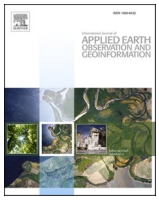
International Journal of Applied Earth Observation and Geoinformation 134 (2024) 104124

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



Assessing topographic features and population abundance in an Antarctic penguin colony through UAV-based deep-learning models

Oleg Belyaev [a,*], Alejandro Román [a], Josabel Belliure [b], Gabriel Navarro [a], Luis Barbero [c], Antonio Tovar-Sánchez [a]

[a] Department of Ecology and Coastal Management, Institute of Marine Sciences of Andalusia, Spanish National Research Council, Spain
[b] Global Change Ecology and Evolution Group (GloCEE), Department of Life Sciences, University of Alcalá, Spain
[c] Department of Earth Sciences, International Campus of Excellence in Marine Science (CEIMAR), University of Cádiz, Spain

ARTICLE INFO

Keywords:
UAV
Deep-learning
Chinstrap penguin
Pygoscelis antarcticus
Breeding site features
Vapour col

ABSTRACT

Penguins play an essential biochemical role in the Antarctic ecosystem, being the study of their dynamics of utmost importance to understand their environment, behaviour and populational trends in the current climate change scenario. In this study, we used multi-rotor Unmanned Aerial Vehicles (UAVs) along the coast of the chinstrap penguin (*Pygoscelis antarcticus*) colony of Vapour Col (Deception Island, Antarctica) to map potential sites of biochemical interactions with the surrounding sea water. Several runoff discharge points were identified, where a precise placing of environmental sampling station is suggested. Additionally, UAVs were used in combination with Object Detection Architectures to obtain the chinstrap colony population size. Applying a simulation for clutch initiation dates due to our off-laying peak count, we obtained an estimated range of 13,250 to 22,000 breeding pairs in the 2021/2022 breeding season, also suggesting an alternative approach using chinstrap chicks as proxy to estimate adult numbers. This research shows the utility of UAV-deep learning for environment characterization and wildlife monitoring, providing a solid framework for upcoming studies in the area.

1. Introduction

Antarctic ecosystems are inhabited by a unique and widely interconnected wildlife, which ultimately makes this continent an idoneous location for conducting research at the biosphere level. Additionally, the underlying dynamics of the species interactions with their environments make some of the Antarctic taxa a suitable proxy to understand the trending of the polar ecosystems health, therefore resulting essential to conduct in-depth research on specific species (Ainley, 2002; Huang et al., 2013). In fact, recent studies have addressed the impacts of climate change on Antarctic wildlife, evaluating the current and future associated risks (Rogers et al., 2020), and highlighting the relevance of assessing its present and upcoming conservation state.

One of the group of species that best characterize the Antarctic continent are penguins. They are key mesopredators that represent the greater portion of bird biomass in this southern polar regions (Woehler, 2002), and are not exempt from the global change effects. Indeed, climate change is one of the main causes of an interspecific disparity among penguin populations in the Antarctic Peninsula, as revealed by their population trends (Clucas et al., 2014; Lynch et al., 2012). Consequently, penguins have been broadly studied from multidisciplinary perspectives, addressing their diet (Lynnes et al., 2004; Rombolá et al., 2010; Volkman et al., 1980), phenology (Black, 2016), behaviour (Juáres et al., 2018; Lynnes et al., 2004), environment (Román et al., 2022; Zmarz et al., 2018) and distribution (Ballard et al., 2010; Santora et al., 2020; Trathan et al., 1996) features. More specifically, the chinstrap penguin (*Pygoscelis antarcticus*) accounts for one of the largest populations of Antarctic penguins, with close to 8 million individuals distributed mainly along the Antarctic Peninsula and adjacent islands (BirdLife International, 2022), contrary to the first species in abundance of the *Pygoscelis* genus, the Adélie penguin (*Pygoscelis adeliae*), whose population is uniformly scattered along the Antarctic coast (Mapping Application for Penguin Populations and Projected Dynamics [WWW Document], 2022; Strycker et al., 2020). The breeding season of chinstrap penguins typically begins in late October to early November, marked by nest building and egg laying. Chicks hatch around December,

* Corresponding author.
E-mail address: o.belyaev@csic.es (O. Belyaev).

https://doi.org/10.1016/j.jag.2024.104124
Received 19 April 2024; Received in revised form 19 August 2024; Accepted 25 August 2024

and fledging occurs by February. Their breeding phenology is influenced by sea ice conditions and prey availability, which are being increasingly impacted by climate change. Furthermore, their geographical pattern of distribution also suggests a higher vulnerability, as local changes in this area could significantly affect a large portion of their colonies.

Among the areas inhabited by chinstraps, Deception Island (South Shetland Islands, Maritime Antarctica) offers unique topographic characteristics due to its volcanic condition, and hosts almost 80,000 breeding pairs, split mainly into two colonies. The largest one (Baily Head) is located in the eastern coast, and houses close to 60,000 breeding pairs (Naveen et al., 2012). In the opposite site of the island, Vapour Col is located at the south-western coast, and has been surveyed from the late 80′s. In the last decades, fieldwork at Vapour Col estimated the population size in almost 20,000 breeding pairs (Mapping Application for Penguin Populations and Projected Dynamics [WWW Document], 2022; Naveen et al., 2012), noting however the existence of a population decline of 36 % between 1991 and 2008 (Barbosa et al., 2012). In terms of chinstrap penguin census at Vapour Col, there have not been any documented attempts since 2011, when Naveen et al. (2012) performed a direct-observation estimation, setting the number at 19,177 breeding pairs. Indeed, such ground-based observational census techniques have been widely used for animal quantification during the last decades, mainly due to technological constraints and the relatively non-complex nature of the observation task. However, using conventional methods for population estimation often results difficult to implement in a frequent, non-invasive and high-precision quantitative assessment of wildlife populations, and more specifically in the case of penguins, due to their morphological and behavioural features. For example, ground-based counting often requires reaching remote breeding areas that suppose a logistical challenge, affecting therefore to quantification accuracy and increasing the economic cost and invested time (Dickens et al., 2021).

To address the limitation of these techniques, Unmanned Aerial Vehicles (UAVs) emerge as a useful tool providing high-quality aerial photographic data that enable a wide range of applications, like precise census of the individuals within a colony (Guo et al., 2018; Hodgson et al., 2017) or ecosystem characterization (Borowicz et al., 2018; Pfeifer et al., 2019; Román et al., 2023; Zmarz et al., 2018). In parallel to the increasing popularity of UAV usage in field surveys, machine learning (ML) has experienced an exponential growth in the very recent years, leading to a development of state-of-the-art models able to incorporate and integrate large amounts of complex, non-parametric information, learning from data extracted from observations. The combination of UAV-based information and ML algorithms such as object detection models, converge in a powerful tool for wildlife populational dynamics and their ecotope characterization. For example, the chinstrap penguin has recently shown to be an important biological source of trace elements such as iron to the Antarctic waters, contributing to the Southern Ocean upper layer fertilization (Belyaev et al., 2023). This role of the species in trace metal ecology opens the way to investigate how precisely these metals are discharged from the breeding sites to the surrounding waters, and which is the correlation between the population distribution and this trace metal release. For instance, the rugged topography of the Vapour Col penguin colony favours the terrestrial surface runoff, driven by the amount of day-to-day rainfall, and containing large quantities of penguin guano and material eroded from the island's bedrock. The runoff dynamics of guano into the ocean directly and indirectly impact the entire Antarctic food web, from primary producers to predators. Therefore, it is important to determine the main discharge routes into the ocean to assess nutrient distribution and residence time in the photic zone. Consequently, UAV-ML approach to the parametrization of intrinsic biophysical features of a chinstrap colony could potentially lead to a better understanding of the mechanisms behind the ecological relevance of the species in the Southern Polar environment.

Therefore, in this study we implemented the use of UAV-mounted optical RGB and multispectral sensors in combination with deep-learning models to characterize the main geomorphological features, the population size and the intraspecific maturity-state based distribution (adults and chicks) at the chinstrap colony of Vapour Col. This broad-site physical and behavioural analysis ultimately enable a more precise, non-invasive and cost-efficient research that helps to understand the ecological role of the chinstrap penguin within the Antarctic ecosystem.

2. Material and methods

2.1. Study area

Deception Island (62°55′S 60°37′W) is a volcanic island located in the South Shetland Islands archipelago, between parallels 61°S and 63°S, and 120 km north of the Antarctic Peninsula. It is constituted by a central caldera known as Port Foster, which is covered by seawater and connects with Bransfield Strait by Neptune's Bellow channel. The island's geomorphology is characterized by volcanic slopes, ash-covered glaciers and smoky beaches (Smith et al., 2003), and has been shaped by continuous and violent volcanic eruptions since its origin (Angulo-Preckler et al., 2021). In this study, the Vapour Col (62°59′S 60°44′W, Fig. 1) penguin colony has been monitored with UAV technology. It is settled on a seasonal ice-free surface surrounded by an abrupt slope, where depressions act as channels both for penguins to go to/from the nests, and for guano runoff to the ocean.

2.2. UAVs and sensors

In this study, two different UAVs were deployed for data collection:

– The DJI (Dà-Jiāng Innovations) Mavic 2 Enterprise Advanced (M2EA) quadcopter, with a takeoff weight of 909 g, has a maximum flight autonomy of 31 min (which can be reduced depending on weather conditions), making this its main limiting factor for covering large areas. However, its 297 g batteries are lightweight, facilitating the transport of some of them to cover larger areas by conducting more flights. Its small size when folded (214 × 91 × 84 mm) makes it easy to transport in the field, while its wind resistance (maximum gusts of 10 m/s) and an operating temperature range between −10 and 40 °C makes it an ideal platform for facing the extreme environmental conditions of Antarctica. It was equipped with its incorporated 48 MP, 1/2″ CMOS optical RGB sensor.

– The DJI Matrice 300 (M300) quadcopter, with a takeoff weight of 9 kg, has a maximum flight time of 55 min (without payload, which results in approximately 35 min with equipped sensors, depending on weather conditions). The weight of its batteries (1.35 kg each, with two equipped) and the dimensions of the folded equipment (430 × 420 × 430 mm) make transportation more challenging under the extreme Antarctic environmental conditions. Additionally, it can withstand wind gusts of up to 15 m/s and operate within a temperature range of −20 to 40 °C. Unlike the M2EA, it can be equipped with larger and heavier sensors, such as the dual multispectral MicaSense RedEdge-MX sensor used in this study, which contains a total of 10 spectral bands centered in the blue (444 nm and 475 nm), the green (531 nm and 560 nm), the red (650 nm and 668 nm), the red edge (705 nm, 717 nm and 740 nm), and the near-infrared (NIR, 840 nm) wavelengths. Light conditions and solar angle changes during the flight are considered and controlled by using the MicaSense's own Downwelling Light Sensor with a built-in GPS. Finally, a calibration panel (RP05-2025214-OB) was used before each UAV flight for radiometric calibration.

2.3. Data collection

The entire Vapour Col penguin colony was surveyed during the Spanish Antarctic Campaign 2021–2022, on January 26, 2022, covering approximately 64 ha. At this time of the year, chinstrap penguin adults

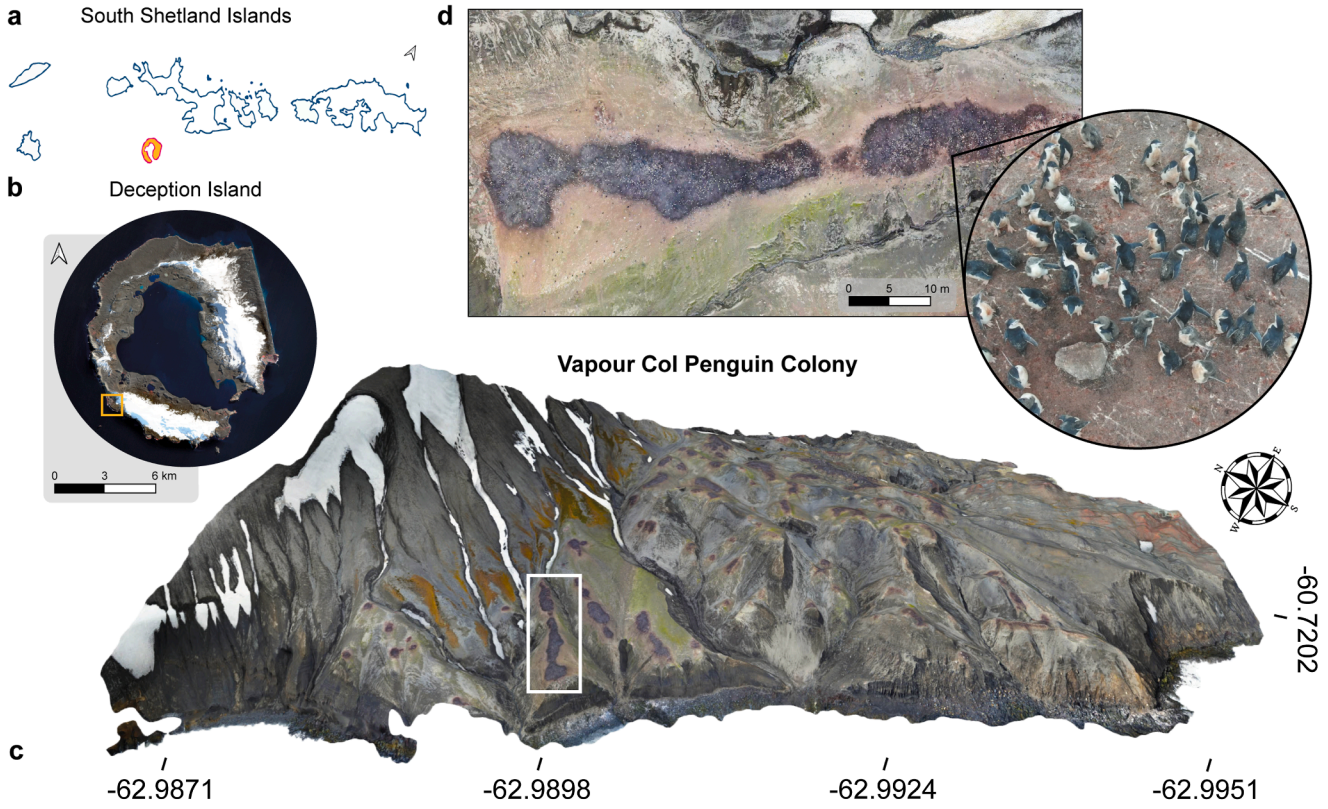


Fig. 1. Vapour Col chinstrap penguin colony. a) South Shetland Islands Archipelago, Deception Island is highlighted in orange. b) Sentinel-2 satellite image of Deception Island, captured on 17 March 2023. Orange box marks the location of Vapour Col, in the Southwestern outer coast of the island. c) 3D photogrammetric composite of Vapour Col, as captured using the DJI Mavic 2 Enterprise Advanced UAV. Within the white box, a typical shape of a guano-rich zone, indicating the presence of chinstrap penguins. d) Zoomed white box from panel c, where a close-up view allows to differentiate each individual from aerial photography. Note that the circular image is a reference photography of the appearance of a typical chinstrap penguin nesting area, having been taken later in the breeding season, when most of the chicks have already undergone moulting. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

are immersed in the reproductive season (October to March; typical peak of egg-laying (PEL) at mid-November; modal clutch size of two eggs), growing nestlings to reach nest independence and subsequent formation of "crèches", a noticeable pattern of chick aggregation in the colony prior to molting and fledging. The meteorological conditions during the UAV surveys were dominated by completely covered skies and strong permanent winds, and the flights were performed in visual line of sight (VLOS). Flight missions were planned using the UgCS desktop software (SPH engineering, Latvia, v.4.14). This software took into account the terrain's topographic features and established a series of fixed parameters for the waypoint plans of the flights conducted in this study, such as the flight altitude above sea level (ASL), speed (4 m/s), duration (between 15 and 20 min per flight), trajectory (parallel to the coast), and a 70–80 % front and side overlap for all sensors across the entire study area. Consequently, the flight altitude varied during the flight to maintain a constant Ground Sample Distance (GSD) with the changing topographic features, averaging 0.96 cm/px in the final RGB orthomosaic and 80.7 cm/px in the Digital Elevation Model (DEM), including the six different flights at heights between 30 and 40 m (AGL) that overlapped to cover the entire Vapour Col colony with the M2EA. On the other hand, the M300 equipped with the MicaSense sensor, achieved an estimated GSD of 13.9 cm/px for each multispectral band. Licensed UAV pilots (A1/A3, A2, and STS licenses according to the European and Spanish Civil Aviation Regulations) followed the guidelines provided by Hodgson & Koh and the Scientific Committee on Antarctic Research (Hodgson and Koh, 2016; SCAR, 2017) to ensure minimal perturbance to seabirds so that UAV surveys posed minimal environmental impacts. In this sense, a research conducted in 2021 where they studied chinstrap penguin (among others) disturbance by UAVs and found that, behavioral responses to UAVs overflights at 30 m were not different from control periods (Krause et al., 2021).

The software Agisoft Metashape v.1.8.4 (Agisoft LLC, St. Petersburg, Russia) was used for a Structure from Motion (*Sf*M) photogrammetry process to generate final reflectance orthomosaics and topographic products for each flight. After importing all UAV captures, a sparse point cloud was built at the "capture alignment" step with the highest accuracy setting. Then, an "aggressive" depth filter and "ultra-high" quality settings were applied for the "3D dense cloud" generation. An interpolated DEM was generated from the "3D dense cloud". Finally, the orthomosaic was rendered using the DEM as a reference surface. The resulting orthomosaics were projected into the coordinate system WGS84 / UTM zone 20S (EPSG: 32720).

2.4. Guano stains delimitation

The characteristic brownish appearance of guano stains, which is visually unmistakable with the resolution provided by the multispectral sensor on board the UAV (and even with finer satellite remote sensing spatial resolutions), was delimited using the machine learning algorithm Support Vector Machine for supervised classification. This algorithm was selected since it applies kernel functions that allow the mapping of statistically large datasets into a higher-dimensional space where a hyperplane located between them aims to correctly divide different classification classes (Bahari et al., 2014; Vapnik, 2000). Miranda et al. (2020) suggested that using a radial basis function as a kernel parameter provides the best results when working with optically complex systems,

as demonstrated in Román et al. (2022) for Antarctic penguin colonies. The software SAGA GIS (Conrad et al., 2015) was used to perform image classification, using a training shapefile consisting of 25 manually drawn polygons of up to 1 m^2 maximum for guano stains. All MicaSense multispectral bands were used as input, and the radial basis function was the selected kernel function.

After image classification, the results were imported into QGIS (QGIS Association, 2024) to filter out misclassification errors (threshold: groups of 10 pixels), and to polygonise the filtered guano stains classification raster. To assess the accuracy of the image classification, an error matrix and statistical coefficients were calculated following Olofsson et al. (2014). This methodology compares the classification raster with "real" maps using a stratified random sampling design. The statistical coefficients calculated were Overall Accuracy (OA), providing the number of pixels properly classified from the total of image pixels; User Accuracy (U-Acc), that indicates the probability that a predicted value is properly classified; Producer Accuracy (P-Acc), that shows the probability that a given value is properly classified; and the Cohen's Kappa Index, that calculates the compromise between the classification and ground truth values manually defined (Congalton, 2009) (See Supplementary Table S1 for confusion matrix and accuracy assessment).

2.5. Surface runoff modelling

Surface runoff modeling was performed around guano stains to calculate the probability of water/guano flow propagation routes and terminal length. The QGIS software was used to perform this task using DEM analysis. QGIS includes an extensive toolbox available for hydrological analyses, located in the SAGA hydrology toolbox. In this part of the methodology, we obtained four different hydrological measurements:

– Catchment area, which represents the drainage basin where water flows over the topographic terrain. For this study, the Multiple Flow Direction (Wolock and McCabe, 1995) model and the original DEM were set as input parameters, assuming that flow acts downslope from any selected point.

– Channel network, which represents the drainage networks where water flows over the topographic terrain. A threshold of 5 m^2 for flow direction and channel initiation, as well asthe original DEM, were used as input parameters. The output map shows pixels with true values where flow drains, while pixels with false values represent no drainage pixels.

– Flow direction, which represents the main route followed by water flow over the study area depending on elevation that determines each pixel location. The original DEM and a minimum slope of 0.01 degrees were set as input parameters. It also uses the "Fill Sinks" tool within the original DEM.

– Watershed basins, which are the different areas (polygons) of the original DEM were water would accumulate (Broster, 2006). It uses the channel network of the DEM, and the original DEM as input parameters.

2.6. Deep learning-based penguin count

The state-of-the-art YOLOv8 object detection architecture was selected to perform an accurate count of the chinstrap penguin individuals within the Vapour Col colony. To prepare the training dataset, a region of interest (ROI) of 8,501 × 7,717 pixels was extracted from the original optical RGB orthomosaic, where both development stages of the chinstrap penguins (adults and chicks) can be clearly identified. Furthermore, to add greater contrast heterogeneity between penguins and the ground, the substrate on the training data also varied between vegetation, guano and bare soil. The ROI was subsequently divided in tiles of 640 × 640 pixels (only those which were of that exact size were used), where each penguin varied between ~ 20 and ~ 40 pixels (Fig. 2). Roboflow online services (Roboflow, Inc. [WWW Document], 2023) were used to annotate each image for 2 classes of individuals: adult, representing the dark-feathered pattern of adult penguins; and chick, representing the gray-feathered young individuals who had not yet undergone molting. Each image was augmented by performing horizontal and vertical flipping, for a total of 2 images per training example. A final training dataset of 349 images was generated, with a train/validation/test distribution of 301/32/16 images, respectively. The dataset was exported using the YOLOv8 annotation format. Training was performed on a local machine, with a 24 GB NVIDIA RTX4090 GPU with CUDA 11.8. The chosen model was YOLOv8x, training a batch size of 32 for 50 epochs. Subsequently, once the model was evaluated, the original Vapour Col orthomosaic was divided using QGIS software to create 21,378 tiles, most of which were 640 × 640 pixels (some were smaller due to the non-perfect rectangular shape of the original colony orthomosaic). Each image was then inferenced with the trained model, using a detection threshold of 0.5. Then, OpenCV (Python resize image) was used to mask the terrain and keep only the detected instances for each class, which were transformed later to a multi-point shapefile in QGIS, obtaining thus a precise location for every individual within the colony (Fig. 2). To calculate the uncertainty associated to the model count, another representative region of Vapour Col was isolated and subjected to 10 detections. The number of adults and chicks was manually counted to obtain the comparative dataset. To obtain a reasonable estimate of the census error, false negatives and false positive data was used, calculated from model metrics (Supplementary Table S2), obtaining a total error of ~ 13 % for adults and ~ 25 % for chicks.

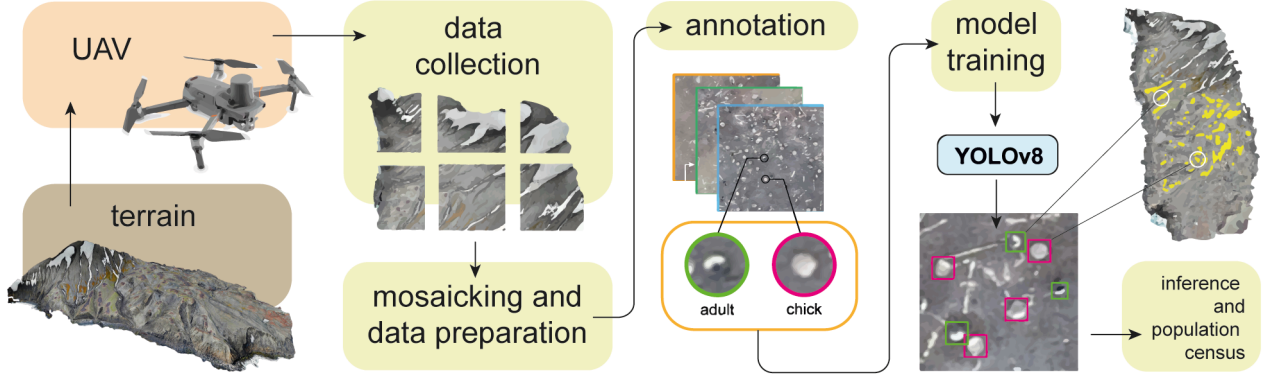


Fig. 2. Schematic description of penguin count using UAV-deep learning combination. Once the photographic data is collected using the UAV, the orthomosaic is created. Specific representative areas within the colony are selected to create a training dataset, where chinstrap penguin adults and chicks are represented over different substrates, adding variability. After training, the model is deployed over the entirety of the colony orthomosaic, obtaining the count and distribution of each penguin.

2.7. Phenology-corrected count

As the aerial survey in this study was performed in late January, the nest occupancy was significantly lower than in the PEL. Therefore, the count of the present research is phenologically adjusted using the approach followed by Naveen et al. (2012), where they correct Shuford and Spear's census in early 1987 (Shuford and Spear, 1988) using a stochastic simulation based on Clutch Initiation Dates (CID) (Lynch et al., 2009) (See Supplementary Table S3 and S4 for calculations). The approximate number of breeding pairs present at the highest nest occupancy is then drawn, typically occurring in mid to late November (Black, 2016). This will ultimately allow to account for representative and comparable numbers of the chinstrap penguins in Vapour Col, the first since Naveen et al. in 2011/2012 breeding season (Naveen et al., 2012). This approach first calculates the mean CID based on mean October temperature, effect of latitude, CID trend and the baseline CID for the chinstrap penguin. Once the mean CID is obtained, the lag between the PEL and the model count (mc) made in this study is calculated.

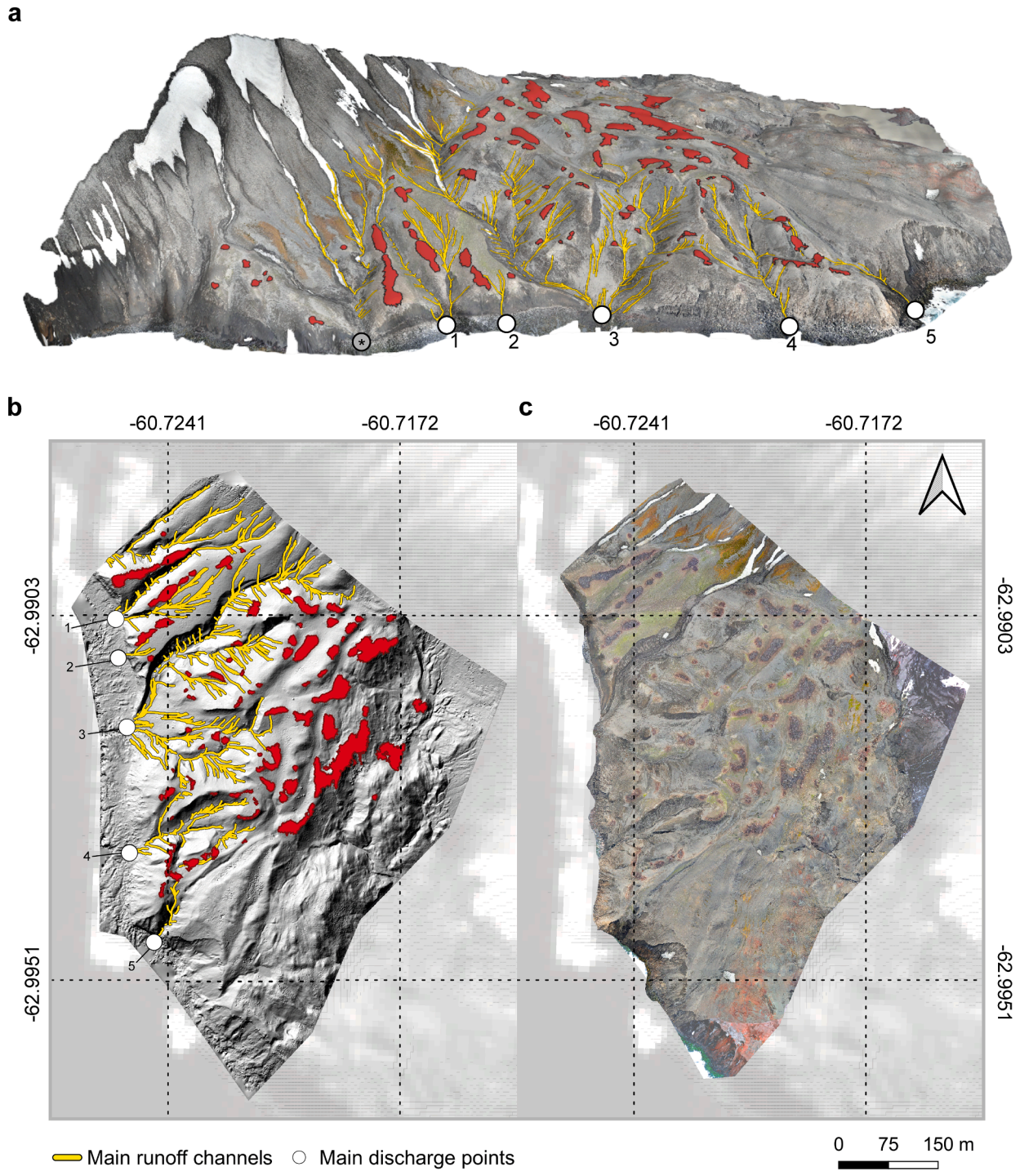


Fig. 3. a) 3D reconstruction of the Vapour Col penguin colony in optical RGB colours. In yellow, the main surface runoff channels retrieved from the DEM information; in red, the main identified guano-stained areas. The labelled white points indicate the main discharge points along the coastline. Separation between: b) the shapefile containing the main runoff channels and the main discharge points, and c) the optical RGB orthomosaic. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The nest attrition rate for the chinstraps (Lynch et al., 2009) is then used to calculate the expected number of breeding pairs during the PEL, i.e., corrected count (cc), which is given by:

$$cc = \frac{mc}{1 - r \cdot d}$$

where r is the nest attrition rate and d is the number of days since the PEL. Although the nest attrition rate and the days since the PEL have an associated error, these are not accounted in error propagation. This is due to the large, exponential weight the nest attrition rate has in the overall error when the count is performed significantly later than the PEL (Supplementary Fig. S1). Three parameters were used to alternatively approximate the number of breeding pairs in Vapour Col: total detected adults, which accounted for every detected instance made by the model within the colony; adults detected only within the guano stains, and all detected chicks, which were used to calculate the number of breeding pairs (assuming the modal clutch size of two offspring per pair).

3. Results

3.1. Surface runoff-based guano release dynamics

Fig. 3 displays the results of applying surface runoff modelling to the UAV-based DEM over Vapour Col penguin colony, highlighting up to five primary surface runoff routes (Supplementary Table S5). The main guano stains (Fig. 3a) are directly connected to these routes, so rainfall would favour the direct deposition of guano onto the surface runoff. In addition, they may be supplemented by secondary channels depending on the relief characteristics, potentially leading to significant discharges into the marine environment and consequently enriching nutrients in the adjacent coastal zone. Interestingly, although the most extense guano stains are located in the center-east of the penguin colony, these reside in a flatter terrain than those located in the western part of Vapour Col, with small drainage channels spread across the terrain, that do not seem to converge at greater canals flowing towards the ocean.

3.2. Chinstrap penguin count and distribution in Vapour Col

According to the model metrics, the model has performed significantly well in detecting true positives, both for adults and chicks of the chinstrap penguin (for a total error of ~ 13 % for adults and ~ 25 % for chicks). However, as expected, the proportion of false positive chicks yielded by the prediction is also high (~9.3 %), mainly due to the colour of their feathers. In this sense, Vapour Col coastal area is completely covered by rocks, most of which have similar sizes and colour to those of the chicks. To therefore obtain an accurate estimation of the actual number of chicks, we visually examined the entire colony orthomosaic to better understand their distribution, concluding that there was almost a complete absence of chicks outside the guano-stained regions. We consequently masked the preliminary obtained chick count yielded by the model using the guano-stained areas, keeping only those individuals that where inside these guano-rich regions (Fig. 4a, Supplementary Fig. S2a). Regarding adults, we re-examined the colony to assess their distribution, comparing it to the model prediction, observing that according to the good model performance, their characteristic dark colour and shape made adult individuals well distinguishable from other environmental features, allowing the precise estimation of their numbers. Notably, unlike the chicks, adults were distributed inside and outside the guano-rich areas (Fig. 4b, Supplementary Fig. S2b), reflecting the characteristic foraging behaviour during the breeding season, when adults continuously move between the nests and the sea. Given the large gap existing between the calculated mean CID (15 November 2021, see Supplementary Table S1 and S2 for the exact data used in mean CID

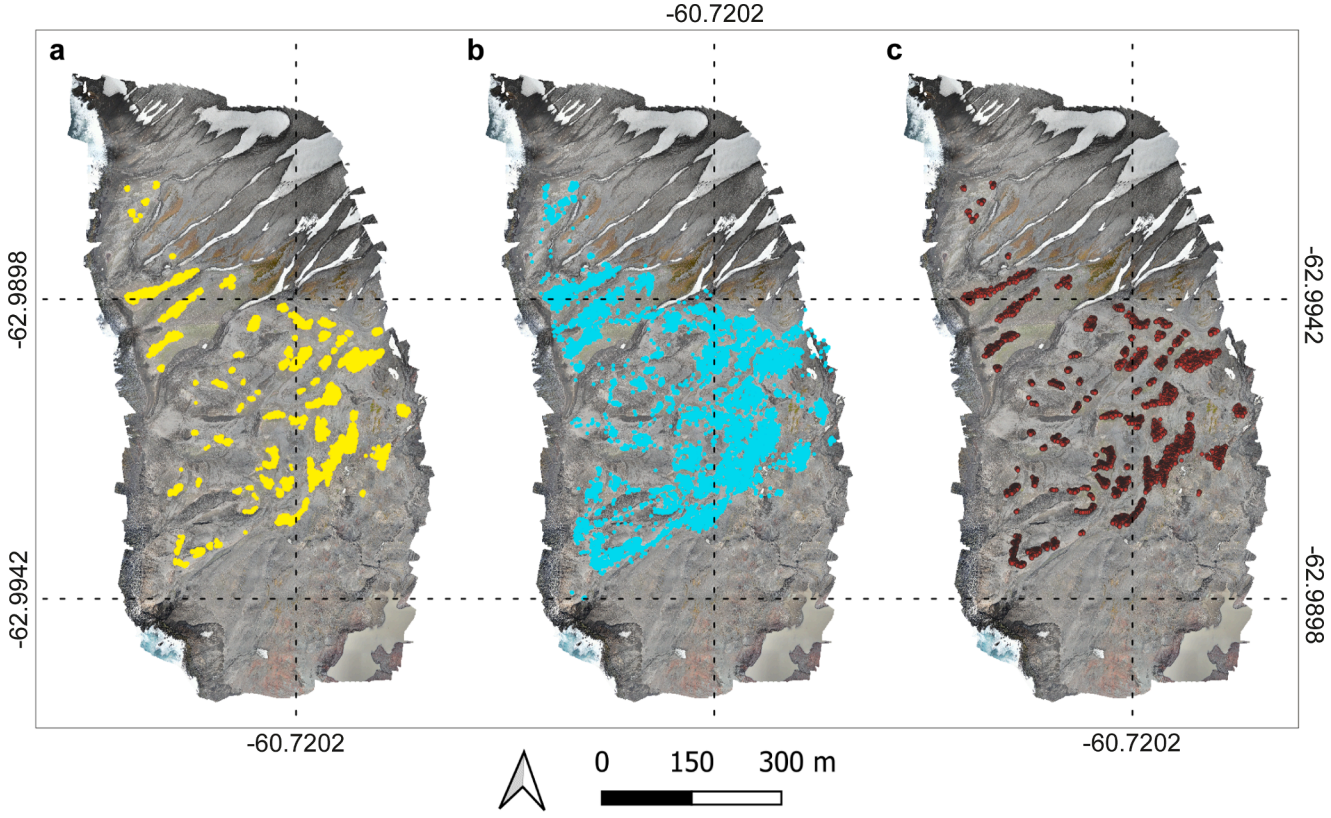


Fig. 4. Three detection outputs considering the distribution of chinstrap penguin individuals. a) Detected chicks are mainly found within the guano stains, with barely no instances neither detected or visually observed outside. b) Adults were detected within the guano stains and also in transit to and from these regions and water. c) Isolated adult individuals in guano stains, if assumed that at least one individual of the breeding pair is always present at the former nest.

Table 1
Comparison between the chinstrap penguin numbers detected by the object detection model at sampling time and the corrected numbers using CID model by Lynch et al. (2009) that can be expected during PEL.

Source of information	Object Detection Model count (penguins recorded at sampling time)	Corrected count (expected penguins at PEL)
Total adults (TA)	12,097 ± 1,573	43,208 ± 2,808
TA Breeding pairs	6,049 ± 787	21,604 ± 1,404
Adults within guano stains (AGS)	7,420 ± 297	26,500 ± 512
AGS Breeding pairs	3,710 ± 297	13,250 ± 256
Total detected chicks	12,357 ± 3,090	–
Adults extrapolated from chicks (AC)	12,357 ± 3,090	44,136 ± 5,516
AC Breeding pairs	6,179 ± 1,545	22,068 ± 2,758

calculations), and the count date, and considering the breeding cycle of the species, it is expectable that by the time the aerial data was acquired one or neither of the parents were present at the nest, being in or in transit to the sea. To address this issue, we used all three types of available information inferred form the count to obtain the corrected counts (Table 1): 12,097 ± 1,573 total adult individuals detected by the model, 7,420 ± 297 adult individuals present only in the guano stains (Fig. 4c), and 12,357 ± 3,090 adults extrapolated from the number of total counted chicks (12,357 ± 3,090 detected chicks) (as our model does not differentiate between breeding and non-breeding individuals, being also the count date late in the breeding season, we considered every adult as part of a breeding pair, consequently the number of breeding pairs will be half of the count presented above). As expected from the original model count, after applying mean CID model using breeding pairs as input, we obtained a similar corrected count of 21,604 ± 1,404 and 22,068 ± 2,758 breeding pairs when using total adults and total chicks as a primary source of information respectively. If only assuming that at least one adult per breeding pair is always present at the guano-stained area and not counting those adults outside the guano stains, we approximate the number of breeding pairs to be 13,250 ± 256.

4. Discussion

Understanding the breeding dynamics of the chinstrap penguins at Vapour Col colony on Deception Island is not only important to evaluate their conservation status and assess their present and long term populational trends, but also to better comprehend the biochemical role they play as fertilizers in the Southern Ocean. In this sense, the identification of potential guano discharge points is key to establish coastal environmental study stations directly linked to nutrient (Bosman and Hockey, 1986; Shatova et al., 2016) and metals such as iron (Belyaev et al., 2023) increases associated with seabird guano. From a logistical perspective, Vapour Col has a cliff zone connected to the sea on its western side, making it inaccessible directly from the coast. Consequently, the existence of water sampling devices with UAVs (Sparaventi et al., 2022; Tovar-Sánchez et al., 2021) requires knowledge of this valuable information to reduce efforts, economic expenses, and time during the surveys. In addition, it has been observed that guano runoffs are not evenly distributed across main identified routes, but vary based on the topography of the area or size of the guano (Gagnon et al., 2013). In this sense, stations 1 and 3 constitute the best candidates to be ideal sampling collection points at the chinstrap colony of Vapour Col. Station 1 concentrates the efflux of the three northern well-defined colonies, where the flow of guano, meltwater and sediments occurs most likely at high velocities due to the inclination of the terrain, probably allowing the accumulation of fresher samples down the hill. On the other hand, station 3 serves as the endpoint for the lixiviates coming from the colonies situated on the less abrupt Vapour Col's central hills, where the flow should be slower than in the northern section, having the efflux enough time to infiltrate in the slope ground and contributing to the development of soils of ornithogenic nature. These assumptions open the opportunity for upcoming research on site, involving the study of Vapour Col soil geochemical characteristics, taking advantage of the suggested sampling points. Additionally, the proposed stations could be extended uphill through transects in order to effectively assess the motion dynamics of the lixiviates.

Regarding penguin population size, similarly to Shuford and Spear, (1988), the UAV surveys performed in this study to determine the number of chinstrap individuals at the colony of Vapour Col occurred significantly later than the peak of egg-laying, implying that the true number of breeding pairs could be far higher. The model used to correct this count and obtain an estimation of the true number of breeding individuals generated a reasonable approximation; however, certain limitations in counting penguins in such a late stage of the breeding season have to be noted. For example, when propagating error derived from the model counting, the nest attrition rate and the number of days since the PEL, a direct correlation is observed between days since PEL and nest attrition rate errors. This propagation results in an exponential increase in uncertainty, as the nest attrition rate contributes most to the overall estimation error (Supplementary Fig. S1). Consequently, this count should not be considered as representative of the number of chinstrap penguins at Vapour Col due to two main reasons. First, the timing of the flights. It is difficult to determine which of the adult individuals belong to the breeding pair and if both adults are in the water, with the chicks left in *chrèches*, i.e., offspring caring by the colony. Furthermore, the nests that were clearly distinguishable during the PEL are no longer visible, which neither contributes to breeding pairs identification. Due to these limitations, the expected breeding pairs during the PEL from total adults and adults in guano stains were calculated, revealing a significant difference of more than 8,000 breeding pairs (Supplementary Table S3 and S4). Second, the previous constraints added to the nest attrition error propagation significantly amplify the uncertainty associated to the final count. However, counting chicks as a proxy to extrapolate the true number of adults yielded an interesting cross-validation to estimate the true number of breeding pairs, as the obtained number of breeding pairs closely matches the estimation made accounting for all adult individuals, 21,604 ± 1,404 and 22,068 ± 2,758, respectively. For this reason, it is considered that the true number of breeding pairs at Vapour Col colony at the peak of the 2021/2022 season is most likely to be close to 22,000 breeding pairs. One important caveat of this assumption that inclines the balance in favour of the lower end breeding pairs estimate is the population decline experienced by this specific colony. This count is consistent with the census of Naveen et al. (2012), during the 2011/2012 season, implying that the population of chinstrap penguins has not declined during the last decade, even experiencing a slight increase. If a 2.11 % linear decline trend in population is assumed (obtained from a 36 % population decrease at Vapour Col in a span of 17 years; (Barbosa et al., 2012), the resulting number of breeding pairs in 2021/2022 breeding season would be close to 15,500, considering the 2011/2012 census of Naveen et al. (2012). From these results and accounting for the uncertainties, a uniform distribution of the chinstrap penguin population at Vapour Col during the 2021/2022 breeding season has been suggested, ranging between 13,250 and 22,000 breeding pairs.

5. Conclusion

Our research provides important insights into the discharge dynamics of the breeding colony, identifying potential guano-discharge points in the surrounding waters, which can help establishing biochemical research sampling stations, guiding also the implementation of UAV water sampling devices, reducing significantly the effort, cost and time spent in the preparation of surveys. This study also presents a practical use of UAVs in combination with Object Detection

Architectures that enable an efficient monitoring and analysis of penguin populations in Antarctica. While our study does not provide an official and definitive count, it offers a reasonably approximated range of chinstrap penguin breeding pairs in Vapour Col during the 2021/2022 season, a range that is also consistent with the previous studies and serves as a perspective on the current population numbers of chinstrap penguins in that colony. To provide an accurate census using the methods mentioned in this study, future research should focus on improving the timing, frequency, and accuracy of the UAV data acquisition, which will help to drastically reduce model uncertainties, allowing the creation of robust colony occupancy curves and analysis of populational trends. Despite the limitations, this research aims to contribute to the population status evaluation and conservation efforts of the chinstrap penguin, improving our understanding of their ecological role.

Credit authorship contribution statement

Oleg Belyaev: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alejandro Román:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Josabel Belliure:** Writing – review & editing, Validation, Supervision, Methodology, Investigation. **Gabriel Navarro:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition. **Luis Barbero:** Writing – review & editing, Validation, Supervision, Methodology. **Antonio Tovar-Sánchez:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research has been funded by the Spanish Government projects PiMetAn (ref. RTI2018-098048-BI00), DICHOSO (PID2021-1257830B-100), University of Cádiz EQC2018-004446-P, CSIC EQC2018-004275-P and EQC2019-005721. O. Belyaev is supported by the Spanish Predoctoral Grant (Ref: PRE2022-103391). A. Román is supported by the Spanish FPU Grant (Ref: FPU19/04557). J. Belliure is supported by the grant PID2019-108597RB-100 (PERPANTAR project) from the Spanish Agency of Research (Spain). This research is part of the POLARCSIC research initiatives. We warmly thank the military staff of the Spanish Antarctic Base Gabriel de Castilla, the crew of the BIO Hespérides oceanographic vessel and the Marine Technology Unit (UTM-CSIC) for their logistic support, without which the XXXV Spanish Antarctic campaign and this research would not have been possible. We also thank David Roque for his support with the UAV operations.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2024.104124.

References

Ainley, D.G., 2002. The Adélie penguin: bellwether of climate change. Columbia University Press, New York.

Angulo-Preckler, C., Pernet, P., García-Hernández, C., Kereszturi, G., Álvarez-Valero, A. M., Hopfenblatt, J., Gómez-Ballesteros, M., Otero, X.L., Caza, J., Ruiz-Fernández, J., Geyer, A., Avila, C., 2021. Volcanism and rapid sedimentation affect the benthic communities of deception Island, Antarctica. Cont. Shelf Res. 220, 104404 https://doi.org/10.1016/j.csr.2021.104404.

Bahari, N.I.S., Ahmad, A., Aboobaider, B.M., 2014. Application of support vector machine for classification of multispectral data. IOP Conf. Ser. Earth Environ. Sci. 20, 012038 https://doi.org/10.1088/1755-1315/20/1/012038.

Ballard, G., Toniolo, V., Ainley, D.G., Parkinson, C.L., Arrigo, K.R., Trathan, P.N., 2010. Responding to climate change: Adélie Penguins confront astronomical and ocean boundaries. Ecology 91, 2056–2069. https://doi.org/10.1890/09-0688.1.

Barbosa, A., Benzal, J., De León, A., Moreno, J., 2012. Population decline of chinstrap penguins (Pygoscelis antarctica) on Deception Island, South Shetlands, Antarctica. Polar Biol. 35, 1453–1457. https://doi.org/10.1007/s00300-012-1196-1.

Belyaev, O., Sparaventi, E., Navarro, G., Rodríguez-Romero, A., Tovar-Sánchez, A., 2023. The contribution of penguin guano to the Southern Ocean iron pool. Nat. Commun. 14, 1781. https://doi.org/10.1038/s41467-023-37132-5.

BirdLife International (2022) Species factsheet: Pygoscelis antarcticus. [WWW Document], 2022. . BirdLife Int. URL http://www.birdlife.org/.

Black, C.E., 2016. A comprehensive review of the phenology of Pygoscelis penguins. Polar Biol. 39, 405–432. https://doi.org/10.1007/s00300-015-1807-8.

Borowicz, A., McDowall, P., Youngflesh, C., Sayre-McCord, T., Clucas, G., Herman, R., Forrest, S., Rider, M., Schwaller, M., Hart, T., Jenouvrier, S., Polito, M.J., Singh, H., Lynch, H.J., 2018. Multi-modal survey of Adélie penguin mega-colonies reveals the Danger Islands as a seabird hotspot. Sci. Rep. 8, 3926. https://doi.org/10.1038/s41598-018-22313-w.

Bosman, A.L., Hockey, P.A.R., 1986. Seabird guano as a determinant of rocky intertidal community structure. Mar. Ecol. Prog. Ser. 247–257.

Broster, B.E., 2006. Geomorphology: a canadian perspective, second edition: (Alan S. Trenhaile). Environ. Eng. Geosci. 12, 83–84. https://doi.org/10.2113/12.1.83.

Clucas, G.V., Dunn, M.J., Dyke, G., Emslie, S.D., Levy, H., Naveen, R., Polito, M.J., Pybus, O.G., Rogers, A.D., Hart, T., 2014. A reversal of fortunes: climate change 'winners' and 'losers' in Antarctic Peninsula penguins. Sci. Rep. 4, 5024. https://doi.org/10.1038/srep05024.

Congalton, R.G., 2009. Accuracy and error analysis of global and local maps: Lessons learned and future considerations., in: Remote Sensing of Global Croplands for Food Security. pp. 47–55.

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geosci. Model Dev. 8, 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015.

Dickens, J., Hollyman, P.R., Hart, T., Clucas, G.V., Murphy, E.J., Poncet, S., Trathan, P. N., Collins, M.A., 2021. Developing UAV monitoring of south georgia and the south sandwich islands' iconic land-based marine predators. Front. Mar. Sci. 8, 654215 https://doi.org/10.3389/fmars.2021.654215.

Gagnon, K., Rothäusler, E., Syrjänen, A., Yli-Renko, M., Jormalainen, V., 2013. Seabird guano fertilizes baltic sea littoral food webs. PLoS ONE 8, e61284.

Guo, X., Shao, Q., Li, Y., Wang, Y., Wang, D., Liu, J., Fan, J., Yang, F., 2018. Application of UAV remote sensing for a population census of large wild herbivores—taking the headwater region of the yellow river as an example. Remote Sens. 10, 1041. https://doi.org/10.3390/rs10071041.

Hodgson, J.C., Koh, L.P., 2016. Best practice for minimising unmanned aerial vehicle disturbance to wildlife in biological field research. Curr. Biol. 26, R404–R405. https://doi.org/10.1016/j.cub.2016.04.001.

Hodgson, A., Peel, D., Kelly, N., 2017. Unmanned aerial vehicles for surveying marine fauna: assessing detection probability. Ecol. Appl. 27, 1253–1267. https://doi.org/10.1002/eap.1519.

Huang, T., Sun, L., Long, N., Wang, Y., Huang, W., 2013. Penguin tissue as a proxy for relative krill abundance in East Antarctica during the Holocene. Sci. Rep. 3, 2807. https://doi.org/10.1038/srep02807.

Juáres, M.A., Casaux, R., Corbalán, A., Blanco, G., Pereira, G.A., Perchivale, P.J., Coria, N.R., Santos, M.M., 2018. Diet of Adélie penguins (Pygoscelis adeliae) at stranger point (25 de Mayo/King George Island, Antarctica) over a 13-year period (2003–2015). Polar Biol. 41, 303–311. https://doi.org/10.1007/s00300-017-2191-3.

Krause, D.J., Hinke, J.T., Goebel, M.E., Perryman, W.L., 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, H.J., Fagan, W.F., Naveen, R., Trivelpiece, S.G., Trivelpiece, W.Z., 2009. Timing of clutch initiation in Pygoscelis penguins on the Antarctic Peninsula: towards an improved understanding of off-peak census correction factors. CCAMLR Sci. 16, 149–165.

Lynch, H.J., Naveen, R., Trathan, P.N., Fagan, W.F., 2012. 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.

Lynnes, A.S., Reid, K., Croxall, J.P., 2004. Diet and reproductive success of Adélie and chinstrap penguins: linking response of predators to prey population dynamics. Polar Biol. 27 https://doi.org/10.1007/s00300-004-0617-1.

Mapping Application for Penguin Populations and Projected Dynamics [WWW Document], 2022. URL. https://www.penguinmap.com/.

Miranda, V., Pina, P., Heleno, S., Vieira, G., Mora, C., E.G.R. Schaefer, C., 2020. Monitoring recent changes of vegetation in Fildes Peninsula (King George Island, Antarctica) through satellite imagery guided by UAV surveys. Sci. Total Environ. 704, 135295 https://doi.org/10.1016/j.scitotenv.2019.135295.

Naveen, R., Lynch, H.J., Forrest, S., Mueller, T., Polito, M., 2012. First direct, site-wide penguin survey at Deception Island, Antarctica, suggests significant declines in breeding chinstrap penguins. Polar Biol. https://doi.org/10.1007/s00300-012-1230-3.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015.

Pfeifer, C., Barbosa, A., Mustafa, O., Peter, H.-U., Rümmler, M.-C., Brenning, A., 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.

QGIS Association, 2024. QGIS Geographic Information System. http://www.qgis.org/.

Roboflow, Inc. [WWW Document], 2023. . Roboflow Inc. URL https://roboflow.com/.

Rogers, A.D., Frinault, B.A.V., Barnes, D.K.A., Bindoff, N.L., Downie, R., Ducklow, H.W., Friedlaender, A.S., Hart, T., Hill, S.L., Hofmann, E.E., Linse, K., McMahon, C.R., Murphy, E.J., Pakhomov, E.A., Reygondeau, G., Staniland, I.J., Wolf-Gladrow, D.A., Wright, R.M., 2020. Antarctic futures: an assessment of climate-driven changes in ecosystem structure, function, and service provisioning in the southern ocean. Annu. Rev. Mar. Sci. 12, 87–120. https://doi.org/10.1146/annurev-marine-010419-011028.

Román, A., Navarro, G., Caballero, I., Tovar-Sánchez, A., 2022. High-spatial resolution UAV multispectral data complementing satellite imagery to characterize a chinstrap penguin colony ecosystem on deception island (Antarctica). Gisci. Remote Sens. 59, 1159–1176. https://doi.org/10.1080/15481603.2022.2101702.

Román, A., Tovar-Sánchez, A., Fernández-Marín, B., Navarro, G., Barbero, L., 2023. Characterization of an antarctic penguin colony ecosystem using high-resolution UAV hyperspectral imagery. Int. J. Appl. Earth Obs. Geoinform. 125, 103565 https://doi.org/10.1016/j.jag.2023.103565.

Rombolá, E.F., Marschoff, E., Coria, N., 2010. Inter-annual variability in Chinstrap penguin diet at South Shetland and South Orkneys Islands. Polar Biol. 33, 799–806. https://doi.org/10.1007/s00300-009-0757-4.

Santora, J.A., LaRue, M.A., Ainley, D.G., 2020. Geographic structuring of Antarctic penguin populations. Glob. Ecol. Biogeogr. 29, 1716–1728. https://doi.org/10.1111/geb.13144.

SCAR, 2017. State of Knowledge of Wildlife Responses to Remotely Piloted Aircraft Systems (RPAS). Scientific Committee on Antarctic Research.

Shatova, O., Wing, S.R., Gault-Ringold, M., Wing, L., Hoffmann, L.J., 2016. Seabird guano enhances phytoplankton production in the Southern Ocean. J. Exp. Mar. Biol. Ecol. 483, 74–87. https://doi.org/10.1016/j.jembe.2016.07.004.

Shuford, W.D., 1988. Spear. Surveys of breeding Chinstrap Penguins in the South Shetland Islands, Antarctica.

Smith, K., Baldwin, R., Kaufmann, R., Sturz, A., 2003. Ecosystem studies at Deception Island, Antarctica: an overview. Deep Sea Res. Part II Top. Stud. Oceanogr. 50, 1595–1609. https://doi.org/10.1016/S0967-0645(03)00081-X.

Sparaventi, E., Rodríguez-Romero, A., Navarro, G., Tovar-Sánchez, A., 2022. A novel automatic water autosampler operated from UAVS for determining dissolved trace elements. Front. Mar. Sci. 9, 879953 https://doi.org/10.3389/fmars.2022.879953.

Strycker, N., Wethington, M., Borowicz, A., Forrest, S., Witharana, C., Hart, T., Lynch, H. J., 2020. A global population assessment of the Chinstrap penguin (Pygoscelis antarctica). Sci. Rep. 10, 19474. https://doi.org/10.1038/s41598-020-76479-3.

Tovar-Sánchez, A., Román, A., Roque-Atienza, D., Navarro, G., 2021. Applications of unmanned aerial vehicles in Antarctic environmental research. Sci. Rep. 11, 21717. https://doi.org/10.1038/s41598-021-01228-z.

Trathan, P.N., Croxall, J.P., Murphy, E.J., 1996. Dynamics of antarctic penguin populations in relation to inter-annual variability in sea ice distribution. Polar Biol. 16, 321–330. https://doi.org/10.1007/BF02342178.

Vapnik, V.N., 2000. The Nature of Statistical Learning Theory. Springer New York, New York, NY. doi: 10.1007/978-1-4757-3264-1.

Volkman, N.J., Presler, P., Trivelpiece, W., 1980. Diets of Pygoscelid Penguins at King George Island, Antarctica. The Condor 82, 373. doi: 10.2307/1367558.

Woehler, E.J., 2002. A statistical assessment of the status and trends of Antarctic and Subantarctic seabirds. Scientific Committee on Antarctic Research, Kingston, Tas.

Wolock, D.M., McCabe, G.J., 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. Water Resour. Res. 31, 1315–1324. https://doi.org/10.1029/95WR00471.

Zmarz, A., Rodzewicz, M., Dąbski, M., Karsznia, I., Korczak-Abshire, M., Chwedorzewska, K.J., 2018. Application of UAV BVLOS remote sensing data for multi-faceted analysis of Antarctic ecosystem. Remote Sens. Environ. 217, 375–388. https://doi.org/10.1016/j.rse.2018.08.031.