson et al. 2018).

To ensure that new census counts from UAV surveys are comparable with historical long-term datasets, UAV methods must be rigorously evaluated. Some penguin colonies have been monitored using traditional on-ground methods for several decades (Santos et al. 2018; Dunn et al. 2019, 2021), so it is vital that these datasets are not compromised. Previous research has demonstrated that UAV and on-ground surveys conducted on the same day are within ± 5% (e.g. Dunn et al. 2021), the generally accepted threshold used to ensure count repeatability for brushtail penguin surveys (Croxall and Kirkwood 1979). Other studies comparing UAV and on-ground surveys for other seabird species have even found that UAV surveys can yield more accurate counts than traditional methods (Hodgson et al. 2016, 2018), though the accuracy is affected by several factors including the quality of the UAV imagery, terrain, weather (precipitation, cloud cover) and species composition (Chabot and Bird 2012; Pfeifer et al. 2025). Yet, to date, there has been no wide-ranging comparison of on-ground and UAV surveys for Antarctic penguins. As UAVs are increasingly used for estimating penguin population size, a broader understanding about methodological limitations is needed.

In addition to better understanding the influence of survey methodology, interpreting long-term population trends also requires an understanding of other factors that influence population sizes across different temporal scales. Traditional on-ground survey methods are generally conducted in a manner that ensures that peak nesting is captured (CCAMLR 2014), so that between-year surveys are comparable (Trathan 2004; Dunn et al. 2021), which is important as gentoos are known to show significant inter-annual fluctuations in phenology (Hinke et al. 2012; Black 2016). However, UAV surveys of penguin colonies are often opportunistic, given the types of ship platforms used. Consequently, it is important to understand how the timing of surveys is likely to influence population estimates, regardless of survey method.

Here, we used data from the 2023/24 Austral summer to compare counts of penguin colonies collected by UAV and traditional on-ground methods. Using data from a range of sites, we identified where differences in counts between methodologies arose, investigated possible causes of those differences, and compared the counts from our surveys with previous counts for each colony. We also investigated changes to breeding phenology at one site on the Antarctic Peninsula, Goudier Island, where on-ground surveys have been undertaken for over 20 years. We discuss the implications this may have when comparing counts between survey methodologies and assessing long-term population trends.

Methods

On-ground survey methods

We used two different approaches to estimate penguin population size using on-ground surveying methods. The first was the Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR) Ecosystem Monitoring Programme (CEMP) standardised method for assessing penguin breeding population size (CCAMLR 2014), which involves selecting one penguin sub-colony that is monitored every two days to determine the number of nests, number of nests with eggs, and then the number of nests with chicks. The observer stands at the edge of the sub-colony, or enters the sub-colony if necessary, and records each nest and how many eggs or chicks are present. Once 95% of nests in the sub-colony have eggs (then later chicks), a colony-wide survey is triggered. The CEMP method therefore provides information on breeding population size, breeding success, and breeding phenology.

Although this method is more informative, it requires nearly continuous researcher presence, which is typically only possible at research stations or at long-term camps. For the majority of our sites, we therefore used a modified CEMP procedure, which involved observers recording nest counts from near or within the colony, and repeating counts three times to ensure that counts were within ± 5%, with additional repeats if the counts were not within this margin. Observers worked in pairs, but counted separate parts of the colony to ensure the entire colony was counted within the time available. The relationship between counts observed and timing of peak egg-laying was therefore not documented, but we believe nonetheless provides a robust and comparable estimate of the peak nesting population in almost all cases.

Un-crewed aerial vehicle survey method

UAV surveys were conducted using DJI Mavic Enterprise 3 Thermal aircraft. All surveys were flown at a minimum height of 50 m. Previous research has flown UAVs at heights of 35 m for gentoo (Dunn et al. 2021) and 30 m for chinstrap (Krause et al. 2021) penguins without reports of disturbance behaviours; however, surveys in the 2023/24 Antarctic field season were all flown higher due to concerns about Highly Pathogenic Avian Influenza (HPAI). The Scientific Committee on Antarctic Research (SCAR) recommendations required a higher flight height to minimise the possibility of any potential additional stress for the birds (Dewar et al. 2023). All flights took place at wind speeds below 20 knots (10.3 m/s). Further, as precipitation (snow, rain or mist) can compromise image quality, most surveys were conducted in dry conditions.

Smaller colonies were surveyed manually, whilst pre-programmed routes were used for larger colonies. Both methods were conducted using the typical "lawnmower" flight pattern to cover the entire colony. For pre-programmed routes, we ensured 70% overlap between photos forwards and to each side, and used the Reference Elevation Model of Antarctica (REMA) digital elevation model (Howat et al. 2022) to ensure a consistent height above ground level (AGL) of 50 m, replicating these conditions as closely as possible for manual flights. Camera resolution was approximately 1.5 cm ground sample distance (GSD), altering with the altitude of the flight. Pre-programmed routes were checked against the terrain upon arrival at a site, and adjusted to ensure the entire colony was captured. All UAV surveys involved two personnel: one individual was responsible for flying the UAV, whilst the second maintained visual line of sight and acted as a spotter to alert the pilot to risks from terrain, wildlife or encroachment of uninvolved humans.

The research was conducted under a permit issued by the United States National Science Foundation Office of Polar Programmes.

Site selection

For the second on-ground counting method and all UAV surveys, teams of two observers were deployed across several expedition cruise ships. Expedition cruise ships, sometimes referred to as 'ships of opportunity', bring tourists to sites across the Antarctic Peninsula and the South Shetland Islands. Site selection was therefore determined by the schedule of the host ships. As expedition ships commonly visit the same sites, some sites were surveyed multiple times throughout the season, whereas less frequently visited sites were surveyed only once.

Research teams were generally ashore first at each site, surveying whilst the ships conducted their normal operations. The time available for surveying was therefore dependent on the length of the host ships' operations, but usually teams had between three and six hours available at each site. Time and weather permitting, teams attempted to conduct both on-ground surveys and UAV surveys during each visit. If weather conditions prevented flying the UAV (sustained windspeeds over 20 knots, heavy precipitation, or fog), then on-ground surveys were prioritised. If there was only a short period of time to survey a large colony, then UAV surveys were prioritised. Here, we include data from surveys where both a UAV survey and a ground count of the number of nests were conducted on the same day, providing

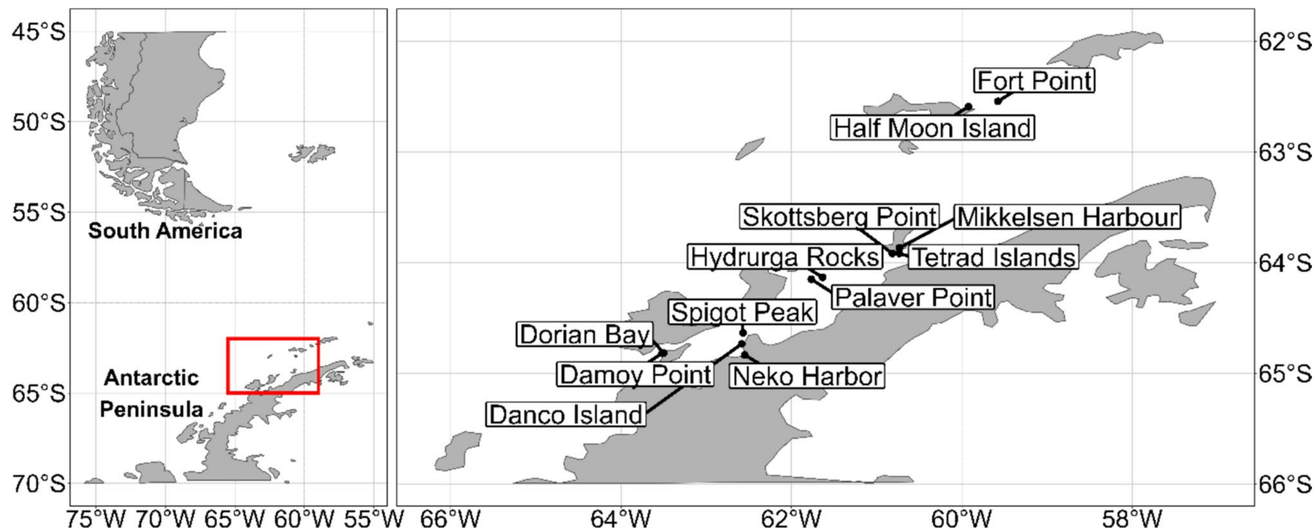

Fig. 1 Location of the 12 Pygoscelid penguin colonies included in the comparison of UAV and on-ground survey methodologies

Table 1 Pygoscelid penguin colonies included in the comparison of UAV and on-ground survey methodologies, the date the survey was conducted, and the species surveyed. See Fig. 1 for map of colony locations

Site name	Survey date	Penguin species surveyed
Damoy Point	2024–01-02	Gentoo (Pygoscelis papua)
Damoy Point	2024–01-22	Gentoo
Damoy Point	2024–01-24	Gentoo
Danco Island	2024–02-03	Gentoo
Dorian Beacon	2024–01-02	Gentoo
Dorian Beacon	2024–01-24	Gentoo
Fort Point	2023–12-16	Gentooa
Half Moon Island	2024–01-07	Chinstrap (Pygoscelis antarcticus)
Hydrurga Rocks	2023–12-17	Chinstrap
Mikkelsen Harbour	2024–01-06	Gentoo
Neko Harbour	2024–01-04	Gentoo
Palaver Pointb	2023–12-17	Chinstrap
Skottsberg Point	2024–01-06	Gentoo
Spigot Peak Pointb	2023–12-02	Gentoo and chinstrap
Tetrad Islands	2024–01-06	Chinstrap

[a]Chinstrap also surveyed but not included in study as surveyed using UAV only

[b]Also had Antarctic Shag (Leucocarbo bransfieldensis) present

16 surveys from 12 unique sites (Fig. 1). Surveys included counts for gentoo (Pygoscelis papua) and chinstrap (Pygoscelis antarcticus) penguins (Table 1).

Analysis

Image analysis

DJI Terra (DJI 2024) was used to stitch the individual vertical images collected during each UAV survey to create a single mosaic image for the whole colony on each survey date. Images were georeferenced to WSG1984 UTM 20S. QGIS (QGIS Development Team 2024) was then used to inspect images to determine whether penguins were either nesting or non-nesting. Each nesting penguin was marked with a point feature to facilitate counting of all marked penguins. Two observers, who had two and 25 years of experience of working with penguins in the field, independently marked nesting penguins for each survey to give two independent counts. Both observers had previous experience of ground and UAV surveys of penguins in Antarctica, and of assessing UAV imagery of penguin colonies.

To determine which penguins were nesting, several visual clues were used, following Dunn et al. (2021). Brushtail penguins build nests out of pebbles, so penguins surrounded by obvious pebbles or a grey-black outline were indicative of a nest. Rings of guano streaks around birds were also used to identify nests, as well as bird posture. Nesting birds usually lie flat with their wings by their sides, or may be hunched over eggs or young chicks. Standing birds were also marked as nesting if surrounded by indications of nests or guano streaks, as the bird may have been briefly standing or stretching at the time the image was taken. At the two mixed species colonies (Table 1), species were differentiated using phenotypic differences, such as the white band on the head of gentoo penguins, and more extensive white around the top of the chest and face on the chinstrap, as in

Pfeifer et al. (2025). Species were usually segregated at a colony, and so prior knowledge on the distribution of species at the site was useful. Examples of imagery from mixed- and single- colonies are provided in the Appendix to highlight these phenotypic differences, and how nesting penguins were marked in QGIS.

Methodological comparison

To examine whether there were differences in counts obtained by UAV and on-ground surveys, for each site on each survey date, we compared the total number of nests counted during the on-ground survey with each of the counts from the UAV imagery. We also compared the total number of nests in the two counts from the UAV imagery.

Phenological analysis

To understand how the timing of opportunistic surveys may impact counts, we used Goudier Island in the Palmer Archipelago of the Antarctic Peninsula (64°49′S, 63°29′W) as a case study to examine nesting phenology. Goudier Island is in Port Lockroy Harbour, and we herein use the term Port Lockroy to refer to the site. We selected Port Lockroy as a case study because it is home to approximately 1000 gentoo penguins, which have been monitored almost every Antarctic summer since 1996/97 using the CEMP methodology (Trathan et al. 2008; CCAMLR 2014; Dunn et al. 2019), and has also had UAV surveys in several recent seasons (Flynn et al. 2023, 2026). The publicly available, long-term dataset (Dunn et al. 2020) allowed us to examine how the date of peak nesting has changed across almost 30 years, and investigate any possible relationship between phenology and nesting population size. Port Lockroy has exhibited a long-term decline in gentoo penguin numbers, and previous research has investigated possible links between the decline and tourism (Trathan et al. 2008; Dunn et al. 2019). Anecdotal evidence suggested that surveys triggered later in the season more often yielded reduced counts compared to surveys triggered earlier in the season, so we aimed to investigate this link.

To investigate whether there have been any phenological changes at Port Lockroy, we calculated the number of days difference between the date of each annual survey and the date of the first survey in the 1996/97 season, which was considered to be a nominal baseline for estimating differences in phenology. We examined differences in the timing of the surveys of the number of nests (Survey One), number of hatched chicks (Survey Two), and the number of crèched chicks (Survey Three). The dates of these surveys in the 1996/97 season were 06/12/1996, 15/01/1997, and 14/02/1997, respectively. We developed linear regression models to examine the influence of the number of years since the first survey on the number of days' deviation compared to the date of the first survey (DaysDeviation ~ YearsSinceFirstSurvey), developing separate models for each of the three surveys. For Survey One (number of nests), we removed surveys from the 1997/98 and 2021/22 seasons. Notes for the 1997/98 survey stated that the survey was conducted late but did not provide a reason. The survey in the 2021/22 season was delayed due to late arrival of the team at the base due to the COVID-19 pandemic.

We then developed linear models to examine whether there was any influence of the date of each survey on the survey outcome, using the number of days' deviation compared to the date of the first survey as the predictor (SurveyResult ~ DaysDeviation), again developing separate models for each survey.

All statistical analyses were conducted in R (R Core Team 2024).

Population changes

To examine changes in population size at each of our 12 study sites (Table 1), we compared the number of nests recorded in our counts with a) the number of nests in the first recorded survey, and b) the number of nests in the most recent ground survey. Data for these comparisons was downloaded from the publicly available dataset Mapping Application for Penguin Populations and Projected Dynamics (MAPPPD; Oceanites 2023).

Results

Methodological comparison

When comparing the number of nests recorded by the on-ground count and two UAV counts, R2 values indicated close comparability (Fig. 2). However, only eight out of the 16 on-ground counts (50%) were within the standard 5% variation for count repeatability (Croxall and Kirkwood 1979) with at least one UAV survey, whereas UAV counts were generally in close agreement (Fig. 2), with 12 counts (75%) within 5% (Fig. 3; Appendix 1). The average variation between the on-ground count and UAV count 1 was 9.9%, whilst the average variation between the on-ground count and UAV count 2 was 11.7%. In comparison, the average variation between UAV counts 1 and 2 was just 1.2%. Counts from UAV surveys tended to give higher numbers of nests than on-ground surveys (12/16 surveys; 75%), especially when the difference was substantial (Fig. 3; Appendix 1).

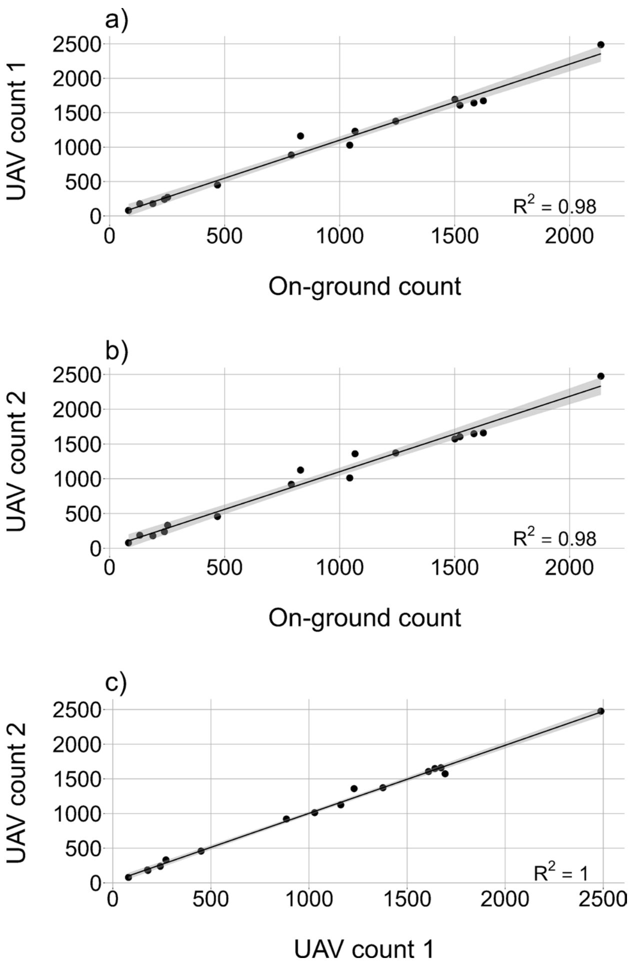

Fig. 2 Comparison of on-ground and UAV surveys of the number of nests in 16 surveys from 12 unique penguin colonies on the Antarctic Peninsula in the 2023/24 season. **A** compares the on-ground count to UAV count 1, **B** compares the on-ground count to UAV count 2, and **C** compares the two UAV counts. Solid line indicates the linear relationship between counts, grey bar around the line indicates 95% confidence interval

Phenological results

There was substantial variation in the date that population size and breeding success surveys were triggered at Port Lockroy, based on CEMP methods (Fig. 4). The 2022/23 season recorded the greatest delay for all three surveys, with 30-, 37- and 22-day delay compared to the 1996/97 survey dates for the number of nests, number of hatched chicks, and number of crèched chicks, respectively. The timing of all three surveys trended later over time (Fig. 4), but we only detected a statistically significant positive effect for Survey One (number of nests), with a moderate proportion of variance explained ($\beta = 0.61$ (CI: 0.16–1.06), $p = 0.01$, $R2 = 0.26$, Table 2).

When examining the influence of survey date on the survey outcome, all three surveys displayed a negative relationship between the number of days deviation and survey outcome (Fig. 4; Table 2). However, the influence was only statistically significant for the number of hatched chicks ($\beta = -13.43$ (CI: $-21.31 - 5.5$), $p = 0.002$, $R2 = 0.33$, Table 2) and number of creched chicks ($\beta = -10.59$ (CI: $-19.55 - -1.64$), $p = 0.02$, $R2 = 0.19$, Table 2). We did not detect a significant effect of survey timing on number of nests ($\beta = -4.01$ (CI: $-8.69 - 0.67$), $p = 0.09$, $R2 = 0.12$, Table 2).

Population changes

All twelve of our study colonies had increased in population size since their previous most recent survey, irrespective of the survey methodology considered (Appendix 1). The majority of sites had also increased or stayed stable since their first survey, except for Neko Harbour, which had declined when compared to the on-ground counts but not UAV counts, and Half Moon Island and Palaver Point, both of which had declined significantly according to all three counts (Appendix 1). As with other areas of the Antarctic Peninsula, we documented a substantial increase in some gentoo colonies, with several colonies found to have more than doubled their populations since their previous survey (Appendix 1; Lynch et al. 2012b). The largest increase was of gentoo penguins at Spigot Peak Point, where gentoos were first recorded in 2007, and increased from 3 nests in 2019 to 82 (79 according to the on-ground count; Appendix 1).

The gentoo population at Port Lockroy also increased since the previous survey, with 600 nests recorded. However, the population has declined since the first survey in 1996, when 726 nests were recorded (17.3% decline).

Discussion

Our study highlights variability between on-ground counts and UAV surveys at 12 sites across the Western Antarctic Peninsula. When comparing survey methodology, eight out of sixteen surveys were within the standard threshold of 5% for count repeatability (Croxall and Kirkwood 1979; Dunn et al. 2021), with differences of up to 42.7% observed (Fig. 3, Appendix 1).

Whilst we did not attempt to explore differences in observer variation, the UAV counts by different observers were largely consistent, although a quarter of counts still had a greater than 5% difference (Fig. 3, Appendix 1). This suggests that the two observers mainly classified nests similarly based on evidence from the imagery.

Although our sample size was too small to conduct statistical analyses to compare methodologies, generally, smaller colonies tended to have the most similar counts between on-ground and UAV surveys (Fig. 2, Appendix 1). This is

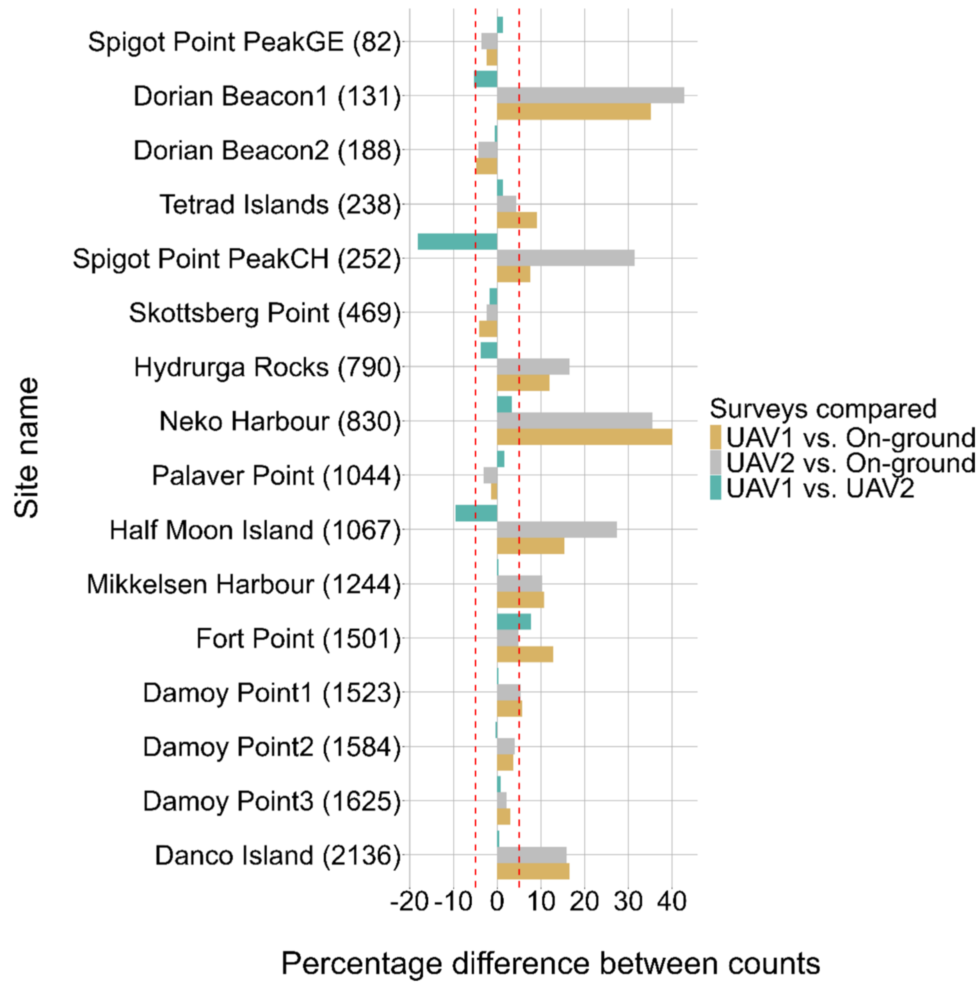

Fig. 3 Percentage differences between counts obtained by on-ground counts and two independent counts of UAV imagery of nests in 16 surveys from 12 unique penguin colonies. Sites are ordered by the result of the on-ground survey, which is indicated in parentheses. Damoy Point and Dorian Beacon were both surveyed multiple times throughout the season; the numbers beside the site names represent the temporal order of surveys with increasing numbers indicating the survey was later in the season. At Spigot Point Peak, CH indicates the result of the chinstrap survey, whilst GE is the result of the gentoo survey. Red vertical dashed lines indicate ± 5%, which is the generally accepted threshold for count repeatability

consistent with previous work comparing on-ground and UAV surveys of penguin colonies, where larger colonies recorded the greatest deviations between counts (Dunn et al. 2021). There were, however, exceptions to this trend, with a survey of the second smallest colony recording the second largest discrepancy in counts (Dorian Beacon; Fig. 3, Appendix 1). We believe this is likely due to part of the colony being missed during the ground count, as the time available to survey was cut short due to changing weather conditions. The large discrepancy between other ground counts and UAV counts at several colonies (Fig. 3, Appendix 1) suggests that UAV surveys captured parts of colonies that were missed during ground counts, as opposed to errors in counting the same parts of the colonies. By using pre-programmed transects based on Google Earth maps and prior knowledge of colonies, UAVs help ensure that the entire colony is surveyed. As in our study, the flight path can be assessed once researchers are on ground before the survey begins, and buffers expanded around colonies to account for any movement of sub-colonies between years, or survey heights increased where numerous flying seabirds exist (see Methods). If there is doubt about the completeness of the survey, a permanent record of the colony has been captured in the imagery, enabling researchers to go back to the image and assess it (Hodgson et al. 2018). Further, several observers can count from the imagery and counts can be repeated if unexpected numbers are returned, thus reducing the opportunity for observer error, highlighting further benefits of surveying by UAV.

Despite the ability of ground observers to detect more fine-scale evidence that would affect nest classification, previous work has demonstrated that UAV surveys are likely to lead to more accurate counts than on-ground counts. For example, during an experimental survey using a known number of replica greater crested terns (Thalasseus bergii), it was found that UAV surveys yielded significantly more accurate counts than ground counts (Hodgson et al. 2018). There are a multitude of factors that may influence ground counts on the day, such as the season, the time of day, the time available for surveying, distribution of the colony, and terrain type (Chabot and Bird 2012). Although there are factors that influence the quality of UAV imagery and subsequent interpretation, such as weather and light conditions, terrain and observer training, the close agreement between

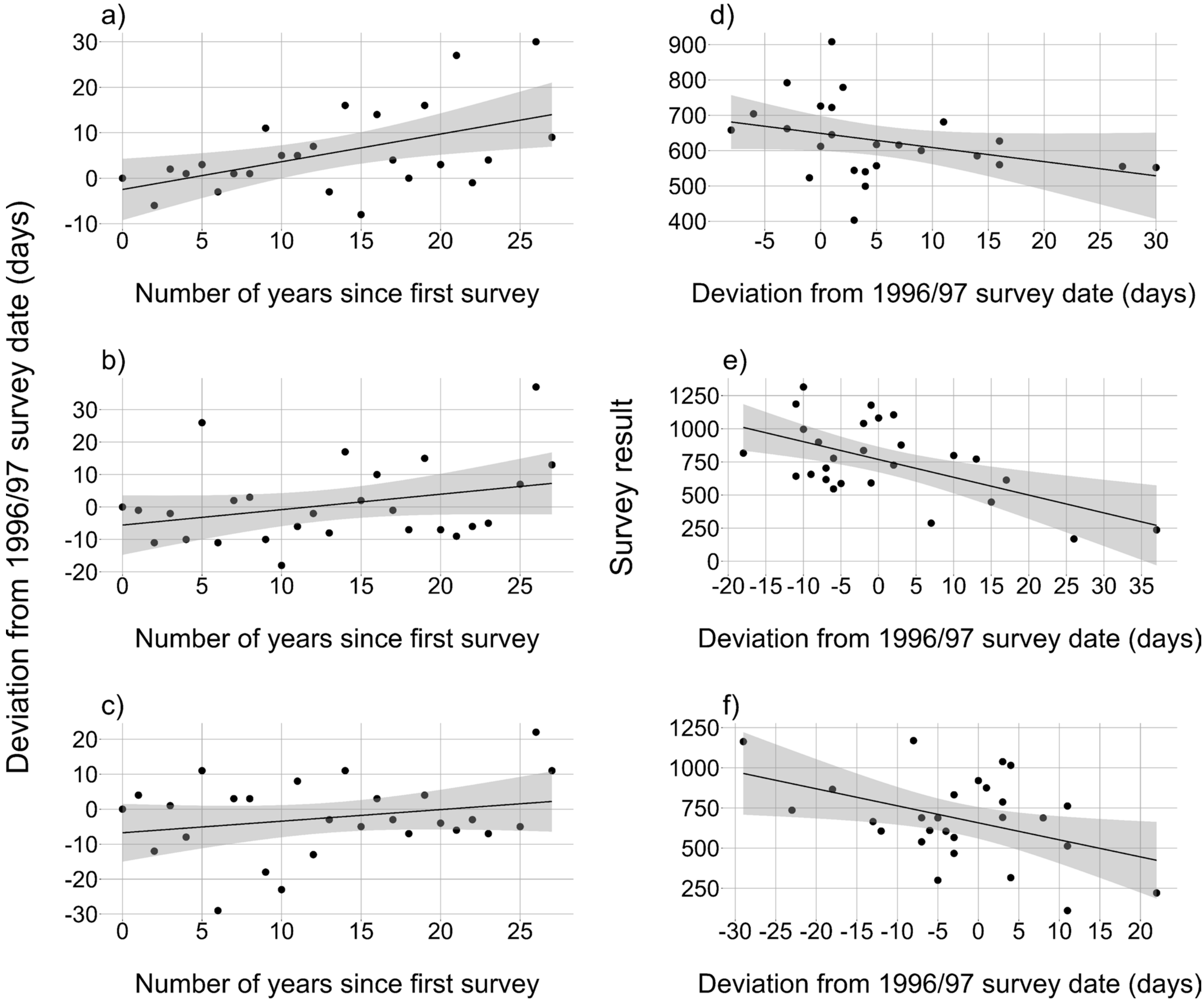

Fig. 4 Variation in the gentoo penguin population counts at Port Lockroy using the Conservation of Antarctic Marine Living Resources (CCAMLR) Ecosystem Monitoring Programme (CEMP) standardised methodology. **A**–**C**: Deviation in date of the CEMP survey since the first survey in the 1996/97 season; **A**: nests; **B**: hatched chicks; **C**: crèched chicks. **D**–**F**: Survey count against the number of days that the survey date deviated since the first survey in the 1996/97 season; **D**: nests; **E**: hatched chicks; **F**: crèched chicks. Solid line indicates the linear regression between variables, grey bar around the line indicates 95% confidence interval

Table 2 Results of analysis into phenology at the gentoo penguin colony at Port Lockroy

Analysis	Survey	*β* (95% CI)	SE	*p*	*R*2
Timing of survey (DaysDeviation ~ YearsSinceFirstSurvey)	Model1: Number of nests	0.61 (0.16 – 1.06)	0.22	0.01**	0.26
	Model2: Number of hatched chicks	0.48 (− 0.12 – 1.07)	0.29	0.11	0.1
	Model3: Number of creched chicks	0.33 (− 0.21 – 0.87)	0.26	0.22	0.06
Influence of survey timing on survey outcome (SurveyResult ~ DaysDeviation)	Model4: Number of nests	− 4.01 (8.69 – 0.67)	2.26	0.09	0.12
	Model5: Number of hatched chicks	− 13.43 (− 21.31 – 5.56)	3.82	0.002**	0.33
	Model6: Number of creched chicks	− 10.59 (− 19.55 – − 1.64)	4.35	0.02*	0.19

Models 1–3 assessed whether the timing of the surveys into the number of nests (Model 1), number of hatched chicks (Model 2) and the number of crèched chicks (Model 3) had changed over time since the first survey in the 1996/97 season. Models 4–6 assessed whether the date of the survey had an influence on the outcome of the surveys into the number of nests (Model 4), number of hatched chicks (Model 6) and the number of crèched chicks (Model 6)

UAV counts suggests that the imagery was of sufficient quality to enable accurate counting.

Although the aerial perspective afforded during UAV surveys is largely advantageous, it can make it more difficult to assess subtle variations in posture to determine whether a bird is nesting, which may result in counts that are overestimated (Pfeifer et al. 2025). For example, a recent study of the gentoo colony at Port Lockroy had to revise their published population estimate derived from a UAV survey after re-analysing the imagery, reducing the count by almost half (Flynn et al. 2023, 2026). The authors discuss how UAV imagery quality can impair an observer's ability to distinguish the breeding status of gentoo penguins and can lead to inflated nest counts (Flynn et al. 2026). This uncertainty can be reduced by multiple observers counting imagery, which is especially important in the case of anomalous count results as were originally reported (Flynn et al. 2023).

All colonies had increased in size since their previous survey irrespective of survey methodology. Ten of the colonies had also increased since their first survey according to UAV counts, but only eight had increased according to ground counts (Appendix 1). This result highlights that perceived increases in colony size may be a consequence of the change in survey methodology, rather than true increases. Whilst our results support the conclusion that UAV surveys generally provide more accurate results than ground counts, it is also possible that numbers reported in UAV surveys may be overestimated for the reasons described above.

Of the four sites where UAV counts in our study differed by more than 5%, two were mixed species colonies (Fort Point and Spigot Point Peak; Table 1), one of which (Spigot Peak Point) also had Antarctic Shags present (Table 1). These were the only two mixed species colonies included in our methodological comparison (Table 1). Our imagery was of sufficient quality to differentiate nesting from non-nesting penguins, but may not have been adequate for distinguishing between species, although the colouration of guano can help when distinguish shags from penguins. Recent work has provided further guidance on approaches to distinguish nesting Pygoscelid penguins from UAV imagery through their markings, such as the extensive white around the face on the chinstrap, and found that a Ground Sample Distance (GSD) of < 5 mm was necessary to identify this feature in all individuals, and a GSD > 20 mm meant it could no longer be clearly identified (Pfeifer et al. 2025). The GSD of our imagery was around 15 mm (see Methods), and so it would be necessary to either reduce the flight height, or use a UAV with a higher quality camera, to clearly identify all individuals. The DJI Mavic Enterprise 3 Thermal aircraft used in our study has a ½-inch CMOS sensor, which is smaller than other comparable models (e.g. the DJI Mavic Enterprise 3 has a 4/3-inch CMOS sensor). Using another model of UAV with a larger sensor would facilitate the capture of higher quality imagery, which would help to improve detection of different species at mixed colonies.

Despite this, there were clear benefits to using UAVs to survey. At Fort Point, part of the chinstrap colony was inaccessible to on-ground survey due to the terrain and temporal constraints, but was able to be quickly surveyed using the UAV. The ability to access previously inaccessible areas using UAVs is a primary benefit of UAV surveys (Pfeifer et al. 2019; Jones et al. 2020) and is especially beneficial in areas such as Antarctica that present several logistical challenges with accessing sites by foot to survey. Further, whilst there is still discussion about the appropriate minimum flight height for surveying gentoo and chinstrap penguins, there is strong consensus that surveys using UAVs, when conducted correctly, are significantly less disturbing to the birds compared to traditional ground survey methods (Krause et al. 2021), which can cause birds to temporarily leave their nests (Dunn et al. 2021).

In our analysis of breeding phenology at Port Lockroy, there was substantial variation in the dates when all three surveys were triggered, as has been previously documented for gentoo penguins (Lynch et al. 2012a; Juáres et al. 2013; Black 2016). All three surveys displayed a trend towards taking place later in the season, although only the change in the timing of the nest survey was statistically significant (Fig. 4, Table 2), possibly suggesting that gentoo penguins may flexibly advance chick fledging date in some circumstances (Hinke et al. 2012). Our findings contrast with other work that has documented an advancement of gentoo phenology in other colonies on the Western Antarctic Peninsula (Lynch et al. 2012a; Juarez Martinez et al. 2026), suggesting changes to phenology are likely to be colony- or region-specific according to local conditions. Surveying at a time in the season that was once considered 'normal' at Port Lockroy may therefore now underestimate the peak population size; however, this may not be true for other colonies.

When assessing the influence of survey timing on survey outcome, our analysis indicated a significant negative effect on the number of hatched chicks and number of crèched chicks for surveys occurring later in the season, but not on the number of nests (Fig. 4; Table 2). This aligns with previous studies, which have shown declines in hatch and crèche rates for clutches initiated later in the season (Hinke et al. 2012). Counts of both the number of hatched and crèched chicks appeared to be influenced by counts at either end of the temporal spectrum, with earlier surveys yielding the highest counts, and the latest surveys yielding the lowest, whilst surveys around the original date had mixed results (Fig. 4). This suggests that there is an optimal window for chicks around the date of the original survey, and that within that window other factors may influence the survey outcome. However, lower chick numbers from surveys occurring later in the season indicates that conditions in those seasons were less favourable towards chick survival. The three seasons with the lowest numbers of hatched and crèched chicks recorded were 2001/02, 2021/22, and 2022/23 (Dunn et al. 2020), and 2001/02 and 2022/23 were also the two latest surveys (Fig. 4). Whilst the cause of this requires further investigation, these were all years that were recorded as having unusually late snowfall (Trathan et al. 2008), which can delay brushtail penguins in establishing their nests (Williams 1990; Hinke et al. 2012; Juáres et al. 2013). It would therefore be beneficial to further investigate the influence of late snowfall on nesting success, and whether late snowfall is becoming more common, which may help explain why the timing of surveys has become later over time. The plasticity of gentoo phenology is likely to be a factor in the increase in their numbers and distribution across the Peninsula (Juáres et al. 2013; Juarez Martinez et al. 2026), however our results suggest that there is a limit to the plasticity, beyond which delayed breeding negatively impacts chick survival (Fig. 4).

Future work

There were four sites where UAV counts differed by more than the 5% threshold for count repeatability (Fig. 3, Appendix 1). It would therefore be beneficial to develop guidelines for UAV survey protocol, image quality and image interpretation to further standardise interpretation and achieve repeatability across all surveys, as are available for traditional on-ground survey methods (CCAMLR 2014). Such guidelines would help prevent further miscounts as occurred during UAV surveys at Port Lockroy (Flynn et al. 2026). For example, minimising the effects of observer bias by specifying that all counts from UAV imagery used in research, or counts being submitted to publicly available databases such as MAPPPD (Oceanites 2023), should ensure that two trained observers reach agreement within 5% before counts are accepted. This is particularly important if counts are used to inform management recommendations.

Our research also documented a shift in the timing of breeding at Port Lockroy (Fig. 4), and negative effects of this trend on the gentoo population (Fig. 4). This result contrasts with previous work (Lynch et al. 2012a; Juarez Martinez et al. 2026), suggesting changes to breeding phenology are likely to be site specific. Renewed investigations into changes to breeding phenology across the Antarctic Peninsula would be beneficial, particularly on the impacts of changing environmental conditions such as late snowfall. Regardless, accounting for breeding phenology is essential when investigating long-term population trends, as phenology is tied to environmental conditions which may help explain inter-annual delays or declines (Hinke et al. 2012; Juáres et al. 2013). Recording the stage in the breeding cycle, such as an estimate of the percentage of nests with eggs or chicks, would therefore provide useful context alongside the date of the survey during opportunistic surveys.

Conclusions

Although all colonies included in our study had increased since their previous survey (Appendix 1), Chinstrap penguin numbers are about 50% of estimates from the 1980s (Strycker et al. 2020), including two of the three Chinstrap colonies included in our study (Half Moon Island and Palaver Point; Appendix 1). Ensuring data collected by UAV surveys is accurate and comparable to historic datasets collected through traditional CEMP methods is therefore vital to track long-term trends.

Our results indicate that an important aspect of ensuring UAV surveys are comparable with ground surveys would involve considering the date of the survey and breeding stage of the colony (Southwell et al. 2015). We documented significant variation in the nesting phenology of the gentoo colony at Port Lockroy, such that surveys conducted on the same date on different years could capture different points in the breeding cycle, making results challenging to compare. As an opportunistic survey method, UAV surveys inevitably miss phenological information, and so we recommend UAV surveys are accompanied by an estimate of the stage in the breeding cycle to ensure the comparisons are robust.

Our findings suggest that, with high quality imagery and properly trained individuals to interpret images, UAV surveys provide an accurate means of surveying brushtail penguin colonies on the Antarctic Peninsula (Fig. 2) and (Fig. 3). It is critical, however, that results obtained from UAV imagery recognize that changes in breeding population numbers obtained from imagery classification be evaluated in the light of the differences and limitations we have described. There will continue to be many instances where only ground counts will be possible, such as when weather or other factors preclude obtaining aerial imagery. Future work should therefore focus on determining how to accurately reflect differences in methodology.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00300-026-03484-z.

Acknowledgements The authors would like to thank all of those who involved with facilitating and assisting with the data collection that made this research possible. We would also like to thank two anonymous reviewers that provided feedback on the manuscript.

Author contribution MH, CN, PNT, GH, and RN conceived and designed research. GH prepared UAV imagery. MH and PNT conducted analysis. MH led the writing of the manuscript. GH, CN, CMF, RN, SF, GCV, and PNT helped write the manuscript. All authors read and approved the manuscript.

Funding This study received funding from the UK Antarctic Heritage Trust.

Data availability The results of all UAV and hand counts of penguin colonies included in this paper can be found at: https://www.penguinmap.com/mapppd/. Data pertaining to the phenology assessment at Port Lockroy can be found at https://portlockroy.data.bas.ac.uk

Declarations

Conflict of interest The authors declare that they have no known competing interests that could have influenced the work reported in this paper.

References

Agnew DJ (1997) Review: the CCAMLR ecosystem monitoring programme. Antarct Sci 9:235–242

Black CE (2016) A comprehensive review of the phenology of Pygoscelis penguins. Polar Biol 39:405–432

Borowicz A, McDowall P, Youngflesh C et al (2018) Multi-modal survey of Adélie penguin mega-colonies reveals the Danger Islands as a seabird hotspot. Sci Rep. https://doi.org/10.1038/s41598-018-22313-w

Bricher PK, Lucieer A, Woehler EJ (2008) Population trends of Adélie penguin (Pygoscelis adeliae) breeding colonies: a spatial analysis of the effects of snow accumulation and human activities. Polar Biol 31:1397–1407. https://doi.org/10.1007/s00300-008-0479-z

CCAMLR (2014) Ecosystem monitoring program standard methods. CCAMLR, Hobart, Australia

Chabot D, Bird DM (2012) Evaluation of an off-the-shelf unmanned aircraft system for surveying flocks of geese. Waterbirds 35:170–174. https://doi.org/10.1675/063.035.0119

Chabot D, Craik SR, Bird DM (2015) Population census of a large common tern colony with a small unmanned aircraft. PLoS ONE 10:e0122588. https://doi.org/10.1371/journal.pone.0122588

Cobley ND, Shears JR (1999) Breeding performance of Gentoo penguins (Pygoscelis papua) at a colony exposed to high levels of human disturbance. Polar Biol 21:355–360. https://doi.org/10.1007/s003000050373

Croxall JP, Kirkwood ED (1979) The distribution of penguins on the Antarctic Peninsula and Islands of the Scotia Sea. British Antarctic Survey, Cambridge, UK

Dewar M, Vanstreels RE, Boulinier T et al (2023) Biological risk assessment of highly pathogenic avian influenza in the Southern Ocean. Scientific Committee on Antarctic Research Antarctic Wildlife Health Network, Australia

Dickens J, Hollyman PR, Hart T et al (2021) Developing UAV monitoring of South Georgia and the South Sandwich Islands' iconic land-based marine predators. Front Mar Sci. https://doi.org/10.3389/fmars.2021.654215

Dunn MJ, Forcada J, Jackson JA et al (2019) A long-term study of Gentoo penguin (*Pygoscelis papua*) population trends at a major Antarctic tourist site, Goudier Island, Port Lockroy. Biodivers Conserv 28:37–53. https://doi.org/10.1007/s10531-018-1635-6

Dunn MJ, Adlard S, Taylor AP et al (2021) Un-crewed aerial vehicle population survey of three sympatrically breeding seabird species at Signy Island, South Orkney Islands. Polar Biol 44:717–727

Dunn M, Nichol C, Forcada J, Trathan P (2020) Population numbers and breeding success of Gentoo penguins (Pygoscelis papua) at Port Lockroy, Goudier Island, 1996 - 2020 (Version 1.0) [Data set]. UK Polar Data Centre, Natural Environment Research Council, UK Research & Innovation. https://doi.org/10.5285/92e92c14-0ff2-4ce2-8305-83d5b3a4071d

Ellenberg U, Edwards E, Mattern T et al (2015) Assessing the impact of nest searches on breeding birds – a case study on Fiordland crested penguins (*Eudyptes pachyrhynchus*). N Z J Ecol 39:231–244

DJI (2024) DJI Terra. DJI Enterprise, v4.0 (EN)

Flynn CM, Hart T, Clucas GV, Lynch HJ (2023) Penguins in the anthropause: COVID-19 closures drive gentoo penguin movement among breeding colonies. Biol Conserv 286:110318

Flynn CM, Wethington M, Hilton M et al (2026) Revisiting "Penguins in the anthropause: Covid-19 closures drive gentoo penguin movement among breeding colonies." Biol Conserv 315:111715. https://doi.org/10.1016/j.biocon.2026.111715

Harris CM, Herata H, Hertel F (2019) Environmental guidelines for operation of remotely piloted aircraft systems (RPAS): experience from Antarctica. Biol Conserv 236:521–531

Herman R, Borowicz A, Lynch M et al (2020) Update on the global abundance and distribution of breeding gentoo penguins (*Pygoscelis papua*). Polar Biol 43:1947–1956. https://doi.org/10.1007/s00300-020-02759-3

Hinke JT, Polito MJ, Reiss CS et al (2012) Flexible reproductive timing can buffer reproductive success of *Pygoscelis* spp. penguins in the Antarctic Peninsula region. Mar Ecol Prog Ser 454:91–104. https://doi.org/10.3354/meps09633

Hodgson JC, Baylis SM, Mott R et al (2016) Precision wildlife monitoring using unmanned aerial vehicles. Sci Rep 6:22574. https://doi.org/10.1038/srep22574

Hodgson JC, Mott R, Baylis SM et al (2018) Drones count wildlife more accurately and precisely than humans. Methods Ecol Evol 9:1160–1167. https://doi.org/10.1111/2041-210X.12974

Howat I, Porter C, Noh M-J et al (2022) The Reference Elevation Model of Antarctica - Strips. Version 4:1

Jones LR, Godollei E, Sosa A et al (2020) Validating an unmanned aerial vehicle (UAV) approach to survey colonial waterbirds. Waterbirds 43(3–4):263–270. https://doi.org/10.1675/063.043.0304

Juáres MA, Santos MM, Negrete J et al (2013) Better late than never? Interannual and seasonal variability in breeding chronology of gentoo penguins at Stranger Point, Antarctica. Polar Res 32:1–11. https://doi.org/10.3402/polar.v32i0.18448

Krause DJ, Hinke JT, Goebel ME, Perryman WL (2021) Drones minimize Antarctic predator responses relative to ground survey methods: an appeal for context in policy advice. Front Mar Sci 8:648772. https://doi.org/10.3389/fmars.2021.648772

Lynch HJ, Naveen R, Fagan WF (2008) Censuses of penguin, blue-eyed shag Phalacrocorax atriceps and southern giant petrel macronectes giganteu populations on the Antarctic Peninsula, 2001–2007. Mar Ornithol 36:83–97

Lynch H, Fagan W, Naveen R et al (2012a) Differential advancement of breeding phenology in response to climate may alter staggered breeding among sympatric pygoscelid penguins. Mar Ecol Prog Ser 454:135–145. https://doi.org/10.3354/meps09252

Lynch HJ, Naveen R, Trathan PN, Fagan WF (2012b) Spatially integrated assessment reveals widespread changes in penguin populations on the Antarctic Peninsula. Ecology 93:1367–1377. https://doi.org/10.1890/11-1588.1

Mustafa O, Barbosa A, Krause DJ et al (2018) State of knowledge: Antarctic wildlife response to unmanned aerial systems. Polar Biol 41:2387–2398. https://doi.org/10.1007/s00300-018-2363-9

Oceanites (2023) MAPPPD v4.1. https://www.penguinmap.com/mapppd/. Accessed 2 Jun 2023

Pfeifer C, Barbosa A, Mustafa O et al (2019) Using fixed-wing UAV for detecting and mapping the distribution and abundance of penguins on the South Shetlands Islands, Antarctica. Drones 3:39. https://doi.org/10.3390/drones3020039

Pfeifer C, Knetsch S, Maercker J et al (2025) Exploring the potential of aerial drone imagery to distinguish breeding Adélie (Pygoscelis adeliae), chinstrap (Pygoscelis antarcticus) and gentoo (Pygoscelis papua) penguins in Antarctica. Ecol Indic. https://doi.org/10.1016/j.ecolind.2024.113011

Pina P, Vieira G (2022) UAVs for Science in Antarctica. Remote Sens. https://doi.org/10.3390/rs14071610

QGIS Development Team (2024) QGIS Geographic Information System. Open Source Geospatial Foundation. Project. http://qgis.osgeo.org

Qian Y, Humphries GRW, Trathan PN et al (2023) Counting animals in aerial images with a density map estimation model. Ecol Evol. https://doi.org/10.1002/ece3.9903

R Core Team (2024) R: A language and environment for statistical computing. Vienna, Austria. https://www.R-project-org

Ratcliffe N, Guihen D, Robst J et al (2015) A protocol for the aerial survey of penguin colonies using UAVs. J Unmanned Veh Syst 3:95–101. https://doi.org/10.1139/juvs-2015-0006

Robinson NM, Scheele BC, Legge S et al (2018) How to ensure threatened species monitoring leads to threatened species conservation. Ecol Manage Restor 19:222–229. https://doi.org/10.1111/emr.12335

Rümmler MC, Mustafa O, Maercker J et al (2018) Sensitivity of Adélie and Gentoo penguins to various flight activities of a micro UAV. Polar Biol 41:2481–2493. https://doi.org/10.1007/s00300-018-2385-3

Rush GP, Clarke LE, Stone M, Wood MJ (2018) Can drones count gulls? minimal disturbance and semiautomated image processing with an unmanned aerial vehicle for colony-nesting seabirds. Ecol Evol 8:12322–12334. https://doi.org/10.1002/ece3.4495

Santos MM, Hinke JT, Coria NR et al (2018) Abundance estimation of Adélie penguins at the Esperanza/Hope Bay mega colony. Polar Biol 41:2337–2342. https://doi.org/10.1007/s00300-018-2373-7

Southwell C, Emmerson L, McKinlay J et al (2015) Spatially extensive standardized surveys reveal widespread, multi-decadal increase in East Antarctic Adélie penguin populations. PLoS ONE. https://doi.org/10.1371/journal.pone.0139877

Strycker N, Wethington M, Borowicz A et al (2020) A global population assessment of the Chinstrap penguin (*Pygoscelis antarctica*). Sci Rep 10:19474. https://doi.org/10.1038/s41598-020-76479-3

Trathan P (2004) Image analysis of color aerial photography to estimate penguin population size. Wildl Soc Bull 32:332–343

Trathan PN, Forcada J, Atkinson R et al (2008) Population assessments of gentoo penguins (*Pygoscelis papua*) breeding at an important Antarctic tourist site, Goudier Island, Port Lockroy, Palmer Archipelago, Antarctica. Biol Conserv 141:3019–3028. https://doi.org/10.1016/j.biocon.2008.09.006

Williams TD (1990) Annual variation in breeding biology of gentoo penguins, Pygoscefis papua, at Bird Island, South Georgia

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.