Text S1: Size Estimation Field Tests

To assess the accuracy of the shark length estimations, we conducted overwater field tests with target objects of known length placed in the water at our study site. With the target at the surface, we flew the UAV at altitudes between 10 and 60 m at 10 m intervals, recording 30 seconds of video at each altitude. We then repeated this procedure with the target submerged on the seafloor at a depth of 1.5 m. Based on results from these trials, we compared the accuracy of three different length estimation methods. Treating both takeoff elevation and target depth as increases to the distance between the UAV and target, we calculated target length using 1) UAV altitude above takeoff only, 2) UAV altitude plus takeoff elevation, and 3) UAV altitude plus takeoff elevation and target depth.

We demonstrated that obtaining accurate estimates of length from UAV video footage is feasible, as long as the appropriate factors are taken into consideration. Perhaps the most important step in calibrating a UAV-based photogrammetric approach to length estimation is determining the characteristics of the UAV and camera being used in a field setting (Whitehead et al 2022), such as camera FOV.

Results from our photogrammic calibration trials offer some useful insight into methods that can improve the precision of shark size estimations generated with this common and accessible UAV hardware platform. Our analyses would suggest that including takeoff elevation above sea level and the depth of the shark or other target being measured should also be incorporated into the length calculation for these estimates to be accurate. An interesting caveat to this conclusion is that these factors, which are fixed values, become relatively less important at higher altitudes. This is demonstrated by the high accuracy of uncorrected length measurements from 50 and 60 m in our field tests (**Figure S4**). Therefore, a tradeoff must be considered between the relative importance of correction factors, which decreases with increasing flight altitude, and the ability to distinguish and measure the animal of interest, which becomes more error-prone as altitude increases. This error is exemplified in the overestimation of target size at altitudes of 40, 50, and 60 m, where the corrected size estimates were, on average, 6.5, 8.2, and 9.5 cm too long **(Figure S4)**. Pixel ground length at 40 m altitude is 1.9 cm, and increases to 2.9 cm at 60 m altitude, so our overestimates correspond to an error in the roaming image-measurement process of one or two pixels on either side of the target.

The most likely source of the variation in our repeat measurements is inaccuracy in the UAV altitude reading. We obtained this value from the UAV barometer, whereas studies using LIDAR to measure UAV altitude have reported measurement errors (of whales) smaller than ours by an order of magnitude (Dawson et al. 2017, Christiansen et al. 2018). However, Durban et al. (2016) also used the UAV's factory barometer and reported variability in whale length measurements of $\leq 5\%$, and LIDAR systems also can contribute to photogrammetric error (Dawson et al. 2017).

Our field tests were conducted when there was high water visibility and with a white PVC target, so higher altitude measurements were relatively accurate. The reduced ability to identify objects in the water at high altitudes is offset by the increased duration that an object remains in the frame at a given flight speed and higher-altitude flights also cover more area for a given UAV flight speed. Therefore, we recommend that UAV surveys be conducted at the highest altitude (within FAA and local regulations). When visibility is poor, reduced altitudes may be necessary to make accurate photogrammetric

measurements, and correction factors must therefore be determined with tests conducted with the specific UAV to be used in the study, and at the study site. Furthermore, very low altitudes tended to underestimate target length, possibly because of lens distortion (Burnett et al. 2018). We did not explicitly correct for this in our study, but any lens distortion would have been constant between our FOV and length calibration field tests and our measurements of sharks, and was therefore unlikely to have affected measurement accuracy. Studies focusing on smaller organisms or where conditions such as water visibility necessitate low altitudes should also account for this effect.

Size frequency estimates of all sharks measured using these methodologies in 2020 and 2021 are shown in **Figure S5.**

Text S2: Inshore vs. offshore transect comparison

While we have binned all inshore and offshore data in our analyses, we do note that our data shows more sharks in the shallow inshore transect than in the deeper offshore transect ($p \le 0.001$). Specifically, across 2020 and 2021, we observed a mean density of 29.4 sharks/km² in inshore sharks compared to 3.7 sharks/km² in the offshore transect. It seems most likely this is a real result of shark preference for inshore transects and not an an artifact of detectability/visibility as inner transects had poor visibility on 62.1% of surveys, compared to 38.7% on offshore transect surveys (χ^2 = 25.308, p < 0.001), inconsistent with a visibility artifact. Furthermore, many sharks at this particular site have been determined to spend the majority of their time in surface waters less than 2 m from the surface, regardless of water column depth (Spurgeon et al 2024). However, as noted above, we can only say that sharks are more likely to use surface waters in inshore habitat; we have no ability to comment on their use of the full offshore habitat, and transects are binned in all other analyses in this manuscript.

Text S3: Survey methods comparison

Observed shark density (sharks/km²) was compared between the two survey methods (roaming vs. belt transects) and no statistically significant difference was observed ($p = 0.286$, $n = 221$). Median observed density from roaming surveys was 6.12 sharks/km² (SD = 19.33) (mean=11.97), compared to a median observed density from transect surveys of 5.80 sharks/ $km^2(SD = 16.35)$ (mean =12.04). These data are shown in **Figure S2.**

Text S4: Determining Shark Length

To determine shark length we first drew a bounding box around each identified shark, choosing a frame where the shark was fully extended and close to the surface . From that bounding box (which was oriented such that the width of the box ran the length of the shark) we calculated shark length by using the box dimensions and pixel size. This provided a raw estimate of shark length not including depth corrections, which were then applied based on categorically defined depth categories (1.125 m for shallow depth sharks and 2.25 m for deep depth sharks). Sharks beyond 3 m were not measured or included in analyses (as already noted in main text). Coding was done in R, and image measurement was done with ImageJ.

Field of View $(FOV) = 1.274090354$ radians (73 degrees) Aspect $(A) = 16/9$ (aspect ratio of video frames)

Calculating pixel size

Frame height (FH) = 2(altitude)*tan(FOV/2))/sqrt(1+aspect)² Frame width $(FW) = FH * A$ Pixel size (PS)= FW/2688[MOU9]

Calculating shark length

Box height (in pixels) of AI generated box $(BH) = (maxY - minY)$, in which maxY and minY are the max and min y coordinates of the box

Box width (in pixels) of AI generated box $(BW) = (maxX - minX)$, in which maxX and minX are the max and min x coordinates of the box

Shark length uncorrected (SL, in m) = sqrt($BH^2 + BW^2$) * PS

A length correction factor was applied based on depth of sharks. While there is some variation in correction factor used based on original altitude and depth of shark, for a standard altitude of 20m this resulted in a shallow water correction of 1.05625 and a deep correction factor 1.1125.

Text S5: Sea Surface Temperatures vs. Sea Floor Temperatures

Relationships between sea surface temperature and observed shark density (**Figure S6**) were weaker (only 2.5% vs. 8.7% of deviance explained) than those for sea floor temperature and were opposed in direction to those between sea floor temperature and shark density, with higher shark density being observed in warmer sea surface conditions rather than cooler conditions. Notably there was no relationship between SST and shark density for large sharks. With regard to other environmental parameters, SST models differed from SFT models in that SST models included no significant effect of visibility or sea state for any group of sharks, but these effects were also very weak in the SFT models, so not qualitatively very different. There was a difference in the effect of year across models, for SST year was significant for small sharks while for SFT models it was significant for large sharks.

Figure S1. UAV images from the Carpinteria aggregation site of A) four white sharks in close proximity to one another in an area offshore of the surfline and B) one white shark in the inshore region (center-left of frame) just outside of the surf zone, with the shore to the right of the frame.

Figure S2: Comparison of raw shark count data, observed shark density (count/km²), and rate (count/minute) between roaming and transect surveys ($n = 221$ for each metric/method combination). The median difference between roaming and transect counts from surveys conducted on the same day was significantly greater than zero (paired Wilcoxon signed-rank test, $p < 0.01$). However, there was no significant difference between roaming and transect surveys with regards to observed shark density or rate.

Figure S3: Total shark count was determined by human manual integration of automated belt transect surveys and roaming surveys. Given that human surveys have some chance of multiple detection we checked for correlation between these values and found them to be very high.

Figure S4. Results of length estimation field testing, showing the effect of altitude and correction method on length estimates of a 1.97 m target (true length indicated by the red dashed line) submerged 1.5 m below the ocean surface. Correction method "None" used only the altitude of the UAV above the takeoff point in the length calculation, while "ASL" incorporated the "altitude above sea level" of the takeoff point and "ASL and depth" incorporated both altitude above sea level of the takeoff point and the depth of the target (see Methods). Box plots are generated with length estimates from 30 still frames of each altitude/correction method combination.

Figure S5: Frequency size distribution of all sharks measured in 2020 and 2021; there is a known biological break around approximately 3 m between juvenile morphology and behavior and sharks below this are classified as "small sharks" (shaded blue) in these analyses while sharks above this are classified as large sharks (shaded grey). The subset of sharks less than 2m could potentially be young of year and are shaded in a lighter blue.

Figure S6: Model outputs for small, large, and all sharks, all based on seafloor temperature models. Each panel within the figure shows each individual smoothed model component across a range of a model parameter. The parameter is on the x axis while the y axis shows the smoothed output with the estimated degrees of freedom shown next to the smoothed covariate. Panels B and C within each figure are the modeled days by year for 2020 and 2021 respectively. The x axis has hash marks indicating observed data values to show the range of the parameter.

Small sharks:

Large Sharks

All Sharks

Figure S7: Relationship between sea surface temperature and modeled density in the overall (all observed sharks), small shark (< 3 m TL), and large shark (> 3 m TL) generalized additive models. Shaded regions represent the 95% confidence interval for each model.

Figure S8: Observed trend in density throughout the year (both years combined) for all, small, and large sharks. Shaded areas are 95% Confidence intervals.

Figure S9: Effects of time of day on shark density.

Figure S10: Smoothed daily Sea Floor Temperature (A) and Sea Surface Temperature (B) during the study period

Table S1: Explanatory variables used in the generalized additive models of white shark density. Sea state was categorized as "low" (Beaufort scale 0-1) or "high" (2+), and visibility was categorized as "low", "medium", or "high".

*Temperature loggers included the Innovasea Rx-LIVE acoustic receiver (0.01°C resolution, +/- 0.2°C accuracy), Innovasea aquaMeasure sensor (0.01°C resolution, +/- 0.2°C accuracy), ElectricBlue EnvLogger (0.1°C resolution, +/- 0.2°C accuracy), HOBO Stowaway TidbiT (0.1°C resolution, 1°C accuracy), and Innovasea VR2Tx (0.1°C resolution, +/- 0.5°C accuracy).

Table S2: Total deviance explained (DE) and DE of each temporal and oceanographic term in Generalized Additive Models of observed density of all sharks and density of observed small (<3 m TL) and large (>3 m TL) white sharks using sea surface temperature (SST) data. DE in bold type indicates that a term was significantly associated with density ($p < 0.05$ in the respective model output). DE is not reported when total DE was higher in the reduced models than in the global model.

*Day of year was significantly associated with large shark density in 2021 only.

Table S3: Total deviance explained (DE) and DE of each temporal and oceanographic term in Generalized Additive Models of observed density of all sharks and density of observed small (<3 m TL) and large (>3 m TL) white sharks using sea floor temperature data from transect surveys only. DE in bold type indicates that a term was significantly associated with density ($p < 0.05$ in the respective model output). DE is not reported when total DE was higher in the reduced models than in the global model.

*Day of year was significantly associated with large shark density in 2021 only.

Table S4: Total deviance explained (DE) and DE of each temporal and oceanographic term in Generalized Additive Models of observed density of all sharks and density of observed small (<3 m TL) and large (>3 m TL) white sharks using sea floor temperature data from roaming surveys only. DE in bold type indicates that a term was significantly associated with density ($p < 0.05$ in the model ANOVA). DE is not reported when total DE was higher in the reduced models than in the global model.

Table S5: To further explore the patterns identified by GAM models and quantified by DE (results in main text) we also analyzed these data with an AIC approach and here provide AIC values and adjusted r square for the Total, Large, and Small shark models (using Sea Floor Temperature (SFT) models) when removing each model term. The two modeling approaches align extremely well in terms of final conclusions with only modest differences. For ease of comparison, DE results (from Table 1) are also included here. For both datasets terms are bolded that were identified as significant.

Adjusted R2

Deviance Explained

