# Classifying sperm whale (*Physeter macrocephalus*) sex and age classes using aerial photogrammetry

Ana Eguiguren1\*, David Gaspard1, Christine Konrad1, Hal Whitehead1

1Biology Department, Dalhousie University, Halifax, Nova Scotia, Canada

\*Corresponding author: [anaeguibur@gmail.com](mailto:anaeguibur@gmail.com)

## 1 | INTRODUCTION

## 2 | METHODS

### 2.1 | Data Collection

We carried out dedicated surveys in the deep waters (> 1000 m) off the Galápagos Islands aboard a 12.03 m sailboat (*Balaena*) between January and May 2023 (under research permit No. PC-86-22). We searched for sperm whales acoustically using a 100 m towed hydrophone and visually during daylight hours. When we encountered groups of females and juveniles, we followed them for as long as possible at a cautious distance to collect behavioural, acoustic, and photo-identification data. When single males were detected, we approached for a brief data collection but abandoned them after they deep dove.

If conditions were permissive (windspeed < 10 kts and no rain), we conducted 1 – 2 hour flight sessions using a DJI Mini 2 drone (249 g) equipped with propeller guards and landing gear. We conducted sessions in the morning and afternoon when glare in the water interfered the least with visibility. Once we approached a group of whales with the drone, we flew between 15 - 120 m above the water and pointed the camera perpendicularly over the whales. During flights, we recorded continuously at 29.79 fps at 1080p or 4K resolution. We alternated a group-follow protocol–during which we kept visual contact with a group of whales by flying high enough to fit all whales in the frame (Altmann 1974)–with brief moments of close approach (15 - 20 m)–to capture individual whales’ distinctive marks and allow for more accurate size estimates. At the end of each flight, we hovered over the research vessel to collect a calibration image (see below).

We quality-rated drone footage on a scale of 0 – 8, with 0 being high quality and 8 being low quality, based on the level of glare, sea-surface disruption, focus, and exposure. Only recordings with a quality rating ≤ 4 were included in the analysis. We extracted where whales were lying mostly flat at the water surface, located near the center of the frame, and where the drone camera was positioned at nadir relative to the water surface stills using the behavioural analysis software BORIS (Friard & Gamba 2016). Individual whales were identified based on observable markings—including visible fluke marks, indentations, rake marks, white patches, and sloughed skin patterns (O’Callaghan et al. 2024).

### 2.2 | Whale morphometry

We used MorphoMetriX V2 (Torres & Bierlich 2020) to measure the length in pixels () of the research vessel and whales. For each whale, we measured total length (TL), head-to-flipper length (HF), and head-to-dorsal fin length (HD) (Figure 1). We converted measures to true lengths (*L*) in meters by applying the following equation, modified from Burnett et al. (2019):

where *H* is the drone altitude above sea level, and α is a scaling corresponding to the DJI Mini 2 drone camera. While this value can be computed based on known camera parameters—focal length and pixel dimensions (Torres & Bierlich 2020)—these values were not available for our drone model. Therefore, we empirically estimated α by obtaining measurements of a known object of known length *L* and known distance *H* using equation 2.

#### 2.2.1 | Quantifying and correcting measurement error

Errors in aerial photogrammetry arise from several sources, of which the most impactful are imprecise altitude estimates (Burnett et al. 2019, Bierlich et al. 2021, Glarou et al. 2022, Napoli et al. 2024). Drones that derive altitude measurements from inbuilt barometers, as was our case, can be inaccurate due to changes in meteorological conditions and internal biases (Burnett et al. 2019, Bierlich et al. 2021). We used measurements of our research vessel collected throughout the field season at various altitudes (27 – 120 m) to quantify the uncertainty in morphometric measurements and correct altitude estimates. We quantified percent measurement error using a modified version of Bierlich et al. 2021 as:

Where is the known length of the calibration object (12.03 m), and is the estimated length in meters of the the calibration object in each image .

To compute a corrected altitude estimate, we first calculated the true altitude given the for each still image of the research vessel and its known length .

Next, we applied a mixed-effects linear regression with date as a random effect to estimate a corrected altitude ( given the barometer altitude across measurements for each day:

We used corrected altitude estimates to compute whale morphometry.

#### 2.2.1 | Assessing the reliability of morphometric measurements

We evaluated the degree to which TL, HD, and HF could be measured reliably between frames and across days, we measured estimated the coefficient of variation (CV) in measurements taken from the same individual in 3 different frames in the same video, and of the same whale observed in different days (Christiansen et al. 2018). We examined the relationship between observed CV and altitude and Q rating to determine an optimal cutoff at which morphometric features could be measured reliably.

### 2.3 | Classification of sex/age classes

#### 2.3.2 | Age/Sex Classification

FlexMix model (Leisch 2004). chrome-

<https://statmath.wu.ac.at/~gruen/BayesMix/bayesmix-intro.pdf>

## Results

## Discussion

* Different nose/body ratios may influence length estimation based on IPI’s (Christine)
* Measurement of uncertainty can be incorporated into demographic models based on aerial photogrammetry
* How can one transfer our findings when using other drone models