Supplementary Material

1 Supplementary Data

S1. Tag deployments and Dive Data

Tags were deployed from a laser air rifle with air pressure between 20-25 bars. Figure S1.1 shows images of one of the tags used, the tag cupped in the arrow at the end of the rifle, and images of tags attached to whales. Water soluble tape was used to lightly attach the tags to the arrow so that in the event a shot missed, the tag wouldn't fall out of the arrow and sink before it could be retrieved. Short 2lb test line was used to loop around the base of the barbs and the cup of the arrow to further prevent the tag from falling out if the arrow missed the whale. On all successful deployments where the tag made contact with the whale, the tape and line easily broke. The arrow was equipped with a small foam float at the end to retrieve the arrow after each shot.

Diving information was recorded and transmitted from a subset of 14 tags that were equipped with a pressor sensor. Tags transmitted dive information daily before switching to a duty cycle, depending upon the project objectives (Table S1.1). Start and end of dives were determined using a wet/dry sensor. Once the tag registered as 'wet', the tag had to cross a depth threshold of 30 m and last more than 30 sec before it was classified as a dive. If it did not exceed the 30 m depth threshold or 30 sec duration, the activity was classified as part of the surface interval. Depth was recorded at a resolution of 0.5 m for tags in 2010 (SWsat12-16) and 1 m for all other tags (2013-2016) (Table S1.1). Dive duration (T) was defined as the time spent below the 30 m qualifying depth or greater than the 30 sec qualifying duration. Duration of surface intervals was defined as the elapsed time between dives greater than the dive threshold qualifiers. The shape of dives was classified by the tag as either square, 'V', or 'U', and transmitted along with the other dive parameters. Assuming the bottom of a dive was any depth reading $\geq$ 80% of the maximum reading observed for the dive, bottom time (B) was defined as the time between the first bottom reading and the last bottom reading (Figure S1.2). The tag estimated B and T values internally and transmitted only the summary of a classification of dive shape for each dive. Shape was classified using the following parameters:

Square	$B > 50\%\ T$
V	$B \leq 20\%\ T$
U	$20\%\ T < B \leq 50\%\ T$

Finally, tags recording dive data also recorded and transmitted low-resolution time-series data, which was used to verify summary information and visualize tag data. Depth was read every 2.5min for a period of 24 hours. The time-series data was on a duty cycle of either starting with 1 day on before switching to a duty cycle of 3 or 5 days off and 1 day on or starting with 19 days on before switching to a duty cycle to 1 day off and 1 day on. SWsat12 was the exception to this, starting with 5 days on and then switching to a duty cycle of 1 day off and 1 day on. These data were recorded at a low resolution but could be used to show more approximate fine-scale movement during dives.

S2. Sablefish Catch Data

To assess sablefish catch-per-unit-effort (CPUE) with respect to whale movement, we used data collected from observers on longline fishing vessels participating in the sablefish fishery in the Southeast (SE) statistical area of the GOA, with permission from the NOAA Fisheries' Ted Stevens Marine Research Institute (TSMRI) in Juneau, Alaska. These data do not include catch data from inside waters (e.g. Chatham Strait), where a separate sablefish fishery is managed by the Alaska Department of Fish and Game. Catch data was collected between 1995 and 2019. We calculated a CPUE index for each sablefish longline set in the database as the logarithm of catch per unit of effort (kg / 1000 hooks + 1) (Mateo and Hanselman, 2014). Additional covariates for each set included year, date, bottom depth, set location, and vessel length. For set location we calculated the mid-point of each set from the start and end locations.

To quantify the average spatial pattern of sablefish CPUE for use in the behavioral model we modeled CPUE as a smooth function of latitude and longitude using a generalized additive mixed effect model (GAMM) of the form:

$$CPUE = \alpha + a_y + f_1(lat, lon) + f_2(depth) + f_3(day) + f_4(V.length) + \varepsilon$$

where α is the intercept, a_y is a random intercept for year y to account for within-year correlations and differences in average catch rates among years, f_1 is a smooth bivariate function of latitude and longitude and f_2-f_4 are smooth univariate functions of bottom depth (*depth*), day-of-the-year (*day*) and vessel length (*V.length*). The residuals (ε) are assumed to follow a spatially autocorrelated random process with a Gaussian correlation structure to account for any remaining within-year spatial autocorrelation. The model was fit via restricted maximum likelihood assuming a Gaussian distribution and smoothing parameters were estimated using the generalized cross validation criterion (GCV) to reduce the likelihood of over-fitting (Wood, 2017). Our 'base model' assumes the spatial pattern of CPUE is consistent from year to year, which we tested by comparing the base model with two alternative models. The first model fit separate smooth surfaces of CPUE by year ($f_{1,y}(lat,lon)$) and a second alternative fit CPUE for two separate time periods before and after tagging data were available for sperm whales (i.e. before 2007 and from 2007 to present). The model with year-specific smooth functions did not converge due to large spatial gaps in the data for individual years. Therefore, we fit separate models by year and visually compared the estimated surfaces among years. All models were fit in the package "mgcv" in R (R Core Team, 2019). Model selection was carried out using residual plots and the Akaike information criterion (AIC) to compare reduced models consisting of all possible subsets of the terms f_2-f_4.

The full model was used to predict sablefish CPUE on a 0.02 decimal degree grid (approximately 1 x 1 nautical mile) spanning the observed fishing sets. The gridded area encompassed most of the recorded sperm whale foraging locations in the study area (75%). Depth at the center of each grid cell was obtained by matching its location to the bathymetric surfaces in ArcGIS. Sablefish CPUE was then predicted based on the latitude, longitude and depth of each grid point, with the values for vessel length and Day of the Year fixed at their medians (66 ft and day 137, respectively, corresponding to the typical size of most longline fishing vessels in the Southeast region and to the midpoint of fishing effort during the longline fishing season). We arbitrarily selected 2016 to visualize the spatial pattern but ran predictions over a variety of years and saw very little variation in the range, mean, and median CPUE predictions for the gridded area (Figure S2.1).

The best model indicated that sablefish CPUE peaked between 57° and 57.5° N Latitude, and -136° to -137° W Longitude. Sablefish CPUE increased as bottom depth increased and decreased as vessel length increased Figure S2). There was a seasonal variability in sablefish CPUE, with decreasing

CPUE from the beginning of the season until late April, and then increasing again through the summer months. Finally, CPUE has been decreasing by year since 1995 (Figure S2.2).

References:

Mateo, I., and Hanselman, D. 2014. A Comparison of Statistical Methods to Standardize Catch-Per-Unit-Effort of the Alaska Longline Sablefish Fishery. 1–71 pp. http://docs.lib.noaa.gov/noaa_documents/NMFS/AFSC/TM_AFSC/TM_NMFS_AFSC_269.pdf.

R Core Team. 2019. R: A language and environment for statistical comuting. R Foundation for Statistical Computing, Vienna, Australia

S3. Full Tag Movement Records

Full tag tracks for all whales show five individual tagged whales moved south of Washington state while tags were still transmitting (Figure S3.1). Whales generally stayed over the continental shelf edge while in the GOA and while migrating south (Figure S3.1), rather than moving out to the deep ocean basin. No whales migrated toward Hawaii or into the central North Pacific (Figure S3.1).

The six tag tracks from three individual whales in 2010, 2014, 2015, and 2016, that used inside waters of Chatham Strait are depicted in Figure S3.2.

S4. Dive Data Analysis

Statistical analyses with this data set would be difficult because the data are highly autocorrelated, but we looked for visual patterns in the dive data below.

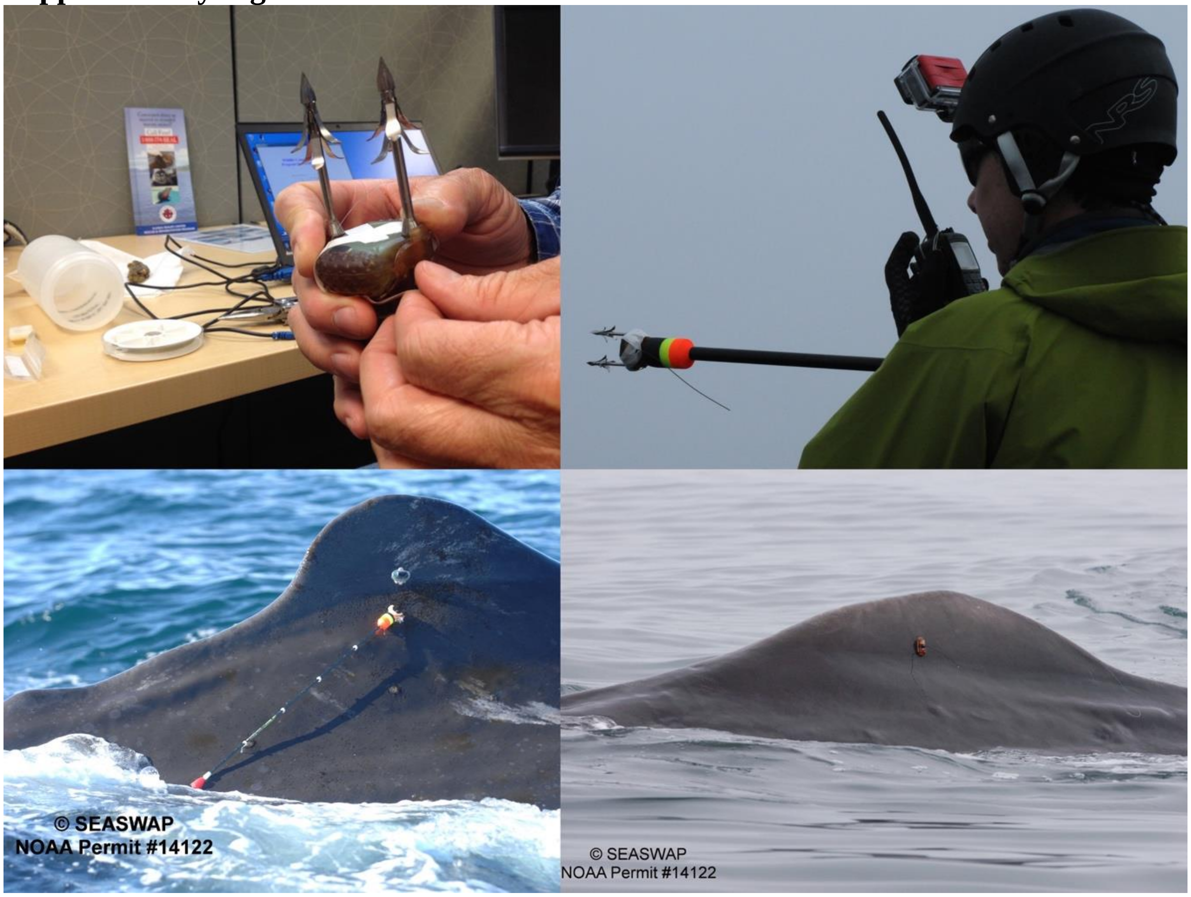
While there was some variability among individuals in the proportion of each dive type, tagged whales in this study primarily exhibited Square-shaped dives, followed by U-shaped dives, and very few V-shaped dives (Figure S4.1). The maximum dive depth, duration, and seafloor depth associated with the dive for each dive type did vary both within and among individuals but there were no striking patterns (Table S4.3). V-shaped dives were often both shallower and shorter than the other two dive types (Table S4.3).

We assessed dive shape with respect to light levels and lunar cycle to explore potential functional differences between dive types (Tables S4.1 & S4.2). The percentage of square, U, and V-shaped dives with respect to light levels was fairly similar, though daytime showed a higher percentage of square-shaped dives than other time periods, and than U or V-shaped dives (Table S4.2). New moons had the highest percentage of Square-shaped dives, though there was generally little variability in dive shape with respect to lunar cycles as well.

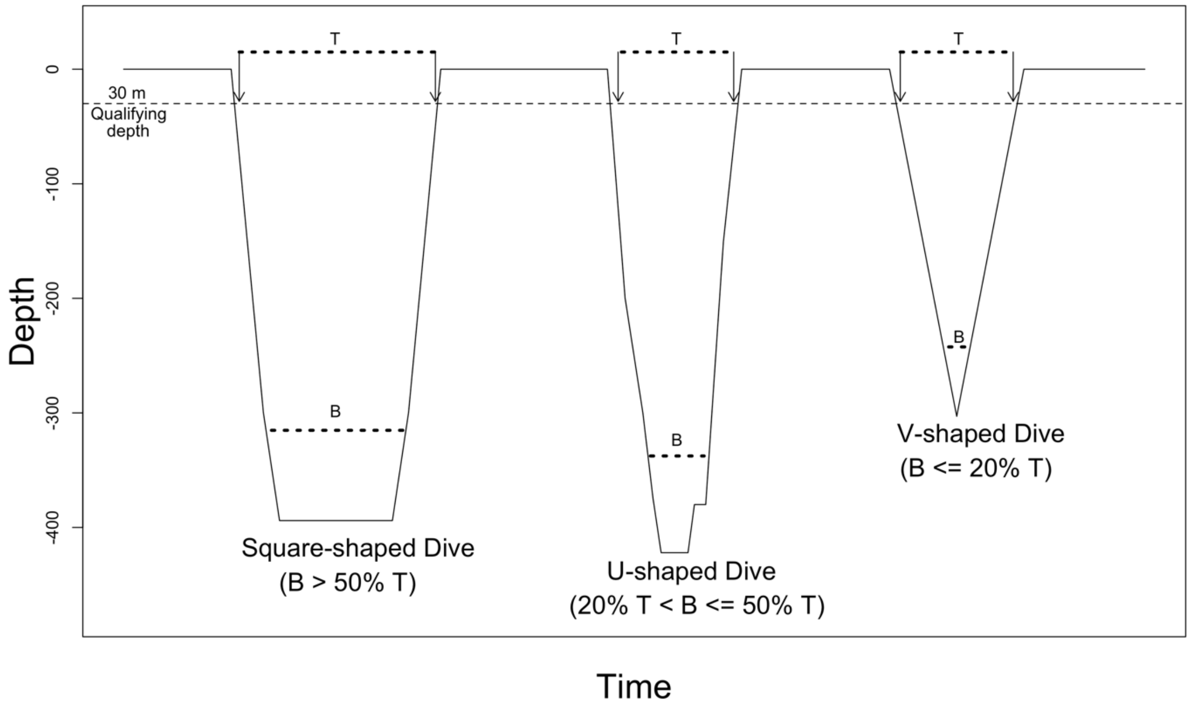
Seafloor depths associated with each dive shape did vary considerably (Table S4.3). Square-shaped dives were typically performed in shallower water, while V-shaped dives were performed in the deepest water (Table S4.3). This result is interesting because V-shaped dives also had the smallest dive depth (Table S4.3).

2 Supplementary Figures and Tables

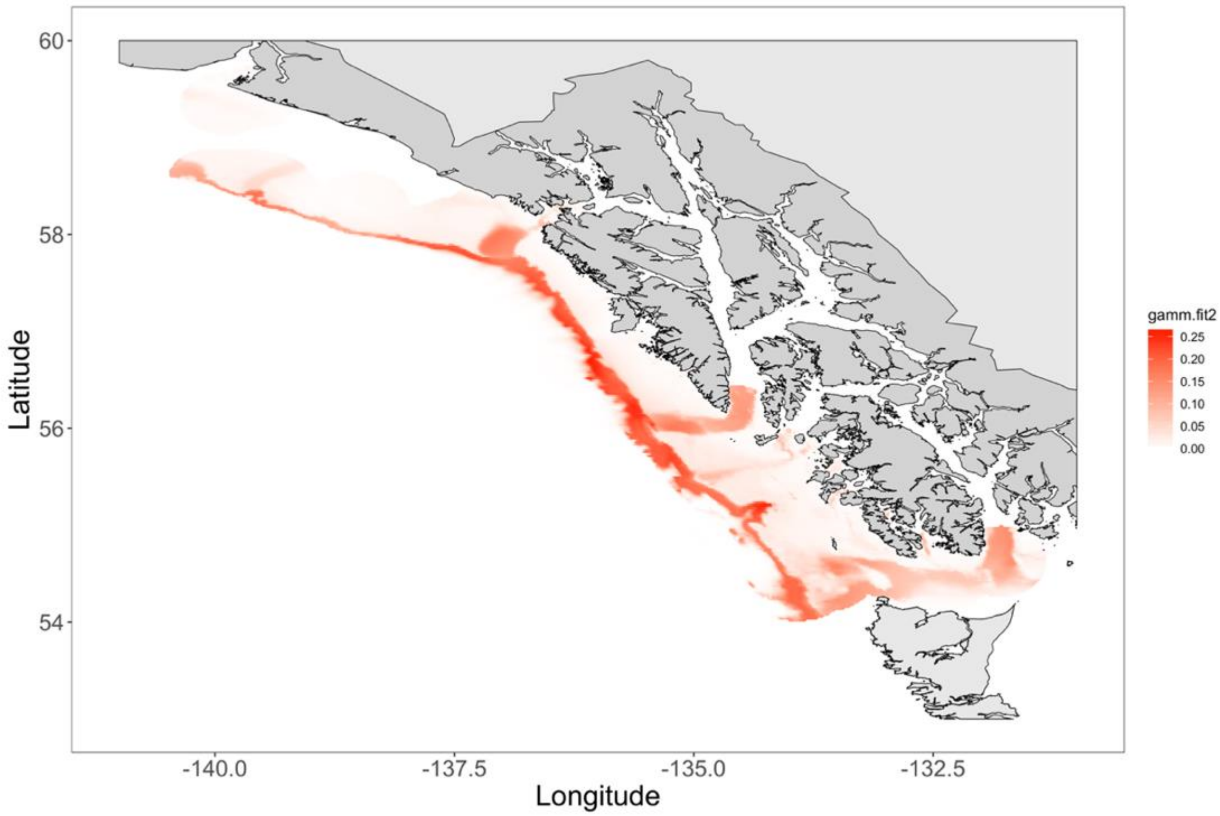
2.1 Supplementary Figures



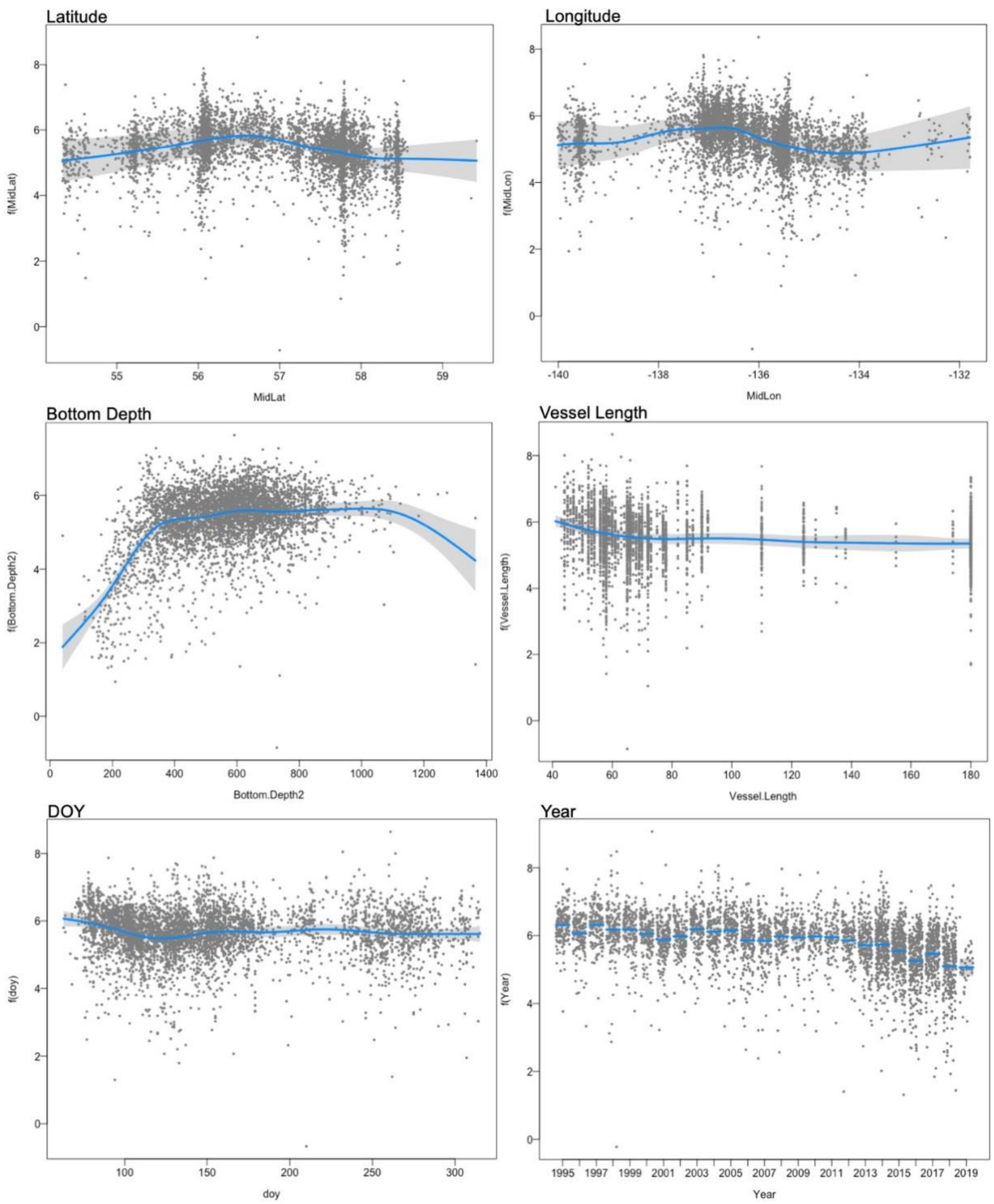
Supplementary Figure S1.1 Photos of satellite tags, clockwise from top left: Water soluble tape being applied to satellite tag in cup holder of arrow with titanium darts pointing up; tag protruding from the end of the air rifle with tape and floats from the arrow visible; tag successfully attached to the left dorsal of a sperm whale, showing size reference of the tag; tag deployment on the right dorsal of a sperm whale with arrow still visible as it detached from tag and was later retrieved from the water.



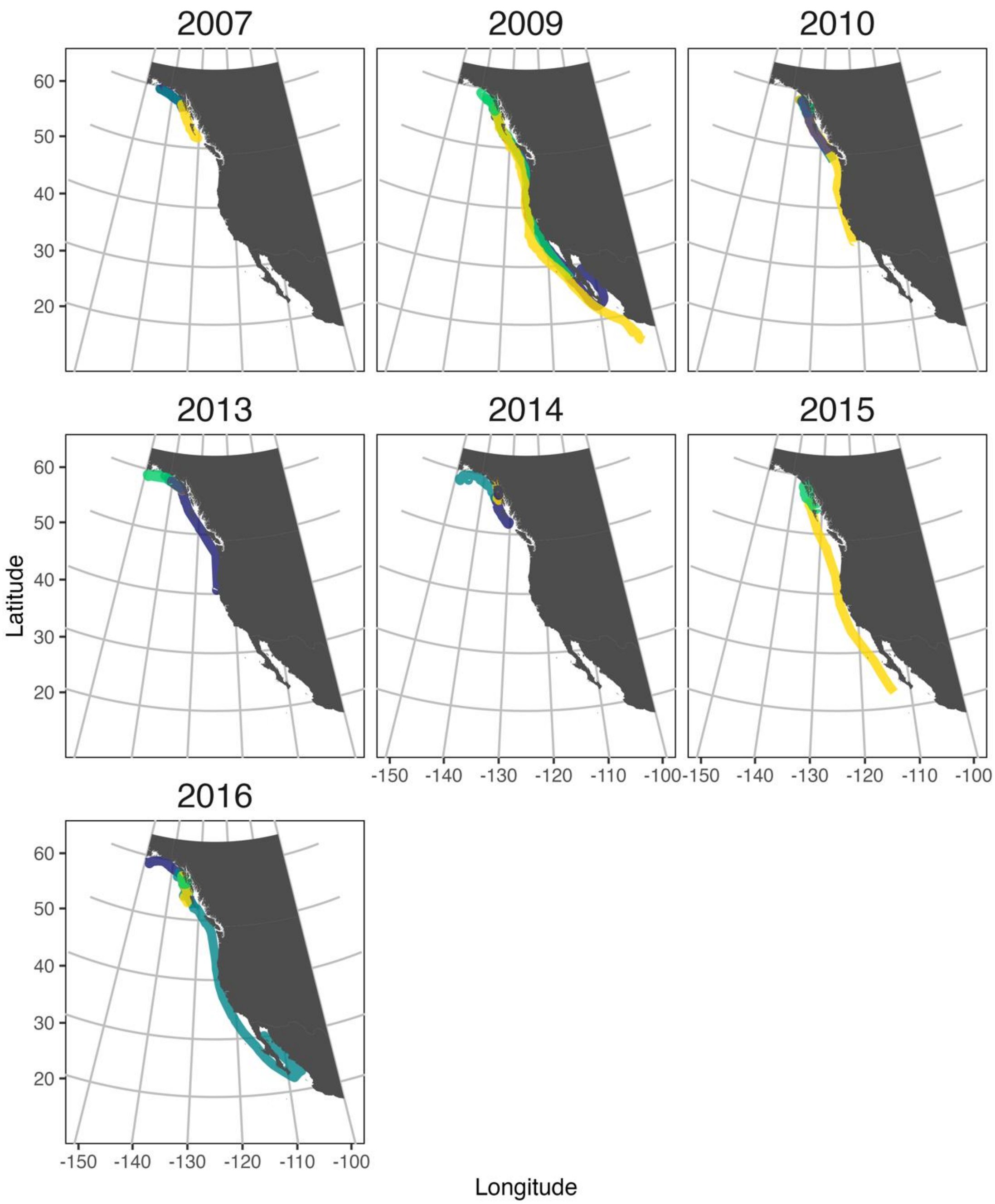
Supplementary Figure S1.2 Classification of the three dive shapes. Bottom time (B) was calculated as the amount of time the tag was below 80% of the maximum dive depth for that dive. Total duration (T) was defined as the time below dive qualifying thresholds of 30 m or 30 sec. Varying dive depths and durations in the figure reflect the averages for each shape (Table 3 in manuscript).



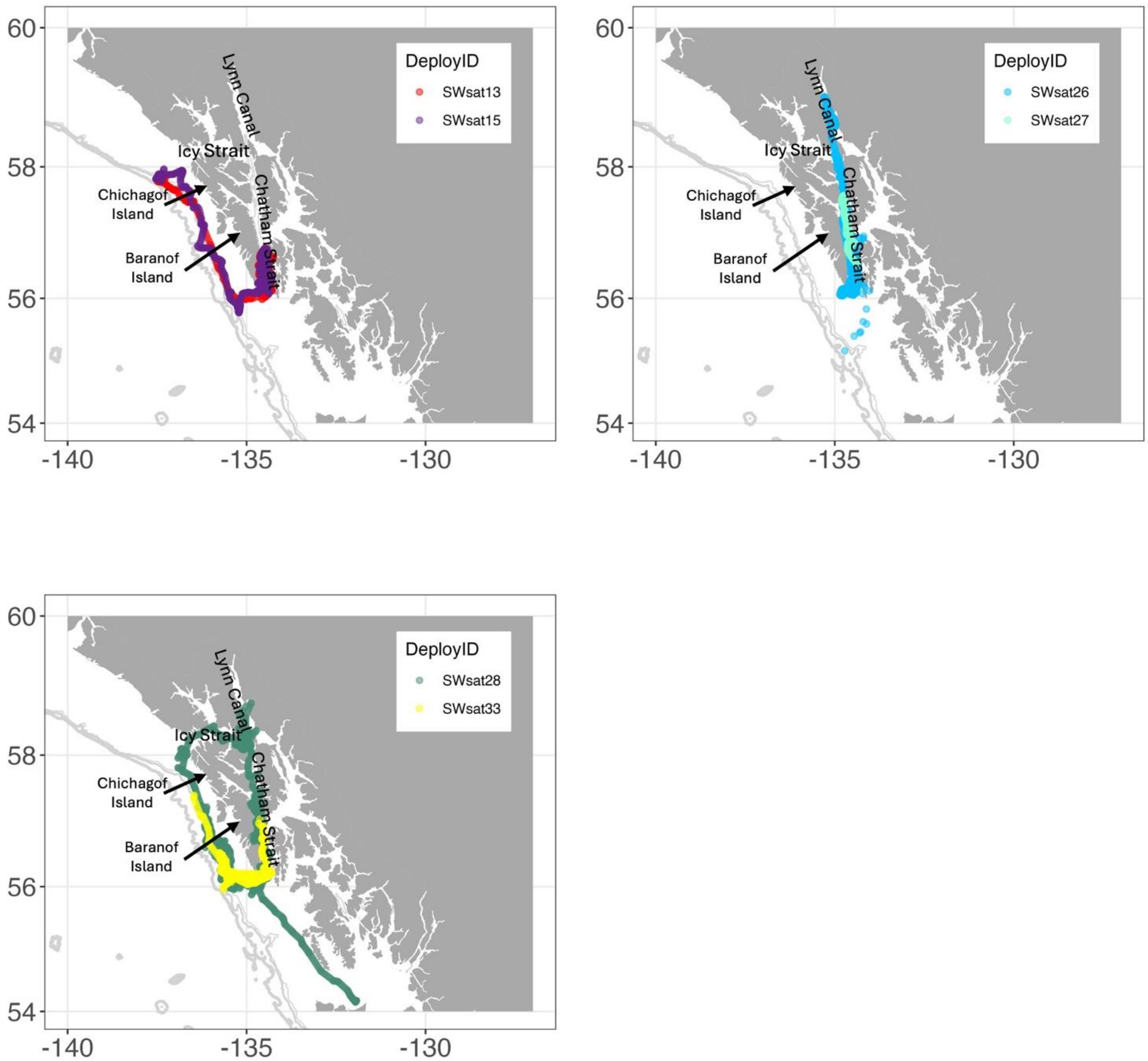
Supplementary Figure S2.1 Predicted sablefish CPUE from modeled observer catch data in the Southeast (SE) statistical area of the GOA. NOTE: gamm.fit2 is the CPUE.



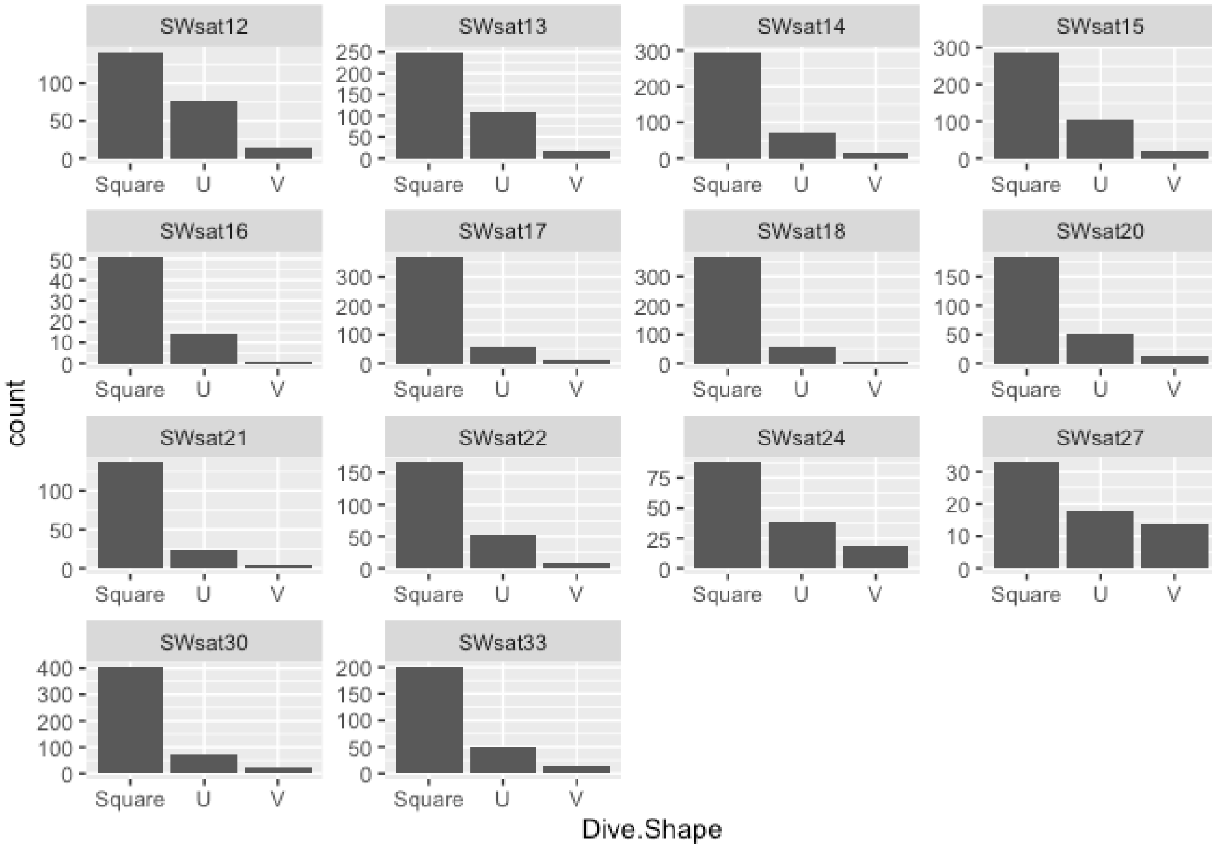
Supplementary Figure S2.2 Gamm output showing how CPUE changes with respect to location (Latitude & Longitude), seafloor depth, vessel length, day of the year (DOY), and Year.



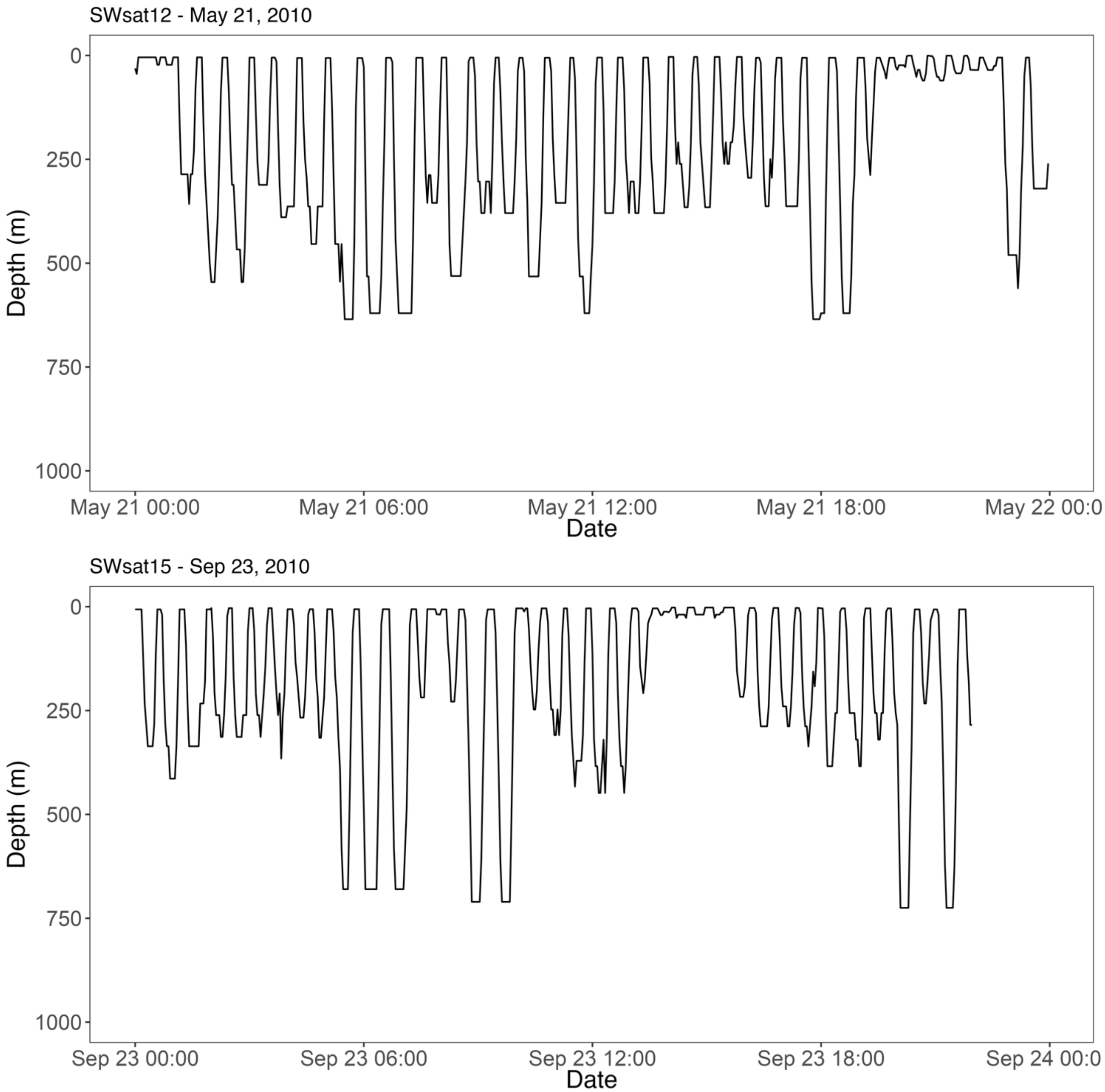
Supplementary Figure S3.1 Full tag tracks for all 29 tags analyzed for this study. Five tagged whales moved south of Washington state while tags were still transmitting; two in 2009, one in 2010, one in 2015, and one in 2016.



Supplementary Figure S3.2 Full tag tracks for all six tags that used inside waters of Chatham Strait.



Supplementary Figure S4.1 Quantity of each dive type displayed by each tagged whale that collected dive information. Y axis shows the number of dives of each shape for each tag. Note difference in the scale of the y-axes. Total number of dives for each tag was 233 (SWsat12), 368 (SWsat13), 410 (SWsat14), 413 (SWsat15), 65 (SWsat16), 366 (Swsat17), 417 (SWsat18), 242 (SWsat20), 184 (SWsat21), 222 (SWsat22), 137 (SWsat24), 63 (SWsat27), 497 (SWsat30), and 260 (SWsat33).



Supplemental Figure S4.2 Time series data (low resolution) from two tags (SWsat12 and SWsat15, both tagged in 2010) showing examples of switching between dive depths. Each plot shows a 24-hour day period.

2.2 Supplementary Tables

Table S1.1 Summary of information recorded for each tag that recorded dive behavior information. Transmit days # Start indicates the number of days the tag would initially transmit dive data after it went on the whale, after which it would switch to a duty cycle. Transmit days: duty cycle refers to the duty cycle on which the tag would transmit data after the first days of transmitting daily. Tags would sometimes collect data more frequently and only transmit occasionally to save battery life, but this only occurred once in the first year of tagging (2010) while the process was being refined.

TagID	Date Tagged	Depth Sensor Resolution (Hz)	Transmit Days: # Start	Transmit Days: Duty Cycle	# Hours to transmit initially	Which hours to transmit each day
SWsat12	5/3/2010	0.5	20	Every other day for 10 days; Then every 10th day	1	0-7; 10-23
SWsat13	8/15/2010	0.5	20	Every 3rd day for 30 days; Then every 10th day	1	0-7; 10-23
SWsat14	8/15/2010	0.5	20	Every 3rd day for 30 days; Then every 10th day	1	0-7; 10-23
SWsat15	8/15/2010	0.5	20	Every 3rd day for 30 days; Then every 10th day	1	0-7; 10-23
SWsat16	8/15/2010	0.5	20	Every 3rd day for 30 days; Then every 10th day	1	0-7; 10-23
SWsat17	5/28/2013	1	20	Every 4th day for 24 days; Then every 8th day	24	0-7; 10-22
SWsat18	5/30/2013	1	20	Every 4th day for 24 days; Then every 8th day	24	0-7; 10-22
SWsat20	5/30/2013	1	20	Every 4th day for 24 days; Then every 8th day	24	0-7; 10-22

SWsat21	5/31/2013	1	20	Every 4th day for 24 days; Then every 8th day	24	0-7; 10-22
SWsat22	6/24/2014	1	19	Every 3rd day for 12 days; Then every 6th day for 30 days; then every 12th day	24	0-7; 10-22
SWsat23	6/24/2014	1	19	Every 3rd day for 12 days; Then every 6th day for 30 days; then every 12th day	24	0-7; 10-22
SWsat24	6/25/2014	1	19	Every 3rd day for 12 days; Then every 6th day for 30 days; then every 12th day	24	0-7; 10-22
SWsat25	6/25/2014	1	19	Every 3rd day for 12 days; Then every 6th day for 30 days; then every 12th day	24	0-7; 10-22
Swsat27	9/16/2014	1	19	Every 3rd day for 18 days; Then every 6 days for 48 days; then every 12th day	48	1-7; 9-23
SWsat30	7/14/2016	1	80	Every 3rd day for 12 days; Then every 10th day	24	0-7; 10-23
SWsat33	9/10/2016	1	39	Every 3rd day for 12 days; Then every 10th day	24	5-8; 11-23

Table S4.1 Dive shape with respect to light levels. Number in parenthesis refers to total number of dives in that category.

Light	# Square (%)	# U (%)	# V (%)
Dawn	125 (71)	40 (23)	8 (5)
Day	2164 (77)	543 (19)	111 (4)
Dusk	107 (71)	35 (23)	9 (6)
Night	580 (73)	175 (22)	44 (6)

Table S4.2 Dive shape with respect to lunar cycle. Number in parenthesis refers to total number of dives in that category.

Lunar	# Square (%)	# U (%)	# V (%)
Crescent	579 (74)	163 (21)	44 (6)
Full	338 (77)	76 (17)	23 (5)
Gibbous	707 (73)	224 (23)	42 (4)
New	415 (82)	77(15)	14 (3)
Quarter	937 (76)	253 (20)	49 (4)

Table S4.3 Dive shape with respect to mean maximum dive depth, duration, and seafloor depth.

Dive Shape	Mean Max Depth (m) (±SD)	Mean Duration (min) (±SD)	Seafloor Depth (m) (±SD)
Square	394 (±135)	34 (±7)	586 (±386)
U	422 (±224)	28 (±11)	640 (±398)
V	303 (±207)	28 (±13)	714 (±476)