Development and parameterization of a data likelihood model for geolocation of a bentho-pelagic fish in the North Pacific Ocean

Julie K. Nielsen[1*] and Cindy A. Tribuzio[2]

[1] Kingfisher Marine Research, LLC, Juneau, Alaska.
[2] Auke Bay Laboratories, NOAA, NMFS, Alaska Fisheries Science Center, Juneau, Alaska.

[*]Corresponding author. Email: Julie.nielsen@gmail.com. Address: 1102 Wee Burn Dr., Juneau, AK 99801.

ABSTRACT

State-space geolocation models feature coupled process (movement) and observation (data likelihood) models to reconstruct fish movement trajectories using electronic tag data. Development of the data likelihood model is therefore a key step in adapting a state-space geolocation model for use with different fish species, geographical regions, or types of electronic data. Here we adapt a discrete hidden Markov model for the geolocation of Pacific spiny dogfish (*Squalus suckleyi*, n = 154) in the North Pacific Ocean by developing a data likelihood model based on Microwave Telemetry X-tag Pop-up Satellite Archival Tag (PSAT) data. The data likelihood model consists of light-based longitude, light-based latitude, sea surface temperature (SST), temperature-depth profile (TDP), and maximum daily depth. Pacific spiny dogfish tend to occupy coastal waters where small-scale local currents and freshwater inputs make SST and TDP variables difficult to map. To address this issue, we introduce an empirical method for parameterizing SST and TDP likelihoods by calculating root mean square difference between PSAT temperature and depth values recorded at known locations (day of tag deployment and tag pop-up) and mapped values at those locations. For SST observations (n = 85), the difference between measured and mapped values did not vary seasonally or monthly and the overall root mean square error (RMSE) used to parameterize the SST likelihood was 0.9 °C. Likelihood values for SST at known locations were higher for likelihoods parameterized with the empirical value compared to variance specification methods from previous studies. For TDP, measured values differed from mapped values (n = 89) by depth, season, and month. Therefore, RMSE values used to parameterize the TDP likelihood were calculated for each depth bin (n =

27) and month. RMSE values were low (< 1°C) for all depths during the winter but increased for depths < 100 m during the summer months. Our work provides an example of adapting state-space geolocation models for specific applications. It demonstrates the value of large numbers of tagged animals for parameterizing the data likelihood model in coastal waters as well as flexible data likelihood models with component likelihoods that can be switched on or off depending on geolocation quality.

Keywords: Fish geolocation, hidden Markov model, Pop-up Satellite Archival Tags, Pacific spiny dogfish

1. Introduction

Knowledge of fish movement patterns is a key component of fisheries management and stock assessment (Goethel et al. 2011, Lowerre-Barbieri et al. 2019). Obtaining detailed information on fish movement over large scales in space and time is challenging, but the development of electronic archival tags and analysis tools such as state-space geolocation models in recent years has greatly improved researchers' abilities to obtain insights into important behaviors such as migration and foraging for highly mobile fish species (Costa et al. 2012).

State-space models reconstruct movement paths of fish tagged with electronic archival tags by coupling a movement model, which describes the way the tagged animal is expected to move through the study area, to a data likelihood model that links the data collected by the tag to specific locations in the study area in a probabilistic way. Benefits of state-space models include the ability to accommodate messy data, accounting for measurement error in the model, allowing multiple movement states (e.g., foraging vs. migrating), and providing uncertainty in location estimates. State-space model approaches can be linear, such as a Kalman filter (Nielsen et al. 2006, Lam et al. 2008), or non-linear, such as a particle filter (Andersen et al. 2007) or a hidden Markov model (HMM, Pedersen et al. 2008).

HMMs are relatively simple discrete state-space geolocation models, and thus their use has rapidly increased in recent years. The HMM approach to geolocation was developed for Atlantic cod in the North Sea (Pedersen et al. 2008, Thygesen et al. 2009) and features a study area that is divided into discrete grid cells with an isotropic diffusion (random walk) movement model. It is ideal for non-linear applications such as nearshore study areas, as no probability is assigned to land. It allows for inclusion of multiple movement states and can be readily adapted

for use with different fish species and study areas. To adapt the HMM for specific applications, it is necessary to choose a model grid size, obtain a value of diffusion for the movement model, and develop a data likelihood model that is tailored to the behavior of the fish, the type of electronic tag, and the available geolocation data maps for the study area.

We adapted the HMM for the geolocation of Pacific spiny dogfish (*Squalus suckleyi*, hereafter termed "dogfish") in the North Pacific Ocean. The dogfish is one of the most common shark species in the coastal waters of the North Pacific Ocean and is distributed throughout the nearshore waters of the U.S., Canada, Russia, and Japan (Figure 1). This species is often bycaught in major fisheries such as those for Pacific halibut and walleye pollock (Tribuzio et al. 2020), and has been subjected off and on to directed fishing (King et al. 2017). Dogfish stocks in Alaskan waters are assessed, and the catch of the species is managed, as part of a complex of all shark species within the Gulf of Alaska (GOA) and Bering Sea-Aleutian Islands (BSAI) fishery management plan areas. As a whole, the shark stock assessments are data-limited, but dogfish are the most data-rich within the complex. Life history data and fishery-independent survey indices inform the assessment. However, information on habitat associations and seasonal and annual movement within and between management jurisdictions is limited.

Previous research on dogfish movement using conventional tags has provided some evidence of seasonal movement between U.S. and Canada (Taylor 2008) and some large-scale movements across the northeast Pacific to Japan and Russia (McFarlane and King 2003, Taylor 2008). However, detailed information on movement patterns is difficult to obtain using only information about release and recovery positions. To learn more about dogfish movement patterns relative to fisheries management areas in the North Pacific Ocean, we initiated a Pop-up Satellite Archival Tag (PSAT) study in 2009 and used the HMM to reconstruct movement paths of tagged dogfish based on PSAT light, depth, and temperature data.

We customized the HMM for our application by developing and parameterizing a data likelihood model that accounts for dogfish behavior, PSAT data type, and the environmental characteristics of our study area. The data likelihood model for Pacific spiny dogfish is based on light-based latitude and longitude, sea surface temperature (SST), temperature-depth profiles (TDP), and maximum daily depth. Although the parameterization of the light-based latitude, light-based longitude, and maximum depth likelihoods is straight-forward, parameterizing the SST and TDP likelihoods is more challenging because dogfish tend to occupy near-shore

103 waters where the influence of small-scale coastal currents and freshwater run-off can reduce
104 accuracy of the mapped values in the study area. We addressed this challenge by leveraging
105 the large number of tagged dogfish in this study to develop a new method for parameterizing
106 data likelihoods for SST and TDP based on empirical differences between PSAT data and
107 mapped data at known (release and pop-up) locations. Our work provides an example of
108 customizing a data likelihood model for a specific application and adapting the model for
109 situations where errors in study area mapped data may be high. In addition, this work can be
110 viewed as a continuation of efforts to explore the sensitivity of state-space geolocation models
111 to choices about fixed parameters and likelihood methods which are needed to ensure
112 robustness and confidence in geolocation results.
113
114 **2. Materials and methods**
115 ***2.1. PSAT data***
116 During 2009 – 2013, 173 Microwave Telemetry (Columbia, Maryland) X-tags were deployed on
117 dogfish in the GOA (Figure 1). Tags were deployed during directed research cruises and
118 opportunistically on the Alaska Fisheries Science Center's annual groundfish longline survey
119 (e.g., Malecha et al. 2019). The PSATs were attached with a method adapted from Carlson et al.
120 (2014) where a hole was drilled through the anterior dorsal fin spine below but near the point
121 where the spine extrudes from the skin. A piece of 300lb test monofilament line was looped
122 through the hole, pulled tight, and clamped. The tag was attached to the monofilament line with
123 a loose loop which allowed the tag to swing freely. The monofilament line was covered with
124 silicone tubing (1/8" inner diameter, 1/4" outer diameter) to prevent it from snagging and keep
125 the clamps from irritating the skin of the dogfish.
126

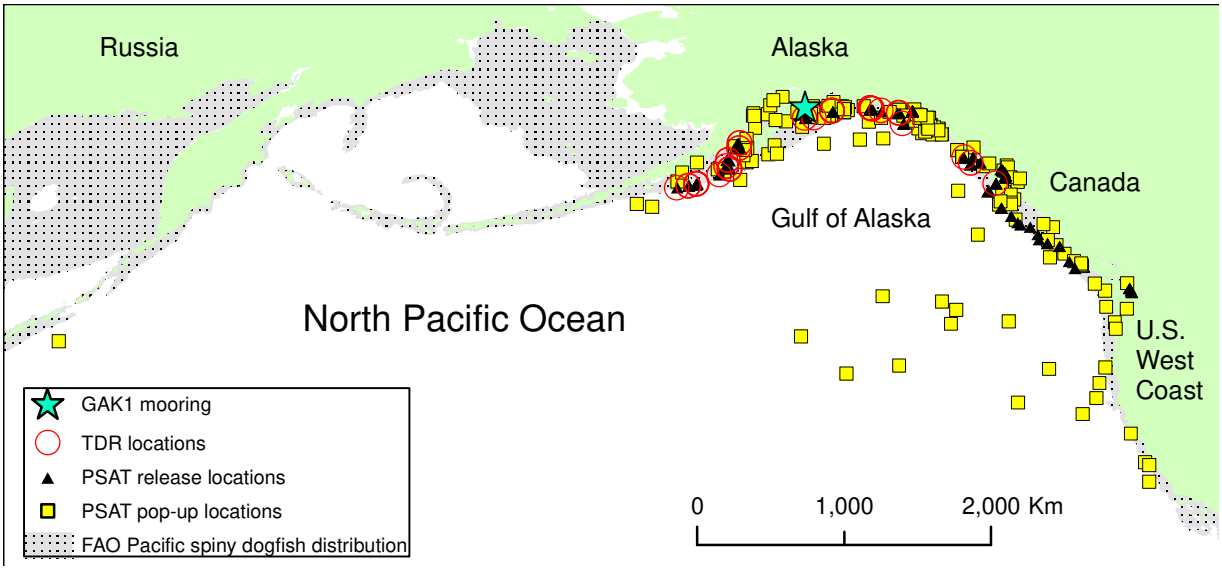


Figure 1. Pacific spiny dogfish study area in the North Pacific Ocean. The Food and Agriculture Organization of the United Nations (FAO) species distribution map for Pacific Spiny Dogfish (pink area) depicts a mostly nearshore distribution.

The PSATs weigh 46 g in air with a diameter of 3.3 cm and a length of 12.2 cm. The tags recorded depth, temperature, and light and were programmed to release ("pop up") from the fish after 6 – 12 months. After the tags popped up, they transmitted daily light-based latitude and longitude estimates and time series of depth and temperature records. If tags released from the fish prior to the pop-up date, data transmission was triggered by a constant depth when the tag floats at the surface. Measurement intervals for the depth and temperature time series varied depending on the length of tag deployment and ranged from 15 minutes to one hour. Information on measurement resolution also accompanied each depth (0.34 m – 5.4 m) and temperature (0.16 °C – 0.23 °C) observation. Physically recovered PSATs provided depth and temperature records every 2 minutes.

Latitude and longitude estimates were derived from light levels at dusk and dawn by the tag manufacturer using a proprietary algorithm. The estimates were available as both raw (daily) and smoothed (multiple day average) locations. The location estimates were accompanied by the depth at which dawn and dusk values were obtained. Because light-based latitude and longitude estimates can produce estimated locations that are far beyond the possible range of movements of the fish in a given time period, we used only raw values and filtered the position

150 estimates manually to remove obviously spurious values (daily difference greater than 2
151 degrees longitude and 4 degrees latitude) prior to running the geolocation model. The filter is
152 larger for latitude compared to longitude because latitude estimates are much less precise.
153 Filtering of extreme values was necessary because spurious estimates that fall outside the
154 study area would negate relevant information from other data sources, such as depth and
155 temperature, at that time step (see "overall likelihood calculation" section below). Latitude and
156 longitude estimates were treated separately for pre-processing and in the model, as longitude is
157 more robust than latitude during equinox and when measurements are obtained from deeper
158 waters (Seitz et al. 2006).
159
160 Temperature and depth records were processed to obtain a data set of concurrent depth and
161 temperature values. Microwave Telemetry tags use an algorithm for the compressed data (i.e.,
162 those data that are transmitted) in which any changes in depth or temperature that are too great
163 to be measured accurately are flagged as "delta limited". For our analyses, we discarded these
164 values, which typically comprised < 2% of the transmitted data points. For periods of time where
165 depth and temperature measurements were offset slightly (e.g., 15 min), depth and temperature
166 records were linearly interpolated using the command "na.approx" from R package "zoo" (Zeileis
167 and Grothendieck 2005) with a maximum record gap of 1 time interval.
168
169 ***2.2. Geolocation model***
170 To reconstruct the movement paths of dogfish in the North Pacific Ocean, we adapted a HMM
171 developed for the geolocation of Atlantic cod in the North Sea (Pedersen et al. 2008, Thygesen
172 et al. 2009). The HMM is a Bayesian state-space geolocation model based on the division of the
173 study area into discrete grid cells. Each grid cell ultimately contains the probability that the
174 tagged fish occupied the grid cell at each time step. First, beginning at the tag release location,
175 a forward filter is implemented that alternately applies an update from the movement model
176 followed by an update from the data likelihood model at each time step (Figure 2). The
177 movement model is a random walk, represented in the model as a two-dimensional diffusion
178 kernel that is convolved with the prior. After the movement model update, the prior is then
179 multiplied elementwise by the data likelihood model values at that time step to obtain the joint
180 probability of the observed tag data and grid cell value. The sum of the joint probability density
181 is referred to as lambda, and this quantity is used to assess the performance of the model and
182 to estimate diffusion. The joint probability density is normalized by lambda to become the
183 posterior and then becomes the prior for the next time step. Once the recovery location is

reached, backward smoothing is performed to update the probabilities with knowledge of the recapture location. In addition to the initial description of the model provided by Pedersen et al. (2008), additional details for all of these steps are available in Thygesen et al. (2009), Pedersen et al. (2011), Le Bris et al. (2013), Woillez et al. (2016), Braun et al. (2018a), and Nielsen et al. (2019).

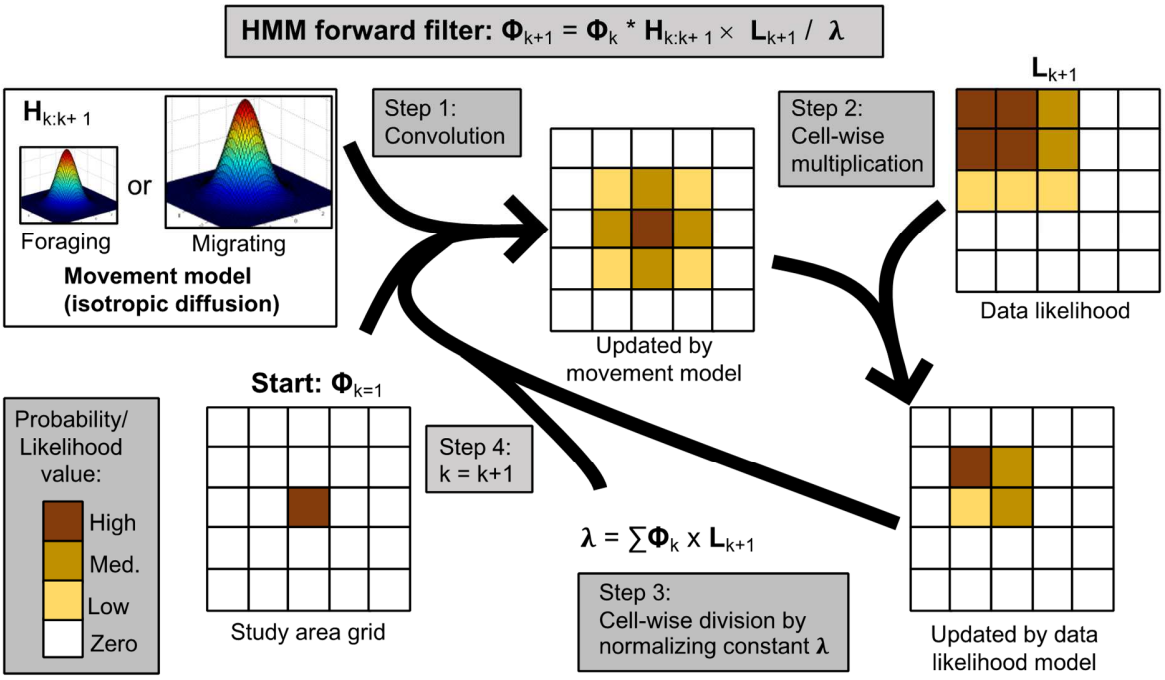


Figure 2. A conceptual model of the forward filter of a Hidden Markov Model (HMM) for fish geolocation (Figure 1 from Nielsen et al., 2019). The matrix $\mathbf{\Phi}_k$ holds the estimated probability distribution in each study area grid cell at time *k*. The model is initiated at time $k = 1$ with all of the probability in the grid cell of the release location. The matrix **H** contains the transition probabilities from time *k* to time *k*+1 derived from the diffusion coefficient (movement model) and can change depending on the movement state (e.g., foraging or migrating, assigned prior to running the model) at time step *k*. The initial probability is updated first by the movement model, through convolution (*) with **H**, and then by the data likelihood model through cell-wise multiplication (x) with the matrix $\mathbf{L}_{k+1}$ that contains the data likelihood at time *k*+1 (see Figure 3 for illustration of individual data likelihood model components for the geolocation of spiny dogfish). Updated probabilities are then normalized by cell-wise division (/) with the normalization constant λ and the recursion continues with alternating movement model and data likelihood updates until the last observation is reached.

205 Development of the data likelihood model is one of the most important steps in the process of
206 adapting the HMM for a new application (e.g., species or geographic region). Data likelihood
207 models specify the way that data collected by the PSAT are linked to known spatial distributions
208 (i.e., maps) of geolocation variables in the gridded study area. The data likelihood models take
209 into account fish behavior (e.g., preference of demersal vs. pelagic habitat), the type and quality
210 of tag data, and environmental gradients of geolocation variables in the study area.
211
212 In order to link data recorded by the PSAT to grid cells in the study area at each time step,
213 specification of the distribution of geolocation variables, usually assumed to be Gaussian, within
214 each grid cell is needed. Deciding how to parameterize grid cell variance is a critical component
215 of the data likelihood model that can affect model performance (Nielsen et al. 2019) and
216 researchers have parameterized variance in different ways. One common approach is to assign
217 variance values in each cell by calculating the standard deviation of adjacent grid cells (Le Bris
218 et al. 2013, Liu et al. 2017, Braun et al. 2018a). When the resolution of the mapped geolocation
219 variable is higher than the model grid cell resolution, it is also possible to assign variance by
220 calculating the standard deviation of small-scale grid cells aggregated to form the larger model
221 grid cells (Nielsen et al. 2019). Other researchers use a constant value for all grid cells in the
222 study area that is based on auxiliary data, prior research, uncertainty values that accompany the
223 data set, or the value that provides the best model performance (Pedersen et al. 2011, Biais et
224 al. 2017). Alternatively, variance may be estimated by the model (Woillez et al. 2016).
225
226 In addition to development of a data likelihood model, adapting the HMM for different
227 applications (e.g., species or geographic regions) involves 1) deciding whether to estimate
228 diffusion in the model or use a pre-determined value, and 2) choosing the optimal grid size.
229 Movement states (e.g., foraging vs. migrating) can be specified prior to model estimation based
230 on auxiliary analyses of the data set or information from previous research, for which different
231 values of diffusion may be used. In this manuscript, however, we focus on the development and
232 parameterization of the data likelihood as the key step in adapting the HMM for our specific
233 application.
234
235 ***2.3. Data likelihood model for dogfish***
236 Dogfish can be found throughout the water column. When they occupy shallow waters, PSATs
237 collect information on light intensity, which provides information on latitude (e.g., day length) and
238 longitude (e.g., time of local noon). Light-based geolocation is the primary means of geolocation

for pelagic fish (Musyl et al. 2001, Schaefer and Fuller 2016). When dogfish spend time near the sea surface, the temperatures recorded by the PSAT can be matched to satellite imagery of sea surface temperature (SST) in the study area (Nielsen et al. 2006, Lam et al. 2008). When dogfish occupy deeper waters, temperature-depth profiles (TDPs) recorded by the PSATs can be matched to mapped TDPs in the study area (Skomal et al. 2009, Braun et al. 2018b). Because dogfish can be anywhere in the water column (in contrast to demersal fishes which are assumed to be on or near the sea floor at least once a time step), the maximum depth recorded by the tag each day is used to rule out geographic areas with shallower depths. Therefore, the data likelihood model for spiny dogfish is composed of light-based longitude and latitude, SST, TDPs, and maximum depth (Figure 3).

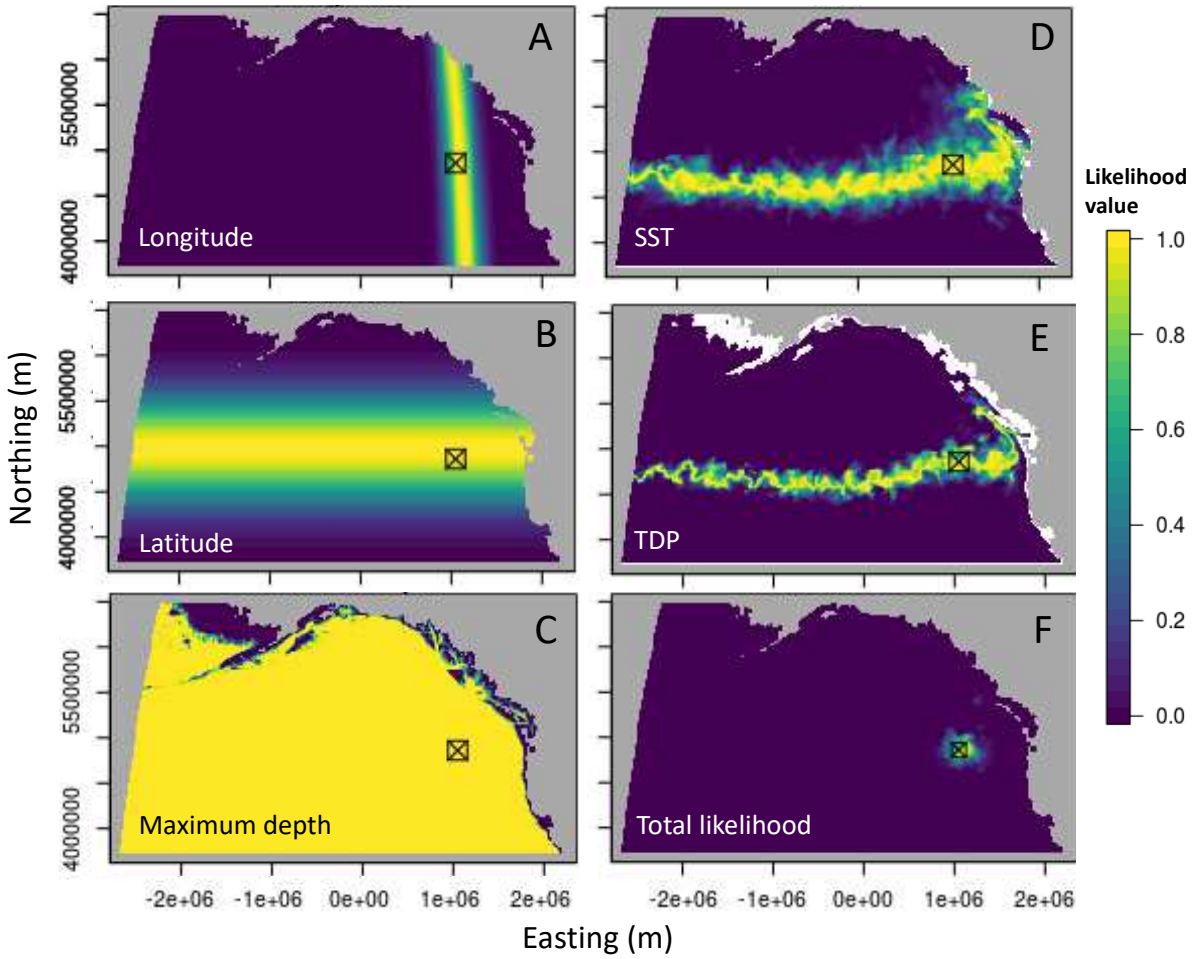


Figure 3. Example of the data likelihood model for a dogfish for the day before the PSAT pop-up date. The pop-up location is indicated by a crossed square symbol in all plots. The data likelihood consists of A) light-based longitude, B) light-based latitude, C) maximum daily depth,

D) sea surface temperature (SST), and E) temperature-depth profile (TDP) components. All 5 components are combined to produce the total likelihood at that time step (F).

2.3.1. Light-based longitude and latitude

We treat longitude (Figure 3A) and latitude (Figure 3B) as separate likelihoods to allow for inclusion of only longitude values when latitude values are spurious (e.g., when the fish occupies deeper waters or during equinox periods). The likelihood value for light-based longitude (or latitude) is the probability of observing the longitude (or latitude) obtained from the PSAT given normal probability density function (PDF) centered on the study area grid cell:

$$L_{light} = N(x; \mu, \sigma), \quad (1)$$

where x is the longitude (or latitude) estimated by the PSAT, μ is the longitude (or latitude) of the grid cell, and σ is the standard deviation of longitude (or latitude) obtained from previous geolocation studies. We specified a standard deviation of 1.5 degrees for longitude and 3.5 degrees for latitude based on values used for other temperate shark species (Biais et al. 2017, Doherty et al. 2017).

2.3.2. Maximum depth

A bathymetric map is used to calculate the maximum depth likelihood (Figure 3C) as well as to assign a probability of zero to land during the estimation process. We use the SRTM30+ Global 1-km Digital Elevation Model (DEM): Version 11 bathymetry data set which provides bathymetry information on a 0.008 degree grid.

The maximum depth likelihood is obtained from a normal cumulative distribution function (CDF) of the mean depth and estimated depth variance within each grid cell (Pedersen et al. 2008). The likelihood value is the CDF quantile represented by the tag depth, normalized by a CDF truncated at a depth of zero and modified to accommodate positive depth values:

$$L_{max\_depth} = \left(1 - \Phi\left[\frac{x-\mu}{\sigma}\right]\right)\left(1 - \Phi\left[\frac{-\mu}{\sigma}\right]\right)^{-1}, \quad (2)$$

where Φ is a Gaussian cumulative distribution function, x is the maximum observed depth during the time step interval from the PSAT, μ is the mean depth value for the grid cell (always

288 positive), and σ is the standard deviation of the bathymetry in the grid cell. The standard
289 deviation for each grid cell was derived from all depths used to aggregate the fine-scale
290 resolution bathymetry map (1-km resolution) to the 20-km resolution model grid (Nielsen et al.
291 2019).
292
293 2.3.3. SST
294 To calculate the SST likelihood, we use the Multi-scale Ultra-high Resolution Sea Surface
295 Temperature (MUR SST), which provides daily SST and error estimates on a 0.01 degree grid
296 (JPL MUR MEaSUREs Project 2010). To link tag temperature at the surface to the SST map,
297 we first define how the X-tag measured SST. For X-tag data, SST must be defined from the time
298 series depth and temperature records. The MUR SST data set provides information about the
299 "foundation" SST, which reflects the base temperature of the top water layer at night or when
300 strong winds mix the surface waters. Under those conditions, differences between SST and
301 deeper waters (e.g., up to the first 10 m) are small relative to tag measurement resolution
302 (Donlon et al. 2007, Kawai and Wada 2007). We conducted a preliminary analysis to determine
303 whether the mean, median, or maximum temperature obtained at depths less than 10 m was
304 closest to known SST values at known locations. Because the maximum temperature performed
305 the best, we define SST measured by the PSAT as the maximum temperature of all
306 measurements obtained when the depth plus tag depth measurement resolution (maximum 5.4
307 m) was less than 11 m each day. This effectively limits the temperature records that can be
308 matched to SST to the two shallowest depth bins (0 – 5.39 m and 5.4 – 10.7 m) at the highest
309 levels of depth uncertainty. This is similar to the definition of SST used by Woillez et al. (2016)
310 for geolocation of a pelagic fish.
311
312 The SST likelihood value in each grid cell (Figure 3D) is obtained by integrating a normal PDF
313 of SST values in the grid cell between the upper and lower values of tag measurement $\pm$ tag
314 measurement uncertainty (Le Bris et al. 2013):
315
316 $$L_{SST} = \int_{T_1}^{T_2} \mathrm{N}(x; \mu, \sigma)\, dx, \quad (3)$$
317
318 where x is the maximum temperature measured by the fish at depths shallower than 11m
319 (including depth measurement uncertainty) during the time interval, T_1 and T_2 are the lower and
320 upper limits of uncertainty in tag temperature measurement, μ is the value of SST in each grid

cell from the MUR SST data set, and σ represents the standard deviation of SST values within the grid cell. Sigma can be derived using different methods (Nielsen et al. 2020). Here, we introduce a new method for determining σ by calculating an empirical value based on the difference between SST measured by the tag and SST provided by the MUR SST data set at known locations (see "Calculation of empirical variance" section below).

2.3.4. TDP

To calculate TDP likelihoods, we use estimated temperatures from the HYCOM global oceanographic model (Wallcraft et al. 2009), which provides daily temperatures for 40 depth bins on a 0.08 degree grid. To link temperature values measured by the PSAT at different depths to the HYCOM map, we calculated the average PSAT temperature for each depth bin and compared those values to the corresponding depth bins for the HYCOM model. For the likelihood calculations, we linked PSAT data to a subset of 27 of the 40 depth bins provided by the HYCOM model to account for depth measurement uncertainty of up to 5.4 m in the PSAT data. For example, for the first 30 m the HYCOM model provides estimates at 0, 2, 4, 6, 8, 10, 12, 15, 20, 25, and 30 m but the depth bins used for the first 30 m of the likelihood were 0, 6, 10, 20, and 30 m.

The TDP likelihood is calculated for each depth bin separately and then combined to obtain the likelihood for all depths at each time step (Figure 3E). We do not reconstruct temperature-depth profile curves as is done with other TDP likelihoods (Braun et al. 2018a, Braun et al. 2018b) because dogfish move rapidly from shallow to deep waters and measurements at intermediate depths were frequently "delta-limited" by the X-tags and thus discarded. Likelihood values for each depth bin are calculated in the same manner as SST likelihood values (eq. 2), where a normal PDF with a mean of the HYCOM temperature in the grid cell and an empirical value for standard deviation (see "Calculation of empirical variance" section below) are integrated between the upper and lower values of the mean tag temperature in the depth bin ± tag measurement uncertainty (Le Bris et al. 2013). After likelihoods for all depth bins are calculated, likelihoods for all depth bins in each grid cell are multiplied (Braun et al. 2018a) to obtain the overall likelihood in each grid cell for that time step:

$$L_{TDP} = L_{TDP\_depth\_1} * L_{TDP\_depth\_2} * \ldots * L_{TDP\_depth\_n} , \quad (4)$$

where depth_1 is the likelihood for first depth bin, depth_2 is the likelihood for the second depth bin, and n is the number of depth bins for which temperature was recorded by the PSAT.

2.3.5. Overall likelihood calculation

All individual likelihoods (longitude, latitude, SST, TDP, and maximum depth) are created using the same grid size. Likelihood data available as geographic coordinates were first aggregated to 0.2 degrees, then projected to meters (Robinson with a meridian at -145º). The 20 × 20 km grid size used in this study was chosen based on the size of the study area, the relatively coarse spatial scale of variation in likelihood components, and the movement speed of spiny dogfish. It is a typical grid size for pelagic fish geolocation studies (Pedersen et al. 2011, Braun et al. 2018b). Then for each time step (one day), all likelihoods are combined by cell-wise multiplication to obtain the overall likelihood for that time step (Figure 3F);

$$L_{Total} = L_{Longitude} * L_{Latitude} * L_{SST} * L_{TDP} * L_{max\_depth} . \quad (5)$$

If a variable (e.g., SST) is not available for that time step, all grid cells in the study area are given a value of 1 for that likelihood.

2.4. Calculation of empirical variance

We parameterized the SST and TDP likelihoods by comparing observed PSAT depth and temperature data at known locations to values from SST and TDP maps at those locations. For each release and pop-up location, we examined all available PSAT records and recorded SST and TDP values up to 3 days after release and 3 days prior to pop-up. Only pop-up locations from PSATs that released on schedule were used for this analysis because tags that detached from the fish prior to the scheduled pop-up date drifted on the water surface for several days before transmission to the Argos satellite network was initiated. We calculated the difference between values measured by the PSAT and mapped values using the values closest to the day of release or pop-up in all calculations. Assuming the location observations would be more accurate on Day 0 compared to Day 3, we assigned ad hoc weights to each PSAT observation based on our confidence that the recorded tag values corresponded to the location of the tagged fish at that time. Observations on days 0, 1, 2, and 3 after release or before pop-up were assigned weights of 0.5, 0.3, 0.15, and 0.05, respectively. Observations from fish released in the same location on the same day were averaged to provide a data set with unique combinations of location and day.

For TDP comparisons, we supplemented tag data with two auxiliary sources of temperature-depth profile information. First, Sea-Bird Scientific (Bellevue, WA) SBE 39 temperature-depth recorders (TDRs) were deployed on the AFSC longline survey near release locations for fish tagged on that survey (Figure 1). These recorders provided TDP information for the full range of depths near release locations (K. Siwicke, NOAA AFSC, unpublished data). Second, we obtained detailed (30 minute) temperature measurements at depths of 20, 30, 60, 100, 150, 200, and 250 m with Sea-Bird Scientific conductivity temperature and depth (CTD) sensors from the GAK1 mooring (University of Alaska College of Fisheries and Ocean Sciences, data available at http://research.cfos.uaf.edu/gak1/) located at 59.845 N, 149.4667 W (Figure 1). The mean temperature for each depth and day was obtained from the GAK1 records, and monthly averages of daily differences between GAK1 measurements and HYCOM values were added to the TDP data set.

We checked for difference between measured and mapped SST and TDP values by season (winter = December, January, and February, spring = March, April, and May, summer = June, July, and August, autumn = September, October, and November) and month using Kruskal-Wallis test for non-parametric analysis of variance (Zar 1999). For the TDP value, we combined depth observations into three larger depth bins for statistical analyses (0 – 50 m, 51 – 100 m, and > 100 m). We plotted difference between measured and mapped values by distance from shore. We then calculated the root mean square error (RMSE) between mapped SST and TDP values and observed tag measurements, weighting each observation by the number of days from the known location, for use in the SST and TDP likelihoods:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Obs_i - Map_i)^2 * w_i}{\sum_{i=1}^{n} w_i}} \quad , \qquad (6)$$

where *Obs* is the observed temperature from PSAT, TDR, or GAK1 measurements, *Map* is the mapped temperature value from MUR SST or HYCOM at the known location, and *w* is the weight corresponding to the number of days between observation and time of release or pop-up. If no temporal variation was observed, one value of RMSE was calculated per depth for use in the model. If temporal variation was observed, RMSE was calculated for each time period and depth, and RMSE values were interpolated with the interp.loess command from the R package

420 “tgp” (Gramacy 2007). To compare the empirical SST RMSE value to offshore values, RMSE
421 was also calculated for observations > 50 km offshore.
422
423 **3. Results**
424 ***3.1. PSAT data***
425 Of the 173 PSATs deployed, 79 PSATs popped up on the scheduled date and thus provided
426 precise pop-up locations. Another 73 PSATs detached from the fish prior to the pop-up date.
427 Eight tags were physically recovered and provided detailed data. Most pop-up locations were
428 scattered along the Pacific coast from California to the Aleutian Islands, ten were 700 - 2000 km
429 offshore in the North Pacific Ocean, and one pop-up location was in Russian waters (Figure 1).
430 An example of a dogfish data set (Figure 4) features periods of rapid change between near-
431 surface waters and depths to more than 400 m, periods of time spent exclusively at shallow
432 depths, and periods of time spent at intermediate depths with few visits to surface waters.
433

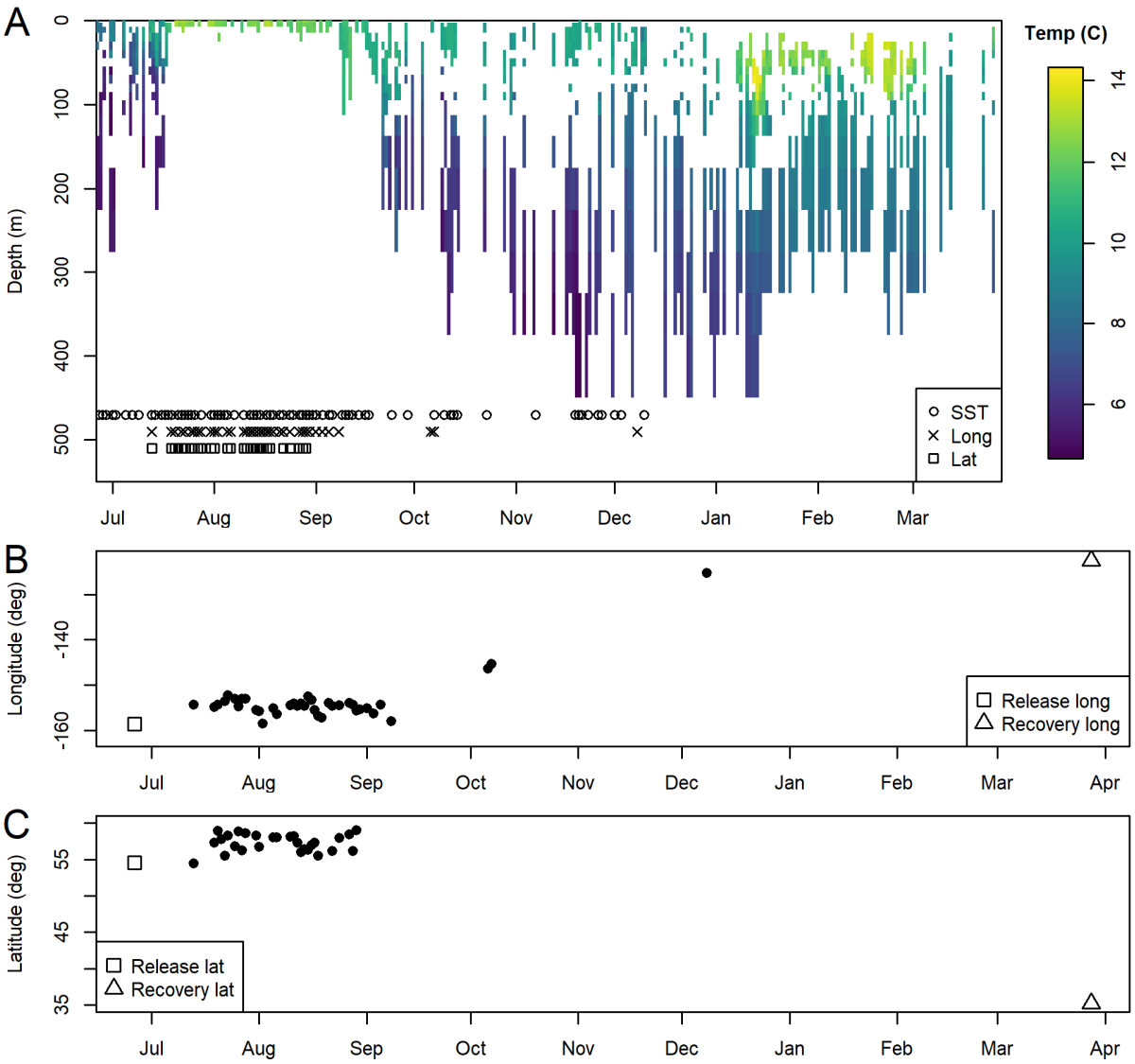
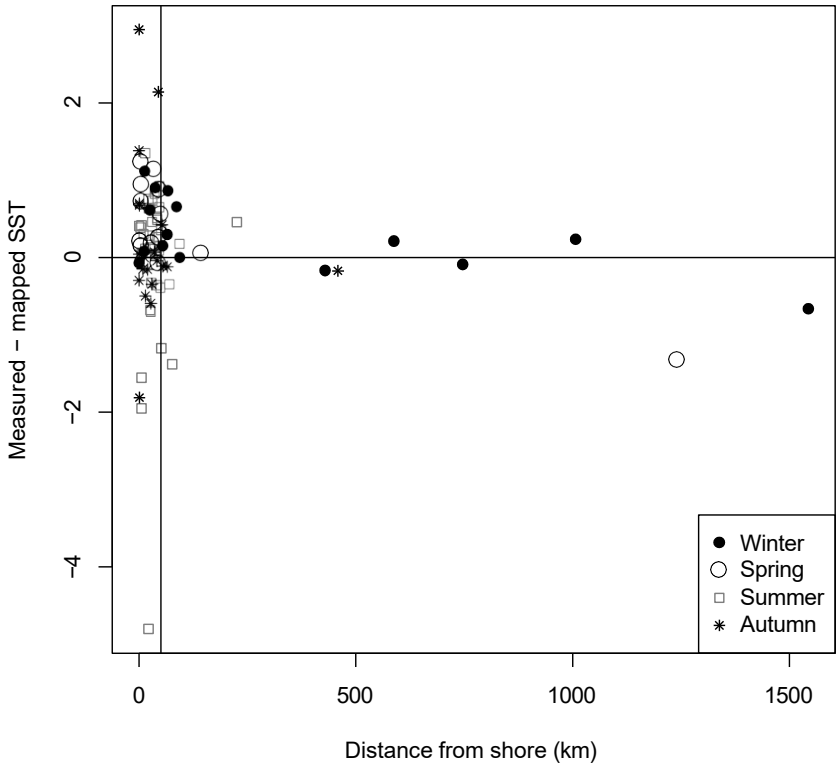


Figure 4. Example of geolocation data available for spiny dogfish. Temperature-depth profile (A) with the HYCOM depth bins used in the model. Circles on the lower portion of the plot indicate days when SST observations are available; x's and squares indicate days with longitude and latitude observations, respectively. Processed (filtered) longitude (B) and latitude (C) observations for use in the model.

3.2. Calculation of empirical variance

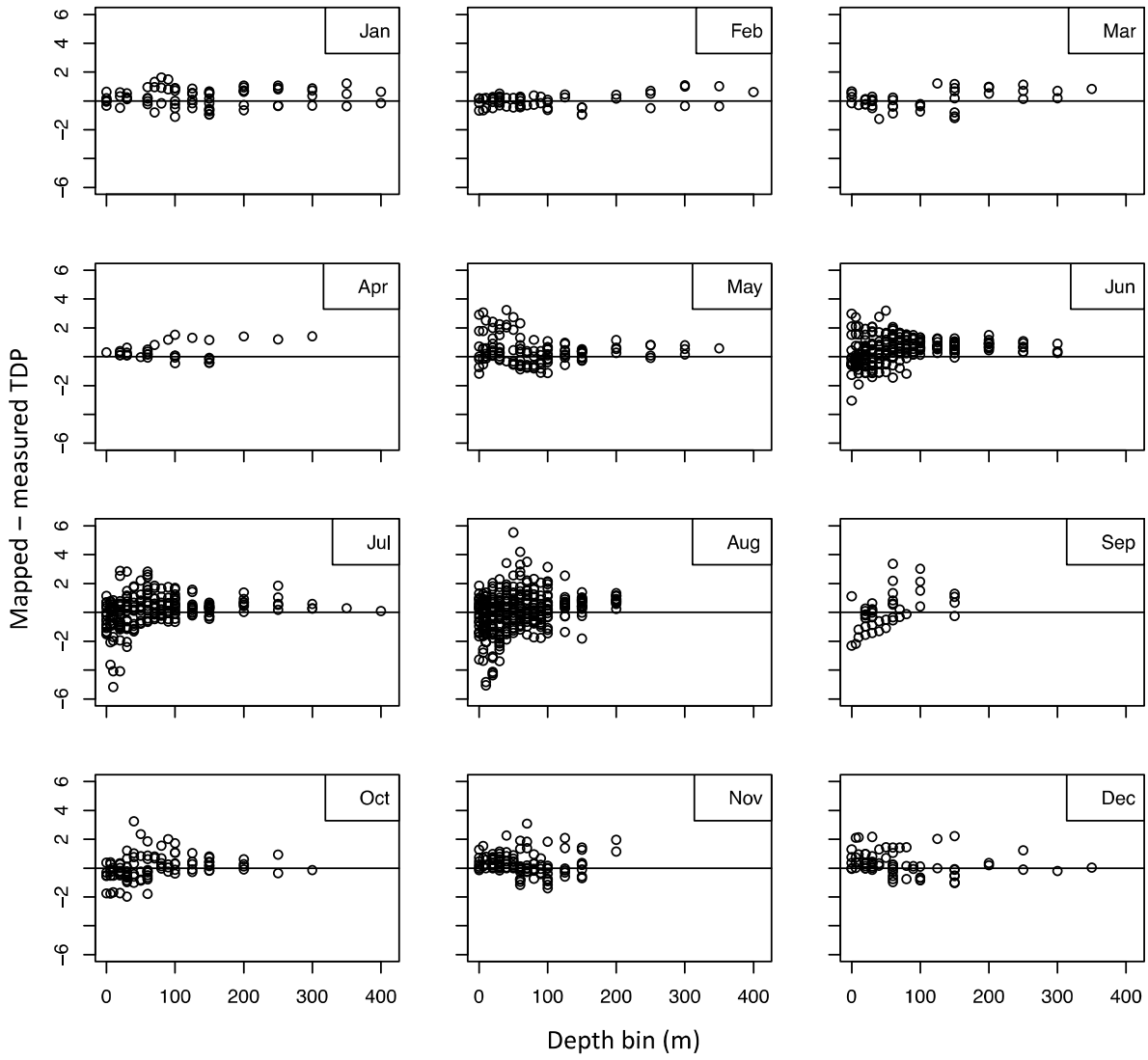
For the SST likelihood, the number of unique observations (locations and times) was 85. No temporal variance in the difference between measured and mapped SST values was observed by season (Kruskal-Wallis, p = 0.2463) or month (Kruskal-Wallis, p = 0.2379). Therefore, a single value of RMSE (0.9) was calculated to parameterize the SST variance in the model. The

446 difference between measured and mapped SST values decreased with distance from shore,
447 however (Figure 5), where the RMSE from locations > 50 km from shore (n=22) dropped to 0.5.



448

449 Figure 5. Difference between sea surface temperature (SST) measured by Microwave
450 Telemetry X-tag Pop-up Satellite Archival Tags (PSATs) and values predicted by the MUR SST
451 satellite-derived map at known locations plotted by distance from shore. Vertical line indicates
452 50 km from shore.



453

454 Figure 6. Difference between temperature measured by Microwave Telemetry X-tag Pop-up
455 Satellite Archival Tags (PSATs) and values predicted by the HYCOM global model at known
456 locations by month and depth bin.

457

458 For the TDP likelihood, the number of unique observations (locations and times) was 89. The
459 sample size was slightly higher than the SST sample size because sometimes TDP values were
460 available only for depths greater than 10 m, and thus SST values could not be calculated. In
461 contrast to SST, the difference between measured and mapped TDP values varied both
462 seasonally (Kruskal-Wallis, p = 2.6 exp-04 for depths 0 – 50 m, p = 1.401 exp-06 for depths 50
463 – 100 m, and p = 1.374 exp-04 for depths > 100 m) and monthly (Kruskal-Wallis, p = 1.74 exp-
464 10 for depths 0 – 50 m, p = 6.35 exp-09 for depths 50 – 100m, and p = 6.47 exp-05 for depths >

100m, Figure 6). Therefore, RMSE was calculated for each depth bin (n = 27) and month of the year. RMSE values were low (<1 °C) at all depths during the winter, but RMSE values almost doubled in shallower waters during the summer months (Figure 7). In contrast to SST, the difference between measured and mapped values did not decrease appreciably with distance from shore (Figure 8).

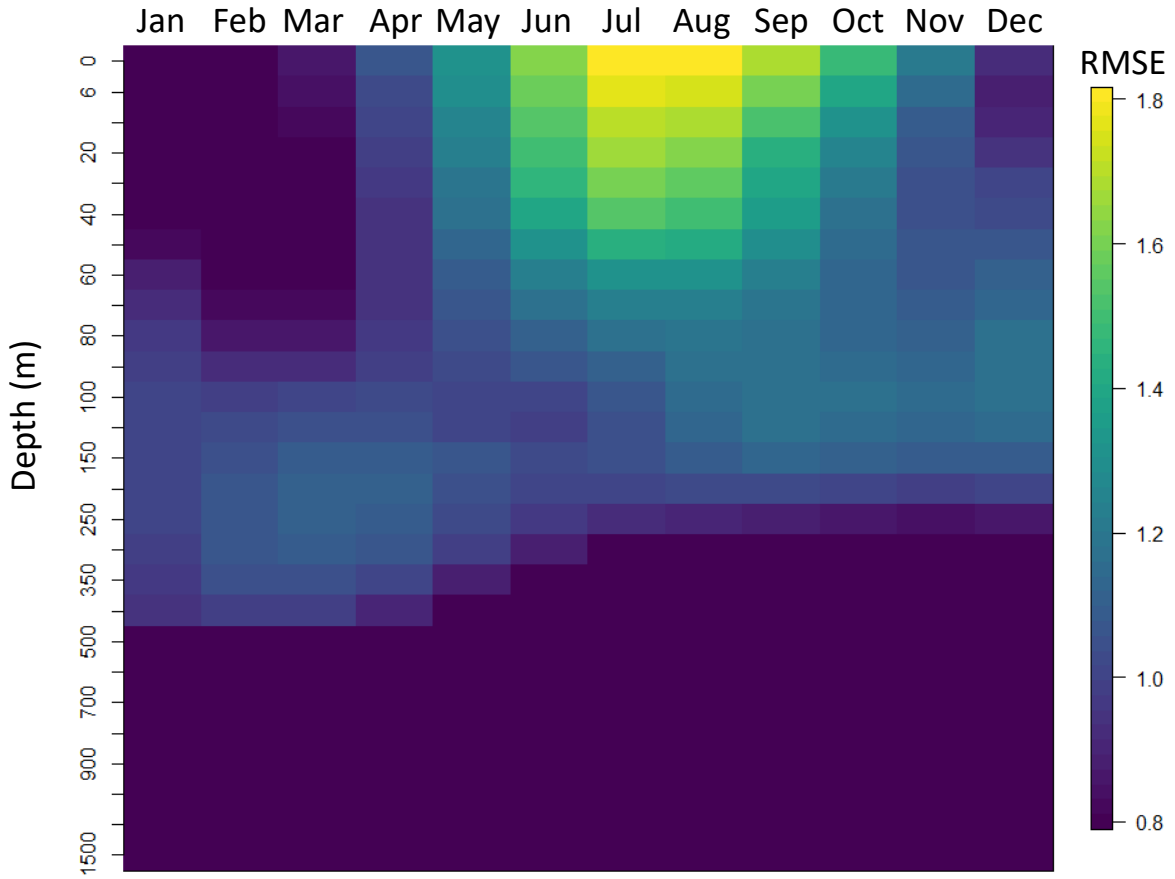


Figure 7. Root mean square error (RMSE) values for the temperature-depth profile (TDP) likelihood. RMSE values reflect differences between temperature at depth values measured by Microwave Telemetry X-tag Pop-up Satellite Archival Tags (PSATs) and values predicted by the HYCOM global hydrodynamic model at known locations. Note that depth bins are not graphed to scale.

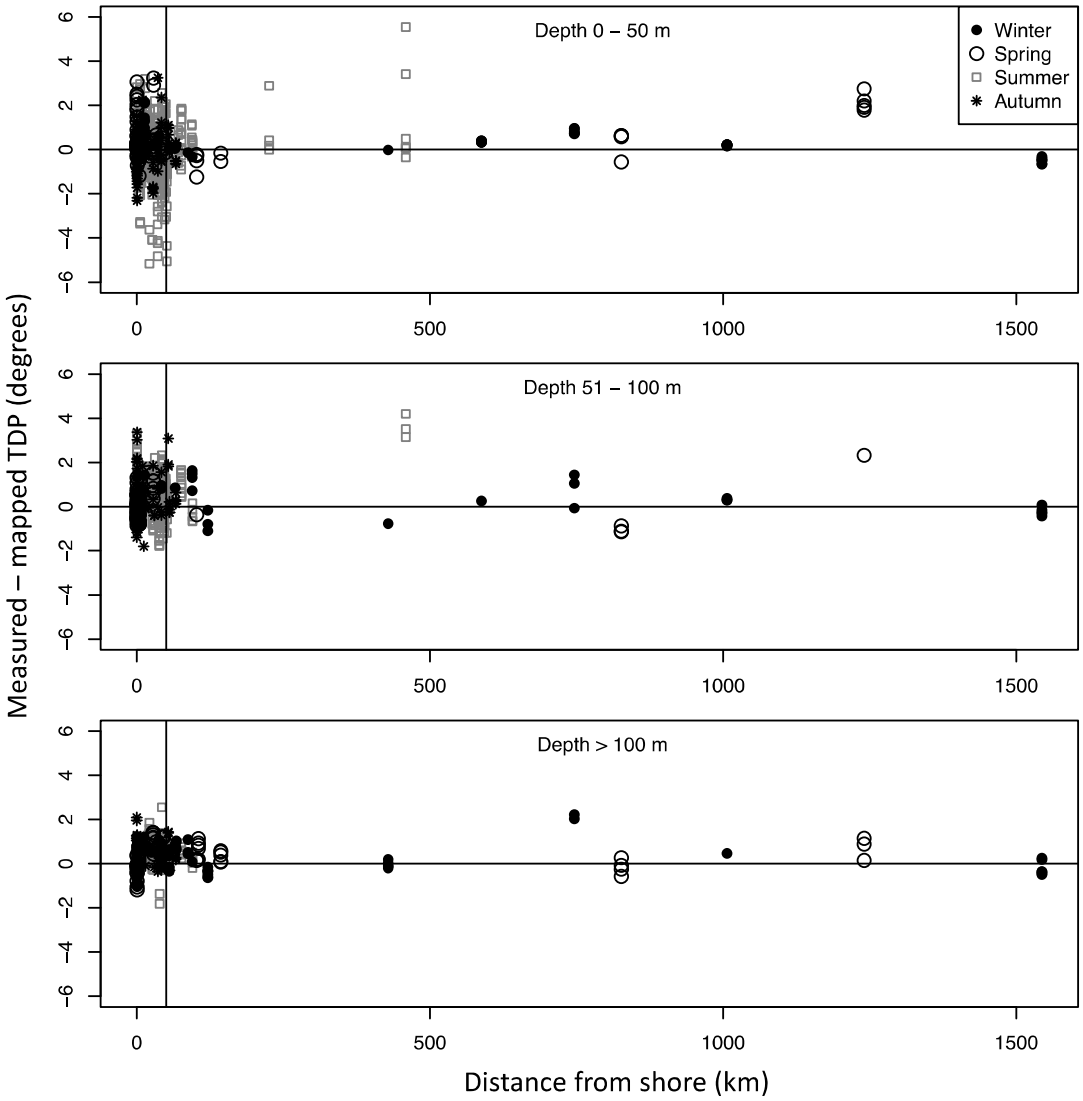


Figure 8. Difference between temperature-depth profile (TDP) measured by Microwave Telemetry X-tag Pop-up Satellite Archival Tags (PSATs) and values predicted by the HYCOM global model at known locations plotted by distance from shore. Observations are grouped into 0 – 50 m, 51 – 100 m, and >100 m depth bins. Vertical line indicates 50 km from shore.

4. Discussion

In this paper, we develop a more flexible application of the HMM which allows for smooth integration and evaluation of accessory data. Due to the large number of tags available, we were able to develop methods for parameterizing the data likelihood model and quantify uncertainty between tag measurement and map data. The HMM demonstration is supported by data, minimizing proxies and assumptions, and is likely to result in an improved estimate of daily locations and overall movement of each tagged animal.

4.1. Data likelihood model

We developed a data likelihood model for the HMM that is tailored to the behavior of our tagged fish, the type of data provided by the PSATs used, and the physical characteristics of our study area. The inclusion of multiple geolocation variables in the data likelihood model is helpful for working with PSAT data, where many gaps may exist due to incomplete signal transmission, and when the behavior of the animal reduces the number of geolocation variables for substantial time periods. For example, dogfish can occupy deep waters or inhabit regions with seasonally limited light for long periods of time, during which the TDP and maximum depth likelihoods are the only source of geolocation information available. Expanding the data likelihood model to include known locations from acoustic telemetry or Smart Position and Temperature Transmitting (SPOT) tags can be easily accomplished by assigning positive likelihood values to grid cells within an acoustic receiver's detection distance or position error radius of the SPOT tag, further expanding the utility of the HMM tool.

Though the data likelihood model is conceptually simple, there are many decisions that need to be made to parameterize it. For example, the semi-pelagic behavior of dogfish means that bathymetry is treated differently than demersal fish for the maximum depth likelihood (Nielsen et al. 2019). In addition, the tendency for dogfish to conduct rapid dives to deep waters and return to shallower waters, combined with the challenges of X-tag compression algorithms for measuring temperature at depth accurately during periods of rapid changes, means that the temperature-depth likelihood for our model may differ slightly from models designed to use binned TDP data from other tag types (e.g., the Wildlife Computers PDT product). Therefore, although likelihood models may seem similar conceptually, there are often important differences in the details that researchers should be aware of before they apply a data likelihood model to their own data. However, open source software such as HMMoce (Braun et al. 2018a) is

transparent, flexible, and can be readily adapted for specific applications (Haase et al. 2021, Hoffmayer et al. 2021).

4.2. Empirical parameterization of SST and TDP likelihoods

Our research expands recent efforts to explore the implications of different methods for grid cell variance specification (Nielsen et al. 2019). The choice of variance parameterization method depends on the characteristics of the study area (such as gradient strength), the resolution and accuracy of available geolocation variable maps in the study area, and the availability of auxiliary data in the study area. Each method has pros and cons. For example, estimation of Gaussian parameters such as variance in the model can lead to errors that affect model performance in state-space models (Auger-Méthé et al. 2016).

Here we present a method for parameterizing variance in situations where map error is high based on empirical comparison of values measured by the tag to mapped values at known locations. The intent of the empirical variance is to estimate the probability distribution of values that could occur within each model grid cell. In effect, the empirical variance estimate accounts for both map error and the spatial scale of variation in the geolocation variable. The resulting variance values are customized for the study because they integrate the behavior of the animal, the type and quality of data provided by the PSAT, and the accuracy of the geolocation map. For example, our empirical variance (0.9 °C) for SST is nearly twice as high as the typical value that accompanies the MUR SST data set (approximately 0.5 °C). This value reflects both the difficulty in mapping SST in nearshore conditions and our definition of SST as the maximum temperature at depths (including tag resolution) less than 11 m. Thus, we simultaneously validated our SST definition for tags that provide only time series data (as opposed to tags that specifically record the temperature when the tag is at the surface) and the difference between that SST measurement and mapped values. Because the variance at locations > 50 km offshore was similar to the typical value provided by the data set, it is likely that the larger variance value is due primarily to map error in nearshore waters. Therefore, researchers with offshore study areas would likely not need to bother with empirical estimation of variance for the SST likelihood. However, some caution is needed in the comparison of nearshore and offshore variance due to low sample sizes for observations > 50 km from shore.

In contrast to SST, our empirical variance estimate for the TDP likelihood varied strongly by month and did not decrease with distance from shore. This indicates that it is a characteristic of

the HYCOM data set that should be considered for all future studies that utilize this geolocation variable. In our situation, the TDP likelihood is most needed in the winter, when spiny dogfish may spend months in deeper waters and therefore the higher variance during the summer months should not cause problems for geolocation accuracy. The TDP provides similar geolocation information to the SST, which is frequently available during the summer months when the HYCOM model error is high. It is worth noting that other temperature-depth profile maps, such as Regional Ocean Monitoring System (ROMS) model maps, may be available in different study areas.

We acknowledge that without independent location information for the tagged fish, accuracy and precision of the geolocation model is unknown. Ideally, a comparison of HMM daily estimated locations with auxiliary data such as SPOT tags, which can provide satellite position estimates when the tagged animal is at the surface, is used to confirm improvements to model performance (Braun et al. 2018a, Gatti et al. 2021). Given our lack of auxiliary location data, additional research with simulated data sets could be beneficial for further understanding the effects of data likelihood model specification methods on performance. Simulations were not attempted in this study because TDP values assigned to simulated locations based on mapped TDP values would likely result in misleading estimates of accuracy and precision given the high map error observed (Gatti et al. 2021). Therefore, future research that features double-tagging with tags that can provide independent location information (such as SPOT or acoustic tags) would likely be the most valuable approach for assessing model performance.

In addition to comparing different variance specification methods, future research could address the effect of including different combination of data likelihood model components. For example, because TDP map error was smaller for waters > 100 m year-round, model performance could be improved if only TDP bins > 100 m are used. In addition, model performance could be improved if the 0 and 6 m TDP depth bins were removed, as these bins have the highest map error during the summer and are also included in the SST likelihood. Performance could also be higher if the TDP likelihood is not included in the model when SST data are available, as likelihood values were much higher for SST compared to TDP at known locations. In addition, a general investigation into assigning weights to each likelihood based on its potential quality could be a valuable contribution for data likelihood model development. Future research could also address the potential effect of spatial and temporal autocorrelation in map errors.

5. Conclusions

Applications of the HMM geolocation model are growing worldwide, yet research on the effects of different parameterization methods on model performance is needed to ensure that models are adapted properly for different study conditions. Our data likelihood model for dogfish in the North Pacific Ocean is specifically customized for Microwave X-tag PSATs, the behavior of the tagged fish, and study area characteristics. It provides an example of data likelihood model customization that should be considered when applying the model to new species or new study areas. The information on SST and TDP variance provided by this study will be helpful for other movement studies of mesopelagic species with Microwave Telemetry X-tag PSATs in the North Pacific Ocean.

Acknowledgements

Tagging for this project involved multiple field parties, agencies, and even countries. For tags released within Alaska, we thank staff at the Alaska Fisheries Science Center (Katy Echave, Dave Clausen, James Murphy, Dana Hanselman, Jon Heifetz, Chris Lunsford, Pat Malecha, Cara Rodgveller, Pete Hulson, Dave Csepp, Kari Fenske), the NMFS Alaska Regional Office (Jason Gasper), the Alaska Department of Fish and Game (Kamala Carroll), the University of Alaska Fairbanks (Thomas Farrugia, Megan Petersen), the University of North Carolina (Ben Edwards), the crews of the fishing vessels F/V Alaskan Leader and Ocean Prowler, and Captain Scott Chadwick, Yakutat Charters. For tags released in Canadian waters, we thank Jackie King at the Department of Fisheries and Oceans Canada. For tags released in Puget Sound, we thank staff at the Northwest Fisheries Science Center (Kelly Andrews, Nick Tolimieri, and Greg Williams) and the Seattle Aquarium (Joel Hollander, Andy Sim, Tim Carpenter, Chris Van Damme, and Bryan McNeil). We thank Kevin Siwicke, Chris Lunsford, Susanne McDermott (AFSC), and two anonymous reviewers for providing valuable feedback on the manuscript.

Funding

This research was funded by the National Oceanic and Atmospheric Administration. The funding source was involved with all aspects of project design, analyses, and manuscript preparation.

References

Andersen, K. H., A. Nielsen, U. H. Thygesen, H. H. Hinrichsen, and S. Neuenfeldt. 2007. Using the particle filter to geolocate Atlantic cod (*Gadus morhua*) in the Baltic Sea, with special

626 emphasis on determining uncertainty. Canadian Journal of Fisheries and Aquatic
627 Sciences **64**:618-627.

628 Auger-Méthé, M., C. Field, C. M. Albertsen, A. E. Derocher, M. A. Lewis, I. D. Jonsen, and J.
629 Mills Flemming. 2016. State-space models' dirty little secrets: even simple linear
630 Gaussian models can have estimation problems. Scientific Reports **6**:26677-26677.

631 Biais, G., Y. Coupeau, B. Séret, B. Calmettes, R. Lopez, S. Hetherington, and D. Righton. 2017.
632 Return migration patterns of porbeagle shark (*Lamna nasus*) in the Northeast Atlantic:
633 implications for stock range and structure. ICES Journal of Marine Science **74**:1268-
634 1276.

635 Braun, C. D., B. Galuardi, and S. R. Thorrold. 2018a. HMMoce: An R package for improved
636 geolocation of archival-tagged fishes using a hidden Markov method. Methods in
637 Ecology and Evolution **doi: 10.1111/2041-210X.12959**.

638 Braun, C. D., G. B. Skomal, and S. R. Thorrold. 2018b. Integrating archival tag data and a high-
639 resolution oceanographic model to estimate basking shark (*Cetorhinus maximus*)
640 movements in the western Atlantic. Frontiers in Marine Science **doi:**
641 **10.3389/fmars.2018.00025**.

642 Carlson, A. E., E. R. Hoffmayer, C. A. Tribuzio, and J. A. Sulikowski. 2014. The use of satellite
643 tags to redefine movement patterns of spiny dogfish (*Squalus acanthias*) along the U.S.
644 east coast: implications for fisheries management. PLoS ONE **9**:e103384.

645 Costa, D. P., G. A. Breed, and P. W. Robinson. 2012. New Insights into Pelagic Migrations:
646 Implications for Ecology and Conservation. Annual Review of Ecology, Evolution, and
647 Systematics **43**:73-96.

648 Doherty, P. D., J. M. Baxter, F. R. Gell, B. J. Godley, R. T. Graham, G. Hall, J. Hall, L. A.
649 Hawkes, S. M. Henderson, L. Johnson, C. Speedie, and M. J. Witt. 2017. Long-term
650 satellite tracking reveals variable seasonal migration strategies of basking sharks in the
651 north-east Atlantic. Scientific Reports **7**:42837. doi: 42810.41038/srep42837.

652 Donlon, C., I. Robinson, K. S. Casey, J. Vazquez-Cuervo, E. Armstrong, O. Arino, C.
653 Gentemann, D. May, P. LeBorgne, J. Piollé, I. Barton, H. Beggs, D. J. S. Poulter, C. J.
654 Merchant, A. Bingham, S. Heinz, A. Harris, G. Wick, B. Emery, P. Minnett, R. Evans, D.
655 Llewellyn-Jones, C. Mutlow, R. W. Reynolds, H. Kawamura, and N. Rayner. 2007. The
656 Global Ocean Data Assimilation Experiment High-resolution Sea Surface Temperature
657 Pilot Project. Bulletin of the American Meteorological Society **88**:1197 - 1213.

658 Gatti, P., J. A. D. Fisher, F. Cyr, P. S. Galbraith, D. Robert, and A. Le Bris. 2021. A review and
659 tests of validation and sensitivity of geolocation models for marine fish tracking. Fish and
660 Fisheries **22**:1041-1066.

661 Goethel, D. R., T. J. Quinn, and S. X. Cadrin. 2011. Incorporating spatial structure in stock
662 assessment: movement modeling in marine fish population dynamics. Reviews in
663 Fisheries Science **19**:119-136.

664 Gramacy, R. B. 2007. tgp: An R package for Bayesian nonstationary, semiparametric nonlinear
665 regression and design by treed Gaussian process models. Journal of Statistical Software
666 **19**:1-46.

667 Haase, S., U. Krumme, U. Gräwe, C. D. Braun, and A. Temming. 2021. Validation approaches
668 of a geolocation framework to reconstruct movements of demersal fish equipped with
669 data storage tags in a stratified environment. Fisheries Research **237**:105884.

670 Hoffmayer, E. R., J. A. McKinney, J. S. Franks, J. M. Hendon, W. B. Driggers, B. J. Falterman,
671 B. Galuardi, and M. E. Byrne. 2021. Seasonal occurrence, horizontal movements, and
672 habitat use patterns of whale sharks (*Rhincodon typus*) in the Gulf of Mexico. Frontiers
673 in Marine Science **7**.

674 JPL MUR MEaSUREs Project. 2010. GHRSST Level 4 MUR Global Foundation Sea Surface
675 Temperature Analysis. Ver. 2. PO.DAAC, CA, USA.

676 Kawai, Y., and A. Wada. 2007. Diurnal Sea Surface Temperature variation and its impact on the
677 atmosphere and ocean: A review. Journal of Oceanography **63**:721-744.

678 King, J., G. A. McFarlane, V. Gertseva, J. Gasper, S. Matson, and C. A. Tribuzio. 2017. Shark
679 interactions with directed and incidental fisheries in the northeast Pacific Ocean: Historic
680 and current encounters, and challenges for shark conservation. Advances in Marine
681 Biology **78**:9-44.

682 Lam, C. H., A. Nielsen, and J. R. Sibert. 2008. Improving light and temperature based
683 geolocation by unscented Kalman filtering. Fisheries Research **91**:15-25.

684 Le Bris, A., A. Frechet, and J. S. Wroblewski. 2013. Supplementing electronic tagging with
685 conventional tagging to redesign fishery closed areas. Fisheries Research **148**:106-116.

686 Liu, C., G. W. Cowles, D. R. Zemeckis, S. X. Cadrin, and M. J. Dean. 2017. Validation of a
687 hidden Markov model for the geolocation of Atlantic cod. Canadian Journal of Fisheries
688 and Aquatic Sciences **74**:1862-1877.

689 Lowerre-Barbieri, S. K., R. Kays, J. T. Thorson, and M. Wikelski. 2019. The ocean's movescape:
690 fisheries management in the bio-logging decade (2018–2028). ICES Journal of Marine
691 Science **76**:477-488.

692 Malecha, P., C. Rodgveller, C. Lunsford, and K. Siwicke. 2019. The 2018 longline survey of the
693 Gulf of Alaska and eastern Aleutian Islands on the FV *Alaskan Leader*: Cruise Report
694 AL-18-01. AFSC Processed Rep. 2019-02, 30 p. Alaska Fish. Sci. Cent., NOAA, Natl.
695 Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

696 McFarlane, G. A., and J. R. King. 2003. Migration patterns of spiny dogfish (*Squalus acanthias*)
697 in the North Pacific Ocean. Fishery Bulletin, U.S. **101**:358-367.

698 Musyl, M. K., R. W. Brill, D. S. Curran, J. S. Gunn, J. R. Hartog, R. D. Hill, D. W. Welch, J. P.
699 Eveson, C. H. Boggs, and R. E. Brainard. 2001. Ability of archival tags to provide
700 estimates of geographical position based on light intensity. Pages 343–367 *in* J. R.
701 Sibert and J. L. Nielsen, editors. Electronic tagging and tracking in marine fisheries.
702 Kluwer Academic Publs, Dordrecht, The Netherlands.

703 Nielsen, A., K. A. Bigelow, M. K. Musyl, and J. R. Sibert. 2006. Improving light-based
704 geolocation by including sea surface temperature. Fisheries Oceanography **15**:314-325.

705 Nielsen, J. K., F. Mueter, M. Adkison, S. McDermott, T. Loher, and A. C. Seitz. 2019. Effect of
706 study area bathymetric heterogeneity on parameterization and performance of a depth-
707 based geolocation model for demersal fish. Ecological Modelling **402**:18-34.

708 Nielsen, J. K., F. J. Mueter, M. D. Adkison, T. Loher, S. F. McDermott, and A. C. Seitz. 2020.
709 Potential utility of geomagnetic data for geolocation of demersal fishes in the North
710 Pacific Ocean. Animal Biotelemetry **8**:17.

711 Pedersen, M. W., T. A. Patterson, U. H. Thygesen, and H. Madsen. 2011. Estimating animal
712 behavior and residency from movement data. Oikos **120**:1281-1290.

713 Pedersen, M. W., D. Righton, U. H. Thygesen, K. H. Andersen, and H. Madsen. 2008.
714 Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and
715 behavioural switching. Canadian Journal of Fisheries and Aquatic Sciences **65**:2367-
716 2377.

717 Schaefer, K. M., and D. W. Fuller. 2016. Methodologies for investigating oceanodromous fish
718 movements: archival and pop-up satellite archival tags. Pages 251-289 *in* P. Morais and
719 F. Daverat, editors. An introduction to fish migration. CRC Press, Boca Raton, FL, USA.

720 Seitz, A. C., B. L. Norcross, D. Wilson, and J. L. Nielsen. 2006. An evaluation of light-based
721 geolocation for demersal fish in high latitudes. Fishery Bulletin, U.S. **104**:571-578.

722 Skomal, G. B., S. I. Zeeman, J. H. Chisholm, E. L. Summers, H. J. Walsh, K. W. McMahon, and
723 S. R. Thorrold. 2009. Transequatorial Migrations by Basking Sharks in the Western
724 Atlantic Ocean. Current Biology **19**:1019-1022.

725 Taylor, I. G. 2008. Modeling spiny dogfish population dynamics in the Northeast Pacific. Ph.D.
726 dissertation, University of Washington, Seattle WA.

727 Thygesen, U., M. Pedersen, and H. Madsen. 2009. Geolocating Fish Using Hidden Markov
728 Models and Data Storage Tags. Pages 277-293 *in* J. Nielsen, H. Arrizabalaga, N.
729 Fragoso, A. Hobday, M. Lutcavage, and J. Sibert, editors. Tagging and Tracking of
730 Marine Animals with Electronic Devices. Springer Netherlands.

731 Tribuzio, C. A., M. E. Matta, K. Echave, and C. Rodgveller. 2020. Assessment of the shark
732 stock complex in the Gulf of Alaska. In: Stock assessment and fishery evaluation report
733 for the groundfish resources of the Gulf of Alaska. North Pacific Fishery Management
734 Council, 1007 W 3rd STE 400, Anchorage, AK 99501. 70pp.

735 Wallcraft, A. J., E. J. Metzger, and S. N. Carroll. 2009. Software design description for the
736 HYbrid Coordinate Ocean Model (HYCOM) Version 2.2. Report number NRL/MR/7320--
737 09-9166, Naval Research Laboratory, Stennis Space Center, MS.

738 Woillez, M., R. Fablet, T.-T. Ngo, M. Lalire, P. Lazure, and H. de Pontual. 2016. A HMM-based
739 model to geolocate pelagic fish from high-resolution individual temperature and depth
740 histories: European sea bass as a case study. Ecological Modelling **321**:10-22.

741 Zar, J. H. 1999. Biostatistical analysis. Prentice Hall, Upper Saddle River, New Jersey.

742 Zeileis, A., and G. Grothendieck. 2005. zoo: S3 infrastructure for regular and irregular time
743 series. Journal of Statistical Software **14**:1-27.
744