

RESEARCH

Open Access



Geolocation of a demersal fish (Pacific cod) in a high-latitude island chain (Aleutian Islands, Alaska)

Julie K. Nielsen^{1*}, David R. Bryan², Kimberly M. Rand³, Martin C. Arostegui⁴, Camrin D. Braun⁴, Benjamin Galuardi⁵ and Susanne F. McDermott²

Abstract

Background Fish geolocation methods are most effective when they are customized to account for species behavior and study area characteristics. Here, we provide an example of customizing a hidden Markov model (HMM) for reconstructing movement pathways of a high-latitude demersal fish species in a remote island chain using Pop-up Satellite Archival Tag (PSAT) data. Adult Pacific cod were tagged with PSATs while occupying winter spawning grounds in the Aleutian Islands in February 2019. We adapted a demersal fish application of the HMM to (1) add light-based longitude to the data likelihood model, (2) account for possible off-bottom behavior of demersal fishes in the maximum daily depth likelihood, and (3) modify the model framework to accommodate convoluted island topography in the study area. A simulation study was conducted to explore the two primary modifications to the model framework, reflecting boundary for the movement kernel and the Viterbi method of pathway reconstruction, under known conditions.

Results Geolocation was performed on satellite-transmitted and detailed archival data sets from 6 adult Pacific cod at liberty for 21–277 days. Migration from winter spawning to summer foraging areas (range 60–395 km) was detected for the 4 tagged fish that were at liberty for more than 90 days. Light-based longitude was the primary geolocation variable for detecting migrations with precision (root mean square error) estimates of 0.56 degrees during winter and 1.3 degrees during the summer. Simulation studies confirmed the effectiveness of model framework modifications and generated guidelines for use in specific applications.

Conclusions This study demonstrates that post-spawning migrations of Pacific cod in the Aleutian Islands can be detected and characterized using PSAT data. Initial insights into migrations, summer foraging areas, and associated development of appropriate analysis tools will support future Pacific cod movement studies in the Aleutian Islands as well as other regions of Alaska. The adaptations to the HMM presented here will benefit current and future research on demersal fish in other regions as well as fish species that occupy areas with convoluted shorelines or island chain topography.

Keywords Geolocation, Hidden Markov model, Fish migration, Satellite tags, Pop-up Satellite Archival Tags, Viterbi algorithm, Reflecting boundary

*Correspondence:

Julie K. Nielsen

Julie.nielsen@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Pacific cod (*Gadus macrocephalus*) is a demersal fish species of great economic and ecological importance in Alaska, where it is managed in three separate management regions: the Gulf of Alaska, the Bering Sea, and the Aleutian Islands. Pacific cod populations have recently experienced declines in abundance in the Gulf of Alaska [1] and northward shifts in distribution in the Bering Sea [2] in conjunction with warming ocean waters. These recent phenomena have highlighted the management need to understand Pacific cod movement patterns and the mechanisms, such as response to varying ocean temperatures, that underlie them [1].

Pacific cod have been observed to conduct large-scale seasonal migrations between winter spawning and summer foraging areas in Alaska based on recovery locations from several conventional and archival tagging studies [3–5]. However, more detailed information on migration timing, extent, pathways, and the proportion of fish that participate in seasonal migrations is needed to understand seasonal movement between management regions and predict changes in migration patterns that could occur in response to different temperature regimes [1]. In response to these information needs, we initiated a satellite tagging program for Pacific cod in 2019 with a pilot study [6] to determine whether satellite tags could provide more detailed information on seasonal movement patterns.

Pop-up Satellite Archival Tags (PSATs) are a type of archival tag that can provide a wealth of detailed information about movement and behavior of migratory fish species without the need to recapture tagged fish [7–9]. PSATs can record depth, temperature, light, acceleration data, or other types of data depending on tag type and manufacturer. They are programmed to release from the fish on a specified date and transmit archived data to the Argos satellite network, which provides information on the location of the tagged fish on the pop-up date (subsequently referred to as the “end location”). Therefore, data can be obtained without recapturing the tagged fish. Transmitted data, together with release and end locations, can then be used to reconstruct movement pathways of tagged fish using geolocation models that link tag data to likely locations within the study area. Reconstructed movement paths consist of estimated locations (usually with associated uncertainty) at regular points in time (e.g., daily) between release and pop-up dates. By reconstructing movement paths, PSATs can provide details on important aspects of fish spatial dynamics, such as migration timing and locations and identification of summer foraging and winter spawning locations [10].

State-space models are a powerful method for reconstructing movement pathways. They generally consist

of two coupled sub-models: (1) a movement model that predicts the rate and type of movement of the tagged fish within the study area, and (2) a data likelihood model that identifies locations within the study area that best match the data (e.g., light, depth, or temperature) recorded by the tag at each time step [11–14]. State-space models are ideal for geolocation, because they can accommodate complex or incomplete data sets, allow for different behavioral states, such as foraging or migrating, and provide estimates of error associated with reconstructed movement paths [15]. State-space geolocation models range from Gaussian and continuous (e.g., a Kalman filter [12]) to nonparametric and discrete (e.g., a hidden Markov model, HMM [14]). HMMs have become a popular approach for geolocation of a wide range of fish species [16] due to their flexibility, as they can incorporate multiple types of non-linear data and do not place probability on land.

The HMM framework, originally developed for geolocation of Atlantic cod, *Gadus morhua*, in the North Sea [17, 18], is similar among studies. Researchers can choose whether to treat time as discrete [17] or continuous [14]. However, the data likelihood model that specifies how the data from the satellite tag are matched to maps of the study area must be adapted to accommodate different tag types, species, or study areas [19]. Therefore, pilot studies to collect preliminary PSAT data and develop data likelihood models can be beneficial for ensuring that PSATs can answer specific movement questions prior to initiating large-scale movement studies.

The satellite tag pilot study for Pacific cod was conducted in the Aleutian Island management area. The Aleutian Islands is a remote, volcanic island chain that stretches west from Alaska toward Russia for more than one thousand miles. It serves as the boundary between the Bering Sea to the north and the Gulf of Alaska to the south and east. Little is known about the movement of Pacific cod in this region. Conventional and archival tags released in this region require tag recaptures by either a commercial fishery or a research platform (e.g., NOAA bottom trawl or longline assessment surveys). Therefore, limited information on seasonal movement has been obtained in this remote region, because the overall commercial catch is low and the majority of fishing occurs during the winter when Pacific cod are aggregated for spawning [20]. To date, only two tagging studies have occurred there: a conventional tag study [3] and an archival tag study [5]. Sample sizes for both studies were small in the Aleutian Islands, as the primary release locations were in the Bering Sea or the Gulf of Alaska. In light of the limited movement data provided by tags that require recapture in the region, alternative methods for assessing movement of Pacific cod are needed to understand

seasonal migration patterns and inform management in the Aleutian Islands.

Here, we describe the methods used to reconstruct movement pathways of Pacific cod in the Aleutian Islands from PSAT data, while details on capture methods, tag attachment methods, and preliminary insights into Pacific cod seasonal migrations produced by this pilot study are reported by Bryan et al. [6]. We adapted a discrete time HMM for geolocation of demersal fish [17, 19] to include: (1) adding light-based longitude to the maximum daily depth data likelihood model, (2) modifying the maximum daily depth likelihood for fish that are demersal but may not contact the sea floor each day, and (3) refinements to the model framework to accommodate convoluted shorelines or island chain topography. We demonstrate the effectiveness of modifications for convoluted shorelines using simulated movement paths. This work provides an example of adapting a HMM for a specific application and introduces additional tools for reconstruction of movement paths in nearshore areas with complex topography.

Methods

PSAT programming and deployment

Fish tags

In February 2019, MiniPAT satellite tags manufactured by Wildlife Computers, Inc. (Redmond, WA) were

deployed on mature Pacific cod at three locations in the Aleutian Islands (Fig. 1) as part of a larger study on seasonal migration that deployed three types of satellite tags [6]. Fish were captured using a combination of pots and trawls at depths ranging from 88 to 118 m. Only fish that showed no signs of injury were selected for tagging. MiniPATs were attached to fish with a “backpack” configuration originally developed for salmonid species [21] that consisted of (1) a padded harness secured to the fish with wires through the dorsal musculature at two locations, and (2) a 150-lb test monofilament tether that connected the tag to the harness. After tagging, fish were carefully lowered to a depth of 50 m with a SeaQualizer™ descender device to facilitate recompression [22].

The tags were programmed to detach from the fish on a given date, after which the tags “popped up” to the sea surface and transmitted their data to the Argos satellite system. The tags provided two types of location data: (1) known locations (to within 250–1500 m) at the end of the deployment from Argos positioning of transmitting tags, and (2) reconstruction of fish movement paths throughout the deployment period using the archived data (see “geolocation” section below). Pop-up dates for the tags were staggered throughout the year to provide known locations during different seasons. Tags were programmed to pop up after 90 days ($n=3$), 180 days ($n=8$), 270 days ($n=2$) and 360 days ($n=8$).

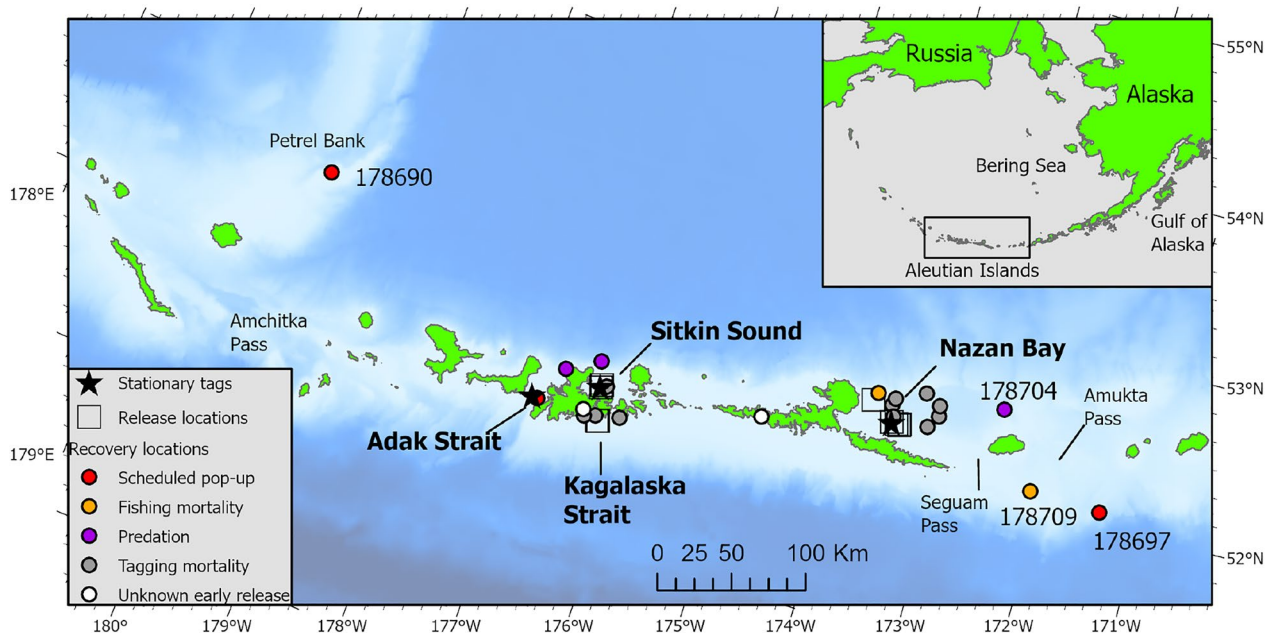


Fig. 1 Study area in the Aleutian Islands. Tagged Pacific cod (hollow squares) were released in Sitkin Sound, Kagalaska Strait, and Nazan Bay. Stationary tags (stars) were deployed in Sitkin Sound, Adak Strait, and Nazan Bay. Fish recovery locations (filled circles) are color-coded by fate: alive until the scheduled pop-up date (red), recaptured in commercial fisheries (orange), marine mammal predation (purple), tagging mortality (gray), and unknown early release (white)

The MiniPAT satellite tags were 124 mm long, 38 mm wide, and weighed 60 g in air. They recorded information on depth (0–1700 m, resolution 0.5 m), temperature (−40 °C–60 °C, resolution 0.05 °C), tri-axial acceleration (−2–2 g, resolution 0.05 g), and light levels ($5 \times 10^{-12} \text{W} \cdot \text{cm}^{-2}$ – $5 \times 10^{-2} \text{W} \cdot \text{cm}^{-2}$) at intervals ranging from 1 to 5 s depending on programmed deployment duration. If a tag was physically recovered, the entire high-resolution data set was available for analysis. Archived data were summarized for transmission to Argos satellites according to the length of programmed deployment. Summarized data for all tags included: (1) light levels during dawn and dusk, from which latitude and longitude were estimated using GPE2 software from Wildlife Computers, (2) daily minimum and maximum temperature and depth, and (3) temperature–depth profiles in 12-h time bins. Time-series depth and temperature were generated at 150-s intervals for 90-day tags, 450-s intervals for 180-day tags, and 600-s intervals for 270-day and 360-day tags. For 360-day tags, depth and temperature time series were generated on alternate days. MiniPATs deployed for 90 days or less also provided accelerometer-derived activity metrics “knockdown” and “% time tilted” [23]. Briefly, these accelerometer activity metrics are derived from changes in the vertical orientation of the tethered tag, where a buoyant PSAT on a fish that is not swimming is vertical, while a PSAT on a swimming fish approaches a horizontal orientation due to the effect of drag as the tag is towed behind the swimming fish. A “knockdown” is recorded when the change in vertical orientation of the tag exceeds a threshold range of movement during a 10-s interval and knockdowns are summed within the summary period (e.g., 1- or 2-h time bins), whereas the “% time tilted” refers to the percentage of time that the tag was tilted beyond a threshold value of tilt during the summary period.

Stationary tags

Three additional MiniPATs were deployed as stationary reference tags in areas, where tagged fish were released (Fig. 1) to assist the process of geolocation (see below). Tags were moored 2 m above the sea floor at a depth of approximately 100 m and programmed to pop-up after 360 days. The stationary tag provided light levels at dawn and dusk as well as time series depth and temperature collected at 600-s intervals. Temperature time series were recorded daily, and depth time series every other day.

Analyses

All analyses were conducted using the R statistical environment, version 4.1.3 [24].

Accuracy of longitude and latitude estimates

Data from stationary tags were analyzed to determine the accuracy and bias of light-based longitude and latitude estimates. Accuracy was calculated as the root mean square error between estimated and known values [16]. Bias was assessed by performing a one-sample Welch’s *t* test on the difference between estimated and known values under the hypothesis that the mean difference was equal to zero.

Geolocation

We adapted a discrete time HMM [17, 25] to reconstruct movement pathways of Pacific cod in the Aleutian Islands. The model framework and process model have been thoroughly described in other manuscripts [17, 19, 25, 26], so it is only briefly outlined here. The model features a study area that is divided into grid cells and ultimately provides the probability that the tagged fish occupied each grid cell on each day of deployment. The model consists of a forward filter followed by backward smoothing. The forward filter begins with a probability of 1 in the grid cell, where the tagged fish were released and a probability of zero in all other grid cells. The probability in each grid cell is then iteratively updated by the movement model (convolution with an isotropic diffusion kernel) followed by cellwise multiplication with the data likelihood model (see below) at each time step. Once the forward filter is completed, backward smoothing is conducted to update all probabilities with the knowledge of the end location. The primary model output consists of a three-dimensional array (latitude, longitude, time) of posterior probability estimates (i.e., for each time step, a matrix of study area grid cells contains the probability that each grid cell was occupied by the tagged fish at that time step). The array can then be summed for specific periods (e.g., the whole trajectory or different seasons) to describe the overall location probability in the study area during that period. Behavioral states can be incorporated by allowing for different diffusion rates for different behaviors (e.g., migrating versus foraging).

Data likelihood models describe the likelihood of the measured variable occurring in a particular model grid cell based on maps of geolocation variables within the study area. They vary based on fish behavior (e.g., demersal or pelagic), tag type (e.g., available sensors and data formats), study area characteristics (e.g., orientation and strength of geolocation gradients), and quality/availability of geolocation variable maps in the region. Developing a data likelihood model is a key step in customizing the HMM for specific applications.

In 2019, we developed a data likelihood model for demersal fish in the North Pacific Ocean [19] based on

maximum daily depth. The maximum daily depth likelihood for demersal fishes is modeled on Pacific halibut and assumes that the tagged fish contacts the seafloor at least once during each time step. Thus, the maximum depth of the tagged fish is assumed to be the depth of the seafloor and can then be linked to bathymetric maps of the study area during the geolocation process. The likelihood value in each grid cell is determined by integrating the distribution of depth values in that cell between the limits of the tag maximum depth plus and minus the tag measurement uncertainty accompanying each depth observation [27]. Note that given two grid cells with the same mean but different depth standard deviations within the cell, this method of calculating likelihoods will produce higher likelihood values for grid cells with a narrow range of potential depth values (e.g., low-gradient areas) compared to high-gradient areas (see caveats in discussion).

Some evidence suggests that Pacific cod may exhibit some degree of off-bottom behavior, so the assumption of contact with the seafloor every day may not always be warranted. Off-bottom behavior may occur during recovery from tagging [28] and also has been observed for Atlantic cod during migration [29] and during summer foraging in areas adjacent to steep drop-offs [30]. Therefore, we modified the maximum depth likelihood to allow the maximum daily depth to be some distance above the seafloor using a split-normal distribution for depth in each grid cell. A split normal distribution allows the specification of separate standard deviations for each half of the probability density function (PDF):

$$f(x) = \begin{cases} A \exp\left[-\frac{(x-\mu)^2}{2\sigma_1^2}\right] & x \leq \mu \\ A \exp\left[-\frac{(x-\mu)^2}{2\sigma_2^2}\right] & x \geq \mu \end{cases} \quad (1)$$

where $A = (\sqrt{2\pi}(\sigma_1 + \sigma_2)/2)^{-1}$, μ is the bathymetry cell mean, σ_1 is the standard deviation for the left (shallow) side of the bathymetry PDF, and σ_2 is the standard deviation for the right (deeper) side of the bathymetry PDF [31]. Therefore, the probability distribution accounts for bathymetric uncertainty on the deeper portion of the grid cell mean and both bathymetric uncertainty and the probability that the fish could be slightly above the seafloor on the shallower portion [32]. A constant value for depth standard deviation was selected to denote off-bottom uncertainty (referred to subsequently as the "off-bottom constant") as part of the model fitting process. Then, for each grid cell, σ_1 (the shallower portion) was calculated as the root sum square of the off-bottom constant and the standard deviation of depth associated with bathymetric uncertainty (i.e., the range of possible depth values likely to be present within the grid cell), while σ_2

(the deeper portion) consisted of only the bathymetric uncertainty.

The likelihood value for depth for each grid cell is thus calculated as

$$L_{\text{Depth}} = \int_{z_1}^{z_2} SN(z; \mu_z, \sigma_1, \sigma_2) dz \quad (2)$$

where z is maximum depth measured by the fish during the time interval, z_1 and z_2 are the lower and upper limits of uncertainty in tag measurement, SN denotes a split-normal distribution, μ_z is the mean of the bathymetry grid cell, σ_1 is the standard deviation used for the shallow portion of the distribution, and σ_2 is the standard deviation for the deep portion of the distribution.

We used a 100-m resolution bathymetry grid of the Aleutian Islands [33] to provide bathymetry information for the study area. Depth data were aggregated into 2-km model grid cells. The resulting mean and standard deviation of depths within each model grid cell were used to calculate the likelihood.

The second modification to the data likelihood model consisted of adding light-based longitude. Though the archived data provide estimates of both latitude and longitude, only longitude is used in the model, because longitude values are more robust than latitude in most light-based geolocation applications [34, 35] and particularly for demersal fish [36]. Furthermore, strong north-south depth gradients provide more precise information about latitude in our study area (a maximum north-south error of approximately 50 km is present using depth data, which accounts for possible occupation of a specific depth contour on the north or south side of the island chain, compared to the RMSE for latitude provided in the results section). Longitude values were filtered manually to remove estimates with a change of more than 2 degrees per day [37].

The likelihood for light-based longitude consists of a normal probability density:

$$L_{\text{Longitude}} = N(x; \mu, \sigma), \quad (3)$$

where the likelihood value in each grid cell is the probability of the longitude estimated by the PSAT (x), given the longitude of the grid cell (μ) and the expected error in longitude estimates (σ). We chose a value of 1.5 degrees for the expected error, as this value was slightly higher than the root mean square error obtained from stationary tag data. The slightly larger variance was chosen to account for slight changes in depth by fish that are not present in stationary tag data.

Likelihoods are processed separately before they are combined. If PSAT data are unavailable on a given day, all cells for that likelihood receive a value of one for that day.

Each likelihood is standardized relative to the maximum value observed. Depth and light-based longitude likelihoods are combined by cellwise multiplication to generate the total likelihood in each grid cell for each day:

$$L_{\text{Total}} = L_{\text{Depth}} * L_{\text{Longitude}} \quad (4)$$

Additional likelihoods can be easily incorporated by cellwise multiplication in this manner. For example, we explored, but ultimately decided against, adding a temperature–depth profile likelihood to the data likelihood model due to map accuracy issues in the study area (Additional file 1).

Likelihood values for the final day of the trajectory consisted of a value of 1 in the grid cell with the end location and values of 0 in all other study area grid cells. For tags that likely drifted in currents before Argos locations could be determined or for tagged fish that were predated upon, likelihood surfaces for the last day of the trajectory were obtained by creating an additional likelihood surface, where a 30 km buffer was placed around the Argos end location and a value of 1 was assigned to all grid cells within the radius and a value of 0 to all grid cells outside the radius. Then, this “known location” likelihood was combined with the maximum depth likelihood, as shown in Eq. 4. [38].

The model incorporated two behavioral states through the use of different diffusion coefficients applied at each time step to reflect periods of differential movement. These behavioral states corresponded to (1) the recovery/spawning period immediately following tagging, when tagged cod were likely to have lower rates of movement due to occupation of spawning grounds and recovery from capture and tagging, and (2) the migration/foraging movement state that assumed a larger rate of movement once fish leave their winter spawning areas. This recovery/spawning movement state was assigned to all days of the data set for fish with pop-up locations in the study area during March. For fish at liberty beyond March, the recovery/spawning state was assigned to each day that patterns in depth typical of barotrauma recovery were evident [28] and temperature–depth profiles of tagged fish matched profiles of stationary tags and tags from fish with tagging mortality in the release locations. It is important to note that the recovery/spawning state defined here does not reflect any specific knowledge of spawning behavior but simply reflects likely occupation of the area, where they were captured and released during the spawning season. For migratory fish, the migration/foraging state was then assigned to all remaining days in the data set after the recovery/spawning period dates were assigned.

Procedures for determining parameter values for diffusion and the off-bottom constant varied by movement

state. Parameter values for the recovery/spawning state were determined based on data from tagged fish known to be in the vicinity of the release location in March. Average parameter values from fish known to be in the release location in March were then applied to migratory fish during days assigned to the recovery/spawning period. Parameter values for the migration/foraging state were then determined for each migratory fish individually. The assumption of similar parameter values for all fish during the recovery/spawning state simplified the estimation process for migratory fish, because parameters were only estimated for one movement state. Diffusion values were estimated using maximum likelihood [19] for a range of off-bottom constants (0–250 m in 50 m increments). Off-bottom constants (and their associated maximum likelihood diffusion values) were then chosen based on examination of step-length histograms from daily location estimates and comparison of observed PSAT data to extracted values from geolocation maps at model-estimated locations to assess goodness of fit [39].

We made two modifications to the model framework to accommodate the topography of our study area, which consisted of a narrow island chain with potential pathways on both the north and south sides of the chain. The first modification involved changes to the movement model update. With a single movement model update per time step, where the movement kernel is convolved with the prior, large diffusion kernels can cross over narrow land masses resulting in movement probability on both sides of the island chain that does not correspond to the actual distance that the tagged animal could have traveled. To minimize this potential artifact, we updated the prior probability with the movement model in a series of n smaller updates before the data likelihood update was performed (Uffe Thygesen, Technical University of Denmark, pers. comm.). Each smaller update used $1/n$ of the nominal value of diffusion assigned to the movement state for that day, and probability on land was removed after each smaller update. Additional file 2 contains additional information about this method, which we refer to as the “expanding kernel movement model”, including advice on selecting the appropriate number of smaller updates for specific studies.

The second modification involved the way the most probable track is calculated. After the posterior probability surfaces have been obtained from the forward filter and backward smoothing, they can be used to estimate a point location for the tagged fish at each time step. The weighted mean of the posterior probability surface is a common local decoding method for estimating a point location based on the posterior probability surface. The weighted mean method works well in open ocean study areas, where posterior

probability surfaces are not multi-modal, but it tends to place locations on land in areas with convoluted coastlines or island topography (e.g., when probability is present on both sides of an island chain). The Viterbi algorithm [40], a global decoding method, provides an alternative to the weighted mean method that will not place estimated locations on land. The algorithm finds the sequence of grid cells that maximizes the product of transition probability multiplied by likelihood at each time step. This method for obtaining the most probable track for fish geolocation was introduced by Pedersen et al. [17] using the data likelihood model as the likelihood in the algorithm. Using the data likelihood model in the algorithm produced a most probable track that was not necessarily related to the posterior probability provided by the forward filter/backward smoothing [14] and could be prohibitively time-consuming. We modified this method using the posterior probability generated by the forward filter/backward smoothing as the likelihood in the Viterbi algorithm [41] instead of the data likelihood model. This constrained the algorithm to find a path through the posterior probability surfaces. The algorithm was further modified by truncating the posterior probability to 95% of the maximum posterior probability observed in the study area each day to reduce processing time (see Additional file 3 for processing time comparisons).

To visualize and interpret model results, we generated gridded probability surfaces, polygons, and point estimates. The residence distribution is a grid surface that depicts which study area grid cells were most likely to have been occupied by the tagged fish during a given period. It is produced by summing the posterior probability in each grid cell over a given time (e.g., a month, season, or the entire trajectory), normalizing by the sum of the total probability for all grid cells, and converting to a cumulative distribution function [14, 26]. We generated residence distributions for the entire trajectory. We visualized daily location estimates by creating polygons that encompassed the grid cells with the highest 50% and 99% of the posterior probability each day. Daily point estimates (most probable track) were generated using both the weighted mean and Viterbi approaches.

Although foraging and migration were combined in one movement state for the HMM, we were able to define migration and foraging periods in the reconstructed pathways using horizontal displacement from the release location over time. Horizontal displacement was calculated using Viterbi daily point estimates. Migration end dates were defined as the date when horizontal displacement from the release locations no longer increased.

Simulation study to assess model adjustments for island topography

To explore the effects of the modifications to the HMM for island topography, the expanding kernel movement model and the Viterbi method of calculating the most probable path, we simulated 100 random walk paths in a subset of the study area and reconstructed movement paths based on the simulated data. Simulated paths were designed to have a diffusion value of $D=50 \text{ km}^2/\text{step}$ (i.e., day), similar to the observed movement speeds of tagged cod in this study. To avoid simulated locations crossing over land masses, simulated paths were generated at smaller time scales and then thinned to obtain the final simulated path. Paths with 10,000 steps were generated using a value of $D=1 \text{ km}^2/\text{step}$. At each time step, a candidate future location was selected by choosing random step lengths in both the X and Y directions obtained from a normal distribution with a mean of 0 and a variance of $\sigma^2 = D \cdot \Delta t$. If the candidate location had a depth between 0 and 2000 m, it was accepted as the location of the next time step; based on depth records from archival tags, cod in this study were not observed to occupy depths deeper than 600 m, so 2000 m was chosen as a conservative threshold to represent locations that cod would not likely occupy. After all steps were simulated, the path was thinned to every 50th location, resulting in a simulated path with 200 steps. Simulated data sets derived from simulated pathways were designed to be similar to observed tag data sets. The maximum depth within each 50-step interval from the un-thinned path was assigned to each time step from the thinned path as the depth inputs for the geolocation model. Longitudes from 30% of the 200 steps were randomly chosen to provide longitude values for the model, and random error with a value of 1.5 degrees was added to each selected longitude.

To assess performance of the expanding kernel method, the simulated data were run in the HMM using the standard method (no “mini expansions”) and the expanding kernel method with 25 “mini expansions”. For each treatment, the most probable path was determined using (1) the weighted mean and (2) the Viterbi pathway (most probable sequence of grid cells occupied) through the posterior probability surface. Performance was assessed using two metrics: (1) the distance mean absolute error (MAE), which summarizes the distance between the known (simulated) and estimated location on each day, and (2) the depth root mean square error (RMSE) between depth at known (simulated) and model-estimated locations. Distance MAE reflects the accuracy of the position estimate, while the depth RMSE reflects the degree to which the estimated location is realistic based on comparison to model inputs. Distance MAE

and depth RMSE for standard versus expanding kernel methods were compared with a paired t test across all 100 simulated pathways for both the weighted mean and Viterbi pathway methods.

Results

Fish tags

Although 21 miniPATs were deployed in the pilot study, only 6 data sets provided enough data to conduct geolocation (Fig. 1, Table 1). Eleven tagged fish experienced tagging mortality in release areas, likely due to barotrauma from capture at depths > 100 m, and 4 tagged fish survived but did not transmit enough data to reconstruct movement paths [6]. Geolocated fish were at liberty for 21–277 days. Total lengths for geolocated fish ranged from 70 to 88 cm. Two tags were physically recovered in the commercial fishery, and thus provided complete, detailed data sets. The 11 post-release mortalities, along with stationary tag data, were used to define temperature depth profiles in release areas.

Stationary tags

Stationary tags were deployed at depths of 117 m (Adak Strait), 105 m (Sitkin Sound), and 110 m (Nazan Bay). All three stationary tags released from their moorings early (possibly due to attachment failure) and transmitted data. One of these tags (Sitkin Sound) had a battery malfunction and transmitted very few records. However, the remaining two tags provided critical information about the accuracy and precision of light-based latitude and longitude and temperature–depth profiles in the release area through November 2019.

Longitude accuracy was high from February through May (RMSE=0.56 degrees) but decreased in the summer months (June–November; RMSE=1.3 degrees). Longitude values were biased to the west from February through May (mean error – 0.22 degrees, $p=0.0016$) and from June through November (mean error – 0.43 degrees, $p=0.0075$). Conversely, latitude values were less accurate from February through May (RMSE=2.6 degrees) compared to June through November (RMSE=1.8 degrees). Latitude estimates were biased to the south June–November (mean error – 0.43 degrees, $p=0.05187$). In general, the accuracy of longitude estimates was much higher than latitude. Few latitude or longitude observations were obtained during May and early June (Fig. 2). We used a value of 1.5 degrees to specify the standard deviation of light-based longitude in the model for the entire period; a more conservative value than the calculated RMSE was chosen for the variance, because accuracy of light-based estimates tends to be slightly lower for free-swimming fish (which can occupy varying

depths during dusk and dawn) compared to stationary reference tags [12, 37].

Temperatures recorded by stationary tags had similar seasonal temperature trends but differed in magnitudes (Fig. 3). For all stations combined, temperatures at approximately 100-m depth ranged from 3.5 to 5.5 °C from February through May and 4 to 7 °C from June through September. Temperatures in Adak Strait and Sitkin Sound tended to be warmer than Nazan Bay.

Geolocation of tagged Pacific cod

Movement pathways were reconstructed for 6 fish. Diffusion value and off-bottom constants derived from tagged fish with end locations in their release areas in March were 5 km²/day and 50 m, respectively (Table 2); these values were then applied to the recovery/spawning state for migratory fish. For migratory fish, diffusion values during the migration/foraging period ranged from 50 to 75 km²/day, and off-bottom constants ranged from 50 to 150 m. Tag 178690 had the largest off-bottom constant due to its travel over waters as deep as 1000 m during migration. Buffered pop-up locations were used for tag 178697, which likely drifted prior to obtaining an Argos location, and tag 178704, which was preyed upon by a marine mammal (tag 178704 recorded temperatures to 37 °C, see Bryan et al.). In terms of model fit, the RMSE between observed and model-estimated depth and longitude ranged from 16 to 48 m and 0.6 to 1.1 degrees, respectively (Table 2).

Movement between winter spawning areas and summer foraging areas could be reconstructed for 4 fish, including one preyed upon by a marine mammal and one captured in the commercial fishery (tag 178709). Migratory fish displacement from release locations ranged from 60 to 395 km (Table 2). Migratory fish were estimated to spend an average of 23 days on the spawning grounds (range 16–34 days), where they were tagged (Table 2). During this period, temperatures recorded by tagged fish were generally within 0.2 °C of temperatures recorded by stationary tags and fish with tagging mortalities (Fig. 4). Migration initiation dates ranged from 3/12 to 3/27. Migration periods lasted an average of 22 days (range 6–43 days). Longitude values during the migration period clearly reflected movement east or west with gradual sequential changes, but during the foraging period, longitude values tended to have greater variability (Fig. 5).

Our modifications to the Viterbi method increased track quality and dramatically decreased processing time. Slight differences were observed between the Viterbi and weighted mean methods for calculating the most probable track for longitude (Fig. 5, 6, 7). However, the Viterbi method clearly produced more realistic estimates

Table 1 Information about tagged fish

Tag ID	Days prog	FL (cm)	Platform	Region	Rel Date	Rel Lat	Rel Long	Rel depth (m)	Rec Date	Rec Lat	Rec Long	Days at liberty	Fate
Geolocated miniPAT data sets													
178690	90	74	Trawl	Nazan Bay	20-02-2019	52.3548	-173.8582	115	23-05-2019	52.5473	-179.6535	91	Scheduled pop-up
178696	180	77	Pot	Sitkin Sound	22-02-2019	51.9238	-176.4302	110	28-03-2019	52.0579	-176.5112	32	Predation
178697	180	74	Pot	Sitkin Sound	22-02-2019	51.9225	-176.4302	93	23-08-2019	52.0770	-171.4197	181	Scheduled pop-up
178701*	270	70	Trawl	Nazan Bay	19-02-2019	52.2565	-173.5720	119	12-03-2019	52.4050	-173.8317	21	Fishing mortality
178704	360	73	Trawl	Nazan Bay	19-02-2019	52.2595	-173.5372	110	24-11-2019	52.5250	-172.5873	277	Predation
178709*	360	88	Pot	Sitkin Sound	22-02-2019	51.9235	-176.4273	98	08-06-2019	52.0900	-172.1230	106	Fishing mortality
Non-geolocated miniPAT data sets													
178689	90	100	Trawl	Nazan Bay	19-02-2019	52.2570	-173.5700	119	09-03-2019	52.3719	-173.1913	3	Tagging mortality
178691	90	104	Pot	Kagalaska Strait	01-03-2019	51.7071	-176.3632	102	31-03-2019	51.7367	-176.3958	2	Tagging mortality
178692	180	72	Trawl	Nazan Bay	18-02-2019	52.2533	-173.5413	119	03-04-2019	52.3596	-173.6625	2	Tagging mortality
178693	180	71	Trawl	Nazan Bay	19-02-2019	52.2563	-173.5730	119	04-03-2019	52.2931	-173.2733	3	Tagging mortality
178694	180	86	Trawl	Nazan Bay	20-02-2019	52.2593	-173.6217	110	29-03-2019	52.0516	-174.8574	7	Unknown early release
178695	180	81	Trawl	Nazan Bay	20-02-2019	52.3570	-173.8565	115	04-03-2019	52.4009	-173.6509	0	Tagging mortality
178698	180	73	Pot	Sitkin Sound	22-02-2019	51.9070	-176.4333	113	12-03-2019	51.9234	-176.3811	13	Tagging mortality
178699	180	83	Pot	Kagalaska Strait	01-03-2019	51.7108	-176.3530	91	26-03-2019	51.9445	-176.8127	18	Predation
178700	270	102	Pot	Kagalaska Strait	01-03-2019	51.7103	-176.3503	91	01-04-2019	51.7472	-176.5238	30	Unknown early release
178702	360	82	Trawl	Nazan Bay	20-02-2019	52.3562	-173.8570	115	13-03-2019	52.2938	-173.6215	4	Tagging mortality
178703	360	98	Trawl	Nazan Bay	20-02-2019	52.3555	-173.8573	115	08-03-2019	52.4879	-173.3722	0	Tagging mortality
178705	360	95	Trawl	Nazan Bay	18-02-2019	52.2537	-173.5430	110	09-03-2019	52.4365	-173.2086	0	Tagging mortality
178706	360	89	Pot	Kagalaska Strait	01-03-2019	51.7077	-176.3598	100	26-03-2019	51.7721	-176.1638	5	Tagging mortality
178707	360	96	Pot	Kagalaska Strait	01-03-2019	51.7085	-176.3542	95	18-03-2019	51.7128	-176.4995	3	Tagging mortality
178708	360	79	Pot	Sitkin Sound	22-02-2019	51.9233	-176.4318	99	19-02-2020	51.7147	-176.9818	362	Scheduled pop-up

*Denotes detailed data sets obtained from physically recovered tags

Number of days programmed, fork length (FL, cm), release platform, release region, release depth, release and recovery dates and locations, days the tagged fish was alive and at liberty, and fate of tagged fish

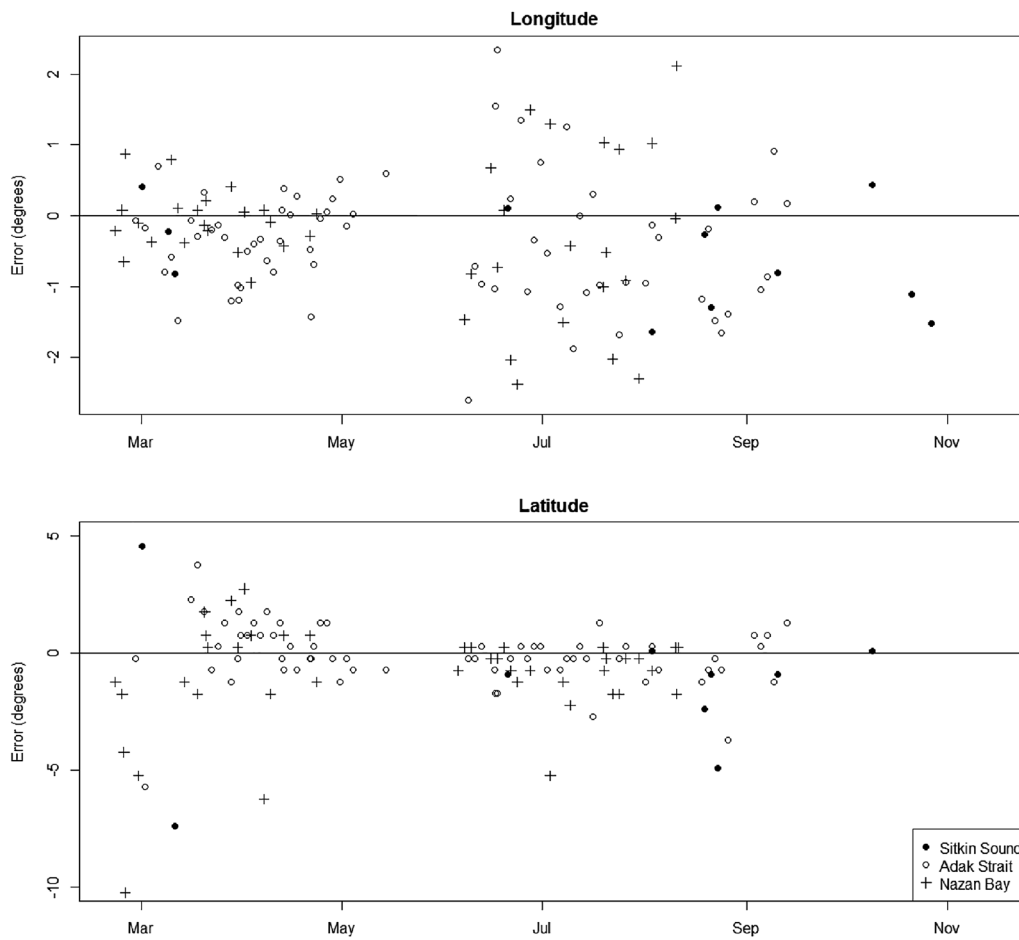


Fig. 2 Error (degrees) in daily light-based longitude and latitude estimates from three stationary tags in the Aleutian Islands at depths of 105 m (Sitkin Sound), 117 m (Adak Strait), 110 m (Nazan Bay). Note that the scale of the error (Y axis) differs between latitude and longitude plots

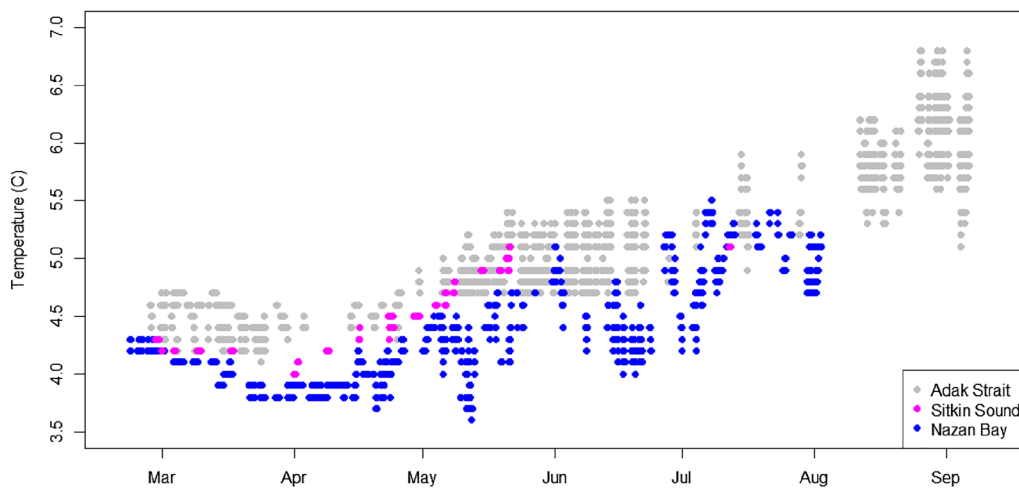


Fig. 3 Stationary tag temperature data sets (10-min interval) at depths of approximately 100 m from Adak Strait (gray), Sitkin Sound (pink), and Nazan Bay (blue) during 2019

Table 2 Geolocation information

TagID	Rel date	Migr start	Migr end	Pop-up	Dist (km)	D (R/S) (km ² /day)	D (M/F) (km ² /day)	OC (R/S) (m)	OC (M/F) (m)	RMSE depth (m)	RMSE long (deg)
178701	2/19/19	NA	NA	3/12/19	24	5	NA	50	NA	16	0.6
178696	2/22/19	NA	NA	3/28/19	16	5	NA	50	NA	25	0.6
178690	2/21/19	3/16/19	4/28/19	5/23/19	395	5	75	50	150	48	1.1
178697	2/22/19	3/16/19	4/2/19	8/23/19	316	5	50	50	50	35	1.0
178704	2/20/19	3/27/19	4/2/19	11/22/19	60	5	50	50	50	29	0.8
178709	2/23/19	3/12/19	4/4/19	6/8/19	298	5	75	50	50	24	0.6

Release date, migration timing, pop-up date, and net distance traveled (km). Geolocation coefficients: diffusion (D, km²/day), off-bottom coefficient (OC, m) used during the recovery/spawning (R/S) movement state and migration/foraging (M/F) movement states, and root mean square error (RMSE) of model residuals (e.g., difference between observed and model-estimated values) for depth (m) and light-based longitude (degrees) based on Viterbi pathways

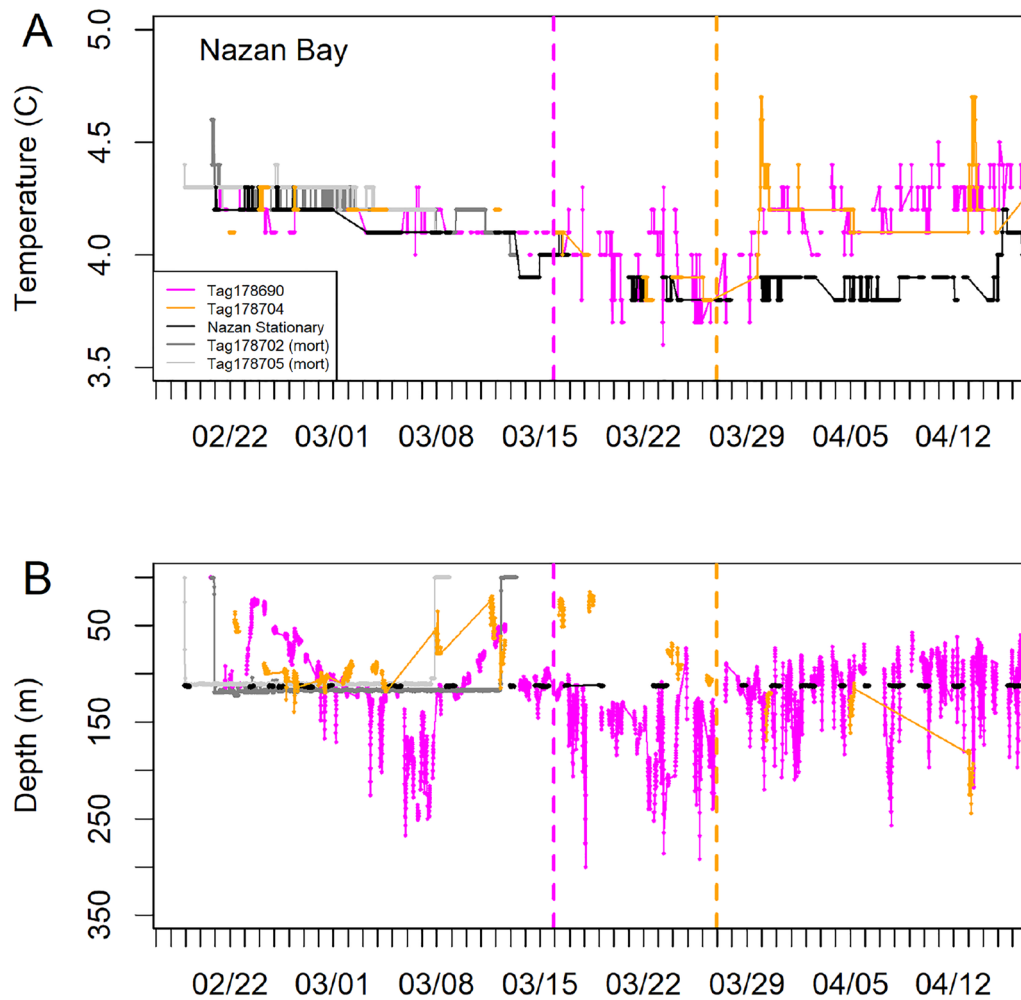


Fig. 4 Example of using temperature–depth profile data to determine onset of Pacific cod migration from release area (Nazan Bay). Temperature observations **A** for geolocated fish (Tag 178690, Tag 178704) match the stationary tag data and two fish (Tag 178702 and Tag 178705) that experienced post-tagging mortality in the release area until likely dates of departure (dashed vertical lines, magenta = Tag 178690, orange = Tag 178704). Temperature observations from geolocated fish **B** were similar (i.e., within 0.2 °C) to stationary reference tags despite depth changes of more than 100 m, while fish likely occupied release areas

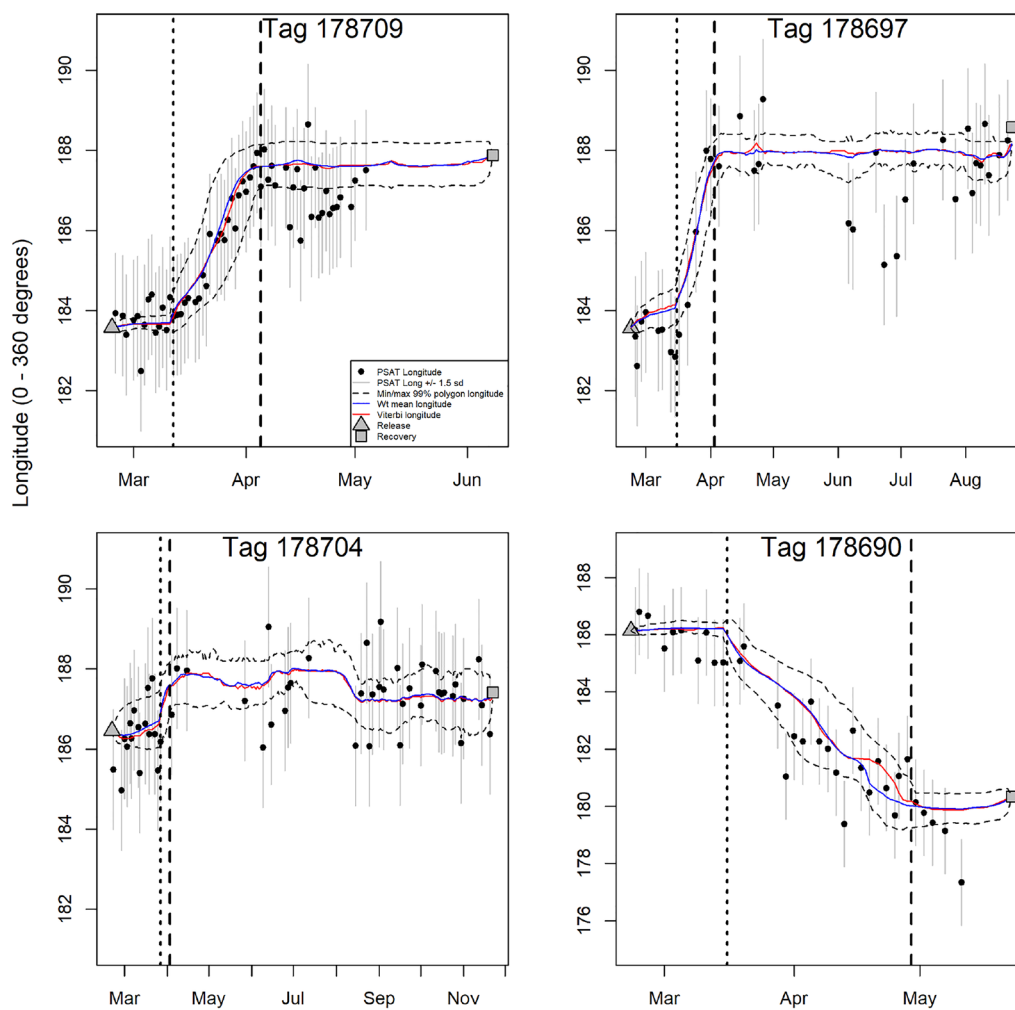


Fig. 5 Longitude values for geolocated Pacific cod. Raw longitude observations (black points) \pm the standard deviation of longitude used in the hidden Markov model (gray vertical lines). Gray triangles and squares indicate release and recovery locations, respectively. Minimum and maximum longitudes from daily 99% location probability polygons are indicated by black dashed lines. Longitude values from reconstructed pathways are shown for the weighted mean (red line) and Viterbi (blue line) methods. Dotted vertical lines indicate the beginning of the migration phase, and dashed vertical lines indicate the beginning of the foraging phase. Note that axes differ across panels

when comparing observed depth to depth at estimated locations (Fig. 8). Therefore, estimates from the Viterbi method were used as point estimates for reconstructed pathways for geolocated fish in subsequent analyses. Viterbi algorithm processing time was reduced from hours to seconds (Additional file 3) after implementing the 95% cumulative probability threshold in the algorithm.

Simulation study to assess model adjustments for island topography

Model performance was dramatically improved for some nearshore simulated pathways using the expanding kernel method (Fig. 9, Additional File 2 Fig. S2.4). For

example, a large proportion of the residential distribution (and thus Viterbi and weighted mean estimated pathway locations) for Path 41 obtained from the standard kernel method was located on the north-eastern side of the island, where no known (simulated) locations occurred, and no probability was present in the north-western side of the islands, where known locations did occur (Fig. 9A). In contrast, the residential distribution from the expanding kernel method is much better aligned with the simulated pathway (Fig. 9B).

Additional insights into the differences between the two methods are provided by examining the daily probability estimates. For example, probability on day 2 using the standard movement kernel (Fig. 10A) crossed

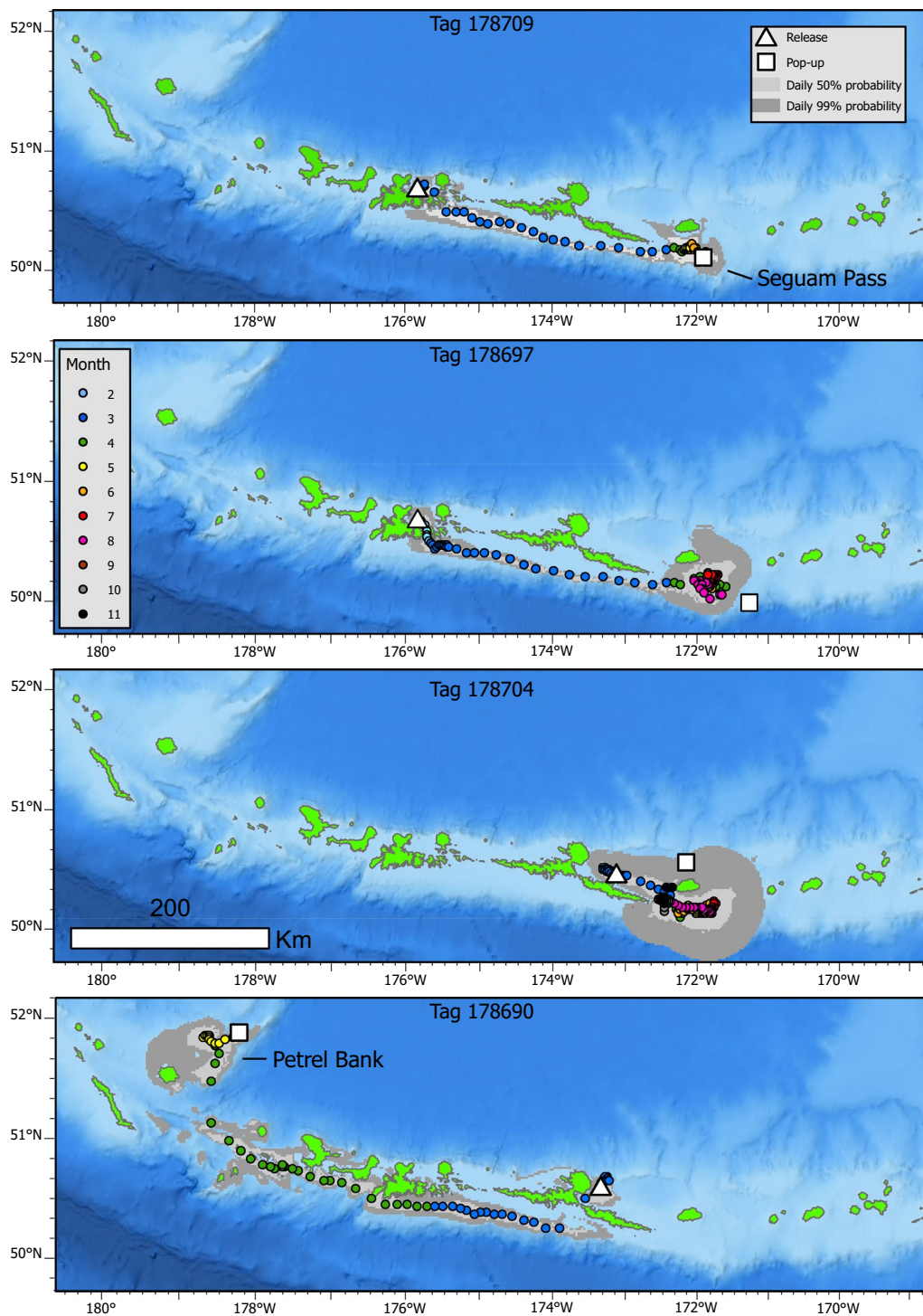


Fig. 6 Daily probability polygons and estimated point locations (Viterbi method) for geolocated Pacific cod. Light gray and dark gray polygons represent daily 50% and 99% location probability polygons. Viterbi daily locations are color-coded by month. Release locations are indicated by a white triangle and recovery locations with a white square

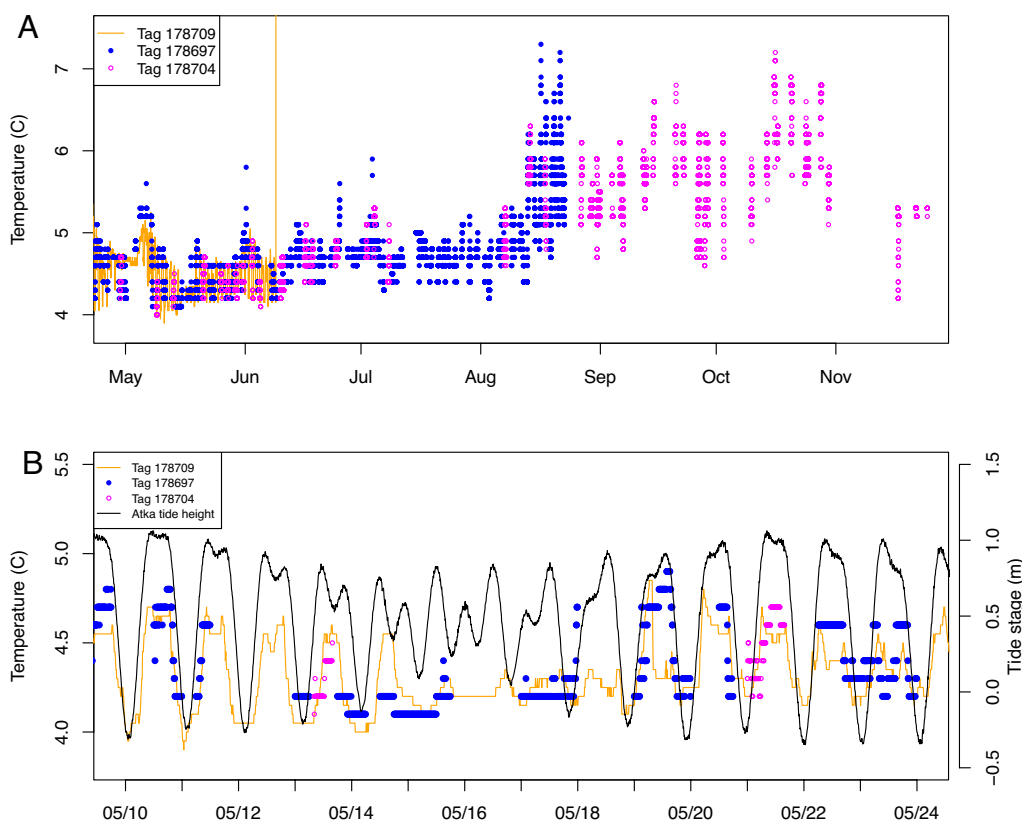


Fig. 7 Comparison of temperature records for all three geolocated fish that likely occupied Seguam pass during summer foraging. **A** Temperature records from May through November (detailed temperature records for Tag 178709, which was captured in the commercial fishery, orange line, and transmitted temperature records for Tag 178697 and 178704, blue and pink circles, respectively). **B** Temperature records during a 2-week period in May, where tagged fish temperatures are shown together with tide stage data from the nearby Atka station (Station 9,461,710, B) at 6-min intervals. The times of high and low tides at Seguam Pass (Finch Cove) are approximately -50 min and -7 min from the Atka station, respectively

land and placed the location of the Viterbi path on the north-eastern side of the island, and the weighted mean path location was on land between the two high probability areas. In contrast, smoothed probability on day 2 using the expanding kernel method (Fig. 10B) did not cross land and was concentrated near the known location. Both Viterbi and weighted mean locations were near the known (simulated) location for that day. On day 100 (the simulated trajectory mid-point), most of the smoothed probability from the standard method, along with the Viterbi and weighted mean path locations, was located on the north-eastern side of the island instead of the true (simulated) location on the south side of the island (Fig. 10C). However, when the expanding kernel was used, all of the probability was located on the south side of the island, and the Viterbi and weighted mean path locations were both close to the known (simulated) location (Fig. 10D).

Despite marked improvements for certain simulated trajectories, no differences in performance between the standard and the expanding kernel method were found

across all 100 simulated data sets. No significant difference in the distance mean absolute error (MAE) between standard and expanding kernel treatments was observed for either the weighted mean or Viterbi pathway methods (Table 3). Of the simulated pathways that were improved by the expanding kernel method, distance MAE was lowered by an average of 4.2 km for the weighted mean method ($n=37$, range 0.001–31.5 km) and 4.5 km for the Viterbi method ($n=42$, range 0.001–34.5 km). In cases where the standard kernel performed better, distance MAE was lowered by an average of 2.1 km (range 0.009–14.3 km) for weighted mean pathways and 2.5 km (range 0.003–17.3 km) for Viterbi pathways.

Across all simulations, distance MAE for weighted mean paths was significantly lower than Viterbi for both the standard ($p=0.0002$) and the expanding kernel ($p=0.0038$) treatments, although the magnitude of improvement was small (1.8 km for the standard kernel and 1.5 km for the expanding kernel). This phenomenon was also observed across a number of mini expansions (Additional file 2 Fig.S2.3). This result likely reflects the

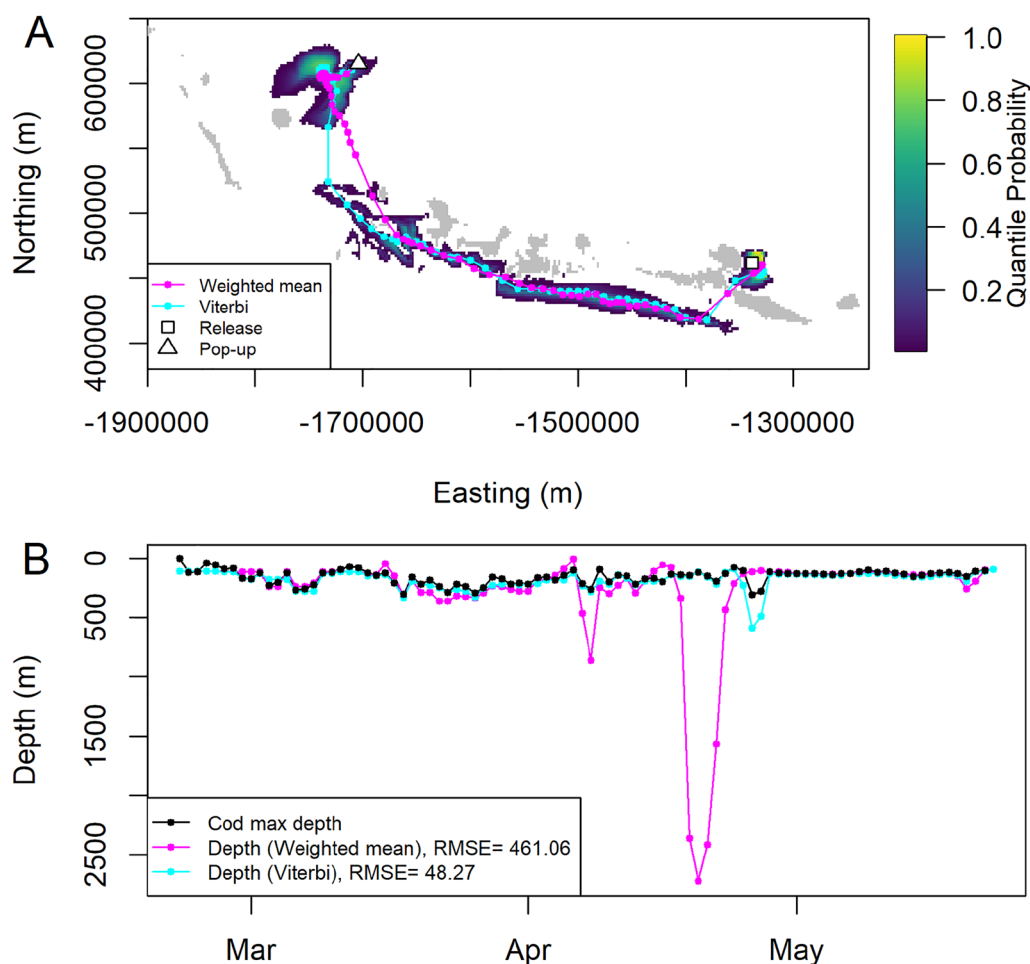


Fig. 8 Comparison of weighted mean and Viterbi methods for estimating the most probable track (daily point estimates) of a tagged Pacific cod. **A** Residence distribution (surface color-coded by quantile probability) for tag 178690 with weighted mean (pink) and Viterbi (blue) pathways (land indicated by gray areas). **B** Maximum daily depth observed for the tagged fish (black) compared to depth at weighted mean (pink) and Viterbi (blue) locations. Root mean square error (RMSE, m) between observed and estimated depths for each method shown in the legend, is higher for locations estimated using the weighted mean method, because the weighted mean of the probability surface falls between two high-probability areas in deep water regions, whereas the Viterbi method is constrained to high-probability grid cells in shallower waters that more closely match the observed depths recorded by the fish

effect of grid size and the method of calculating each pathway (the weighted mean is a continuous average of latitude and longitude, whereas the Viterbi location is taken as the center of the most probable grid cell).

In addition, the two pathway reconstruction methods differed significantly in their ability to produce realistic locations. Viterbi paths out-performed the weighted mean method of calculating point locations in terms of depth RMSE at model-predicted versus known (simulated) locations ($p < 2.2 \times 10^{-16}$) for both standard and expanding kernel weighted mean versus Viterbi comparisons (Table 3). This reflects the adherence of the Viterbi pathway to the highest probabilities from the smoothed posterior probability surface, whereas weighted mean

locations can be located between posterior probability modes in either deeper water or land (Fig. 10).

Discussion

This study is the first to deploy PSATs on Pacific cod. It demonstrates that it is possible to reconstruct plausible movement of tagged Pacific cod from winter spawning areas to summer foraging areas using PSAT data. Mortality of tagged animals following capture was high and likely due to barotrauma of fish captured at deep locations [6]. However, with refined capture and release methods, PSATs can provide important details about Pacific cod post-spawning migrations in the Aleutian Islands in future studies.

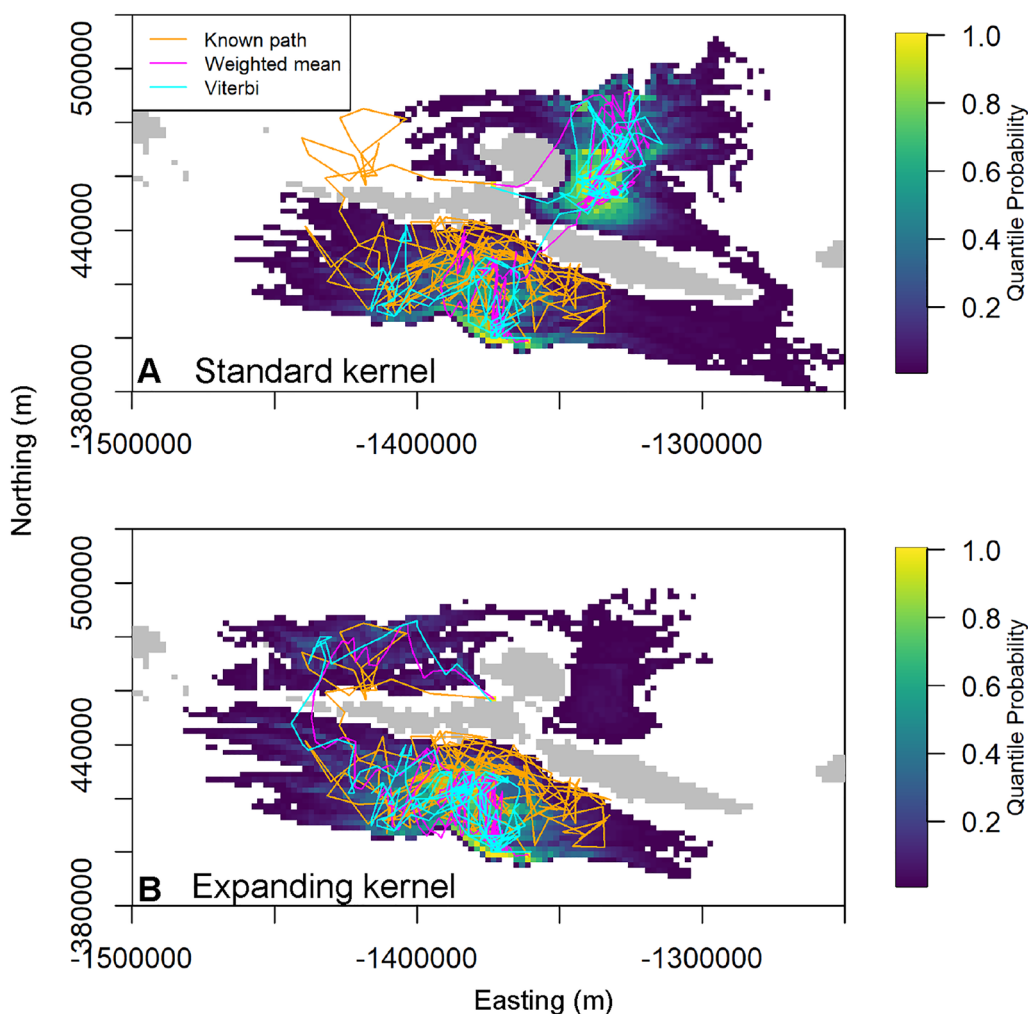


Fig. 9 Residential distributions (color-coded surfaces) resulting from geolocation of simulated pathway (Path 41) using standard (A) and expanding (B) movement kernels. Lines indicate simulated (orange), weighted mean (pink) and Viterbi (blue) pathways

Prior to this study, information on seasonal movement of cod in the Aleutian Islands was obtained from recapture locations of conventional and archival-tagged fish only [3, 5]. The distance and directions traveled by migrating fish in this study are similar to results from previous studies. However, reconstructed pathways from PSAT data have provided the first information on migration timing (i.e., departure from spawning areas and arrival in summer foraging areas) and possible pathways. Thus, the use of PSATs appears to be a promising method for providing insights into the nature of seasonal migration of Pacific cod in Alaska.

Geolocation of Pacific cod

Light-based longitude was the most important geolocation variable for reconstructing cod migratory pathways in the Aleutian Islands, as it clearly detected east–west movement during the winter and early spring. Accuracy

of longitude estimates from the stationary test tags during this period (RMSE = 0.56 degrees) was comparable to stationary test tags at lower latitudes [12] despite being placed at water depths of 100 m or more. This result may be due to excellent water clarity in the study area during the winter and early spring, as this region does not have glacial inputs that can impact water clarity as in other regions of Alaska. The lack of longitude estimates during May for both stationary tags and tagged fish could be caused by decreased water clarity during the spring phytoplankton bloom. The mean bloom timing in the Aleutian Islands between 1998 and 2006 was April 28 (± 24 days) with an average duration of 39 ± 25 days [42]. This timing matches the approximate gap in light-based data for both stationary tags and fish. During summer, primary productivity has been observed to be low in the Aleutian Island passes as the Alaska Coastal Current brings depleted water from the Gulf of Alaska to the

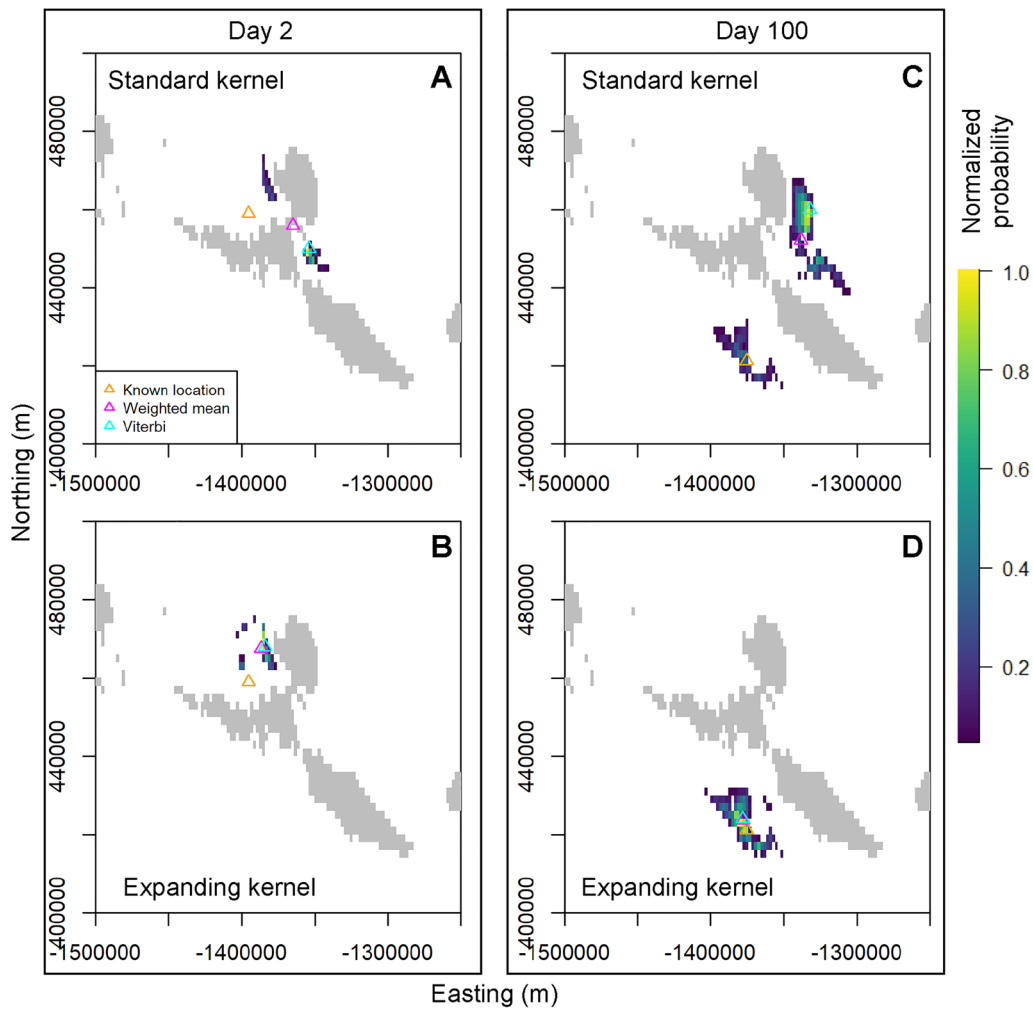


Fig. 10 Comparison of standard kernel (A, C) and expanding kernel (B, D) daily probability surfaces for simulated pathway #41 on days 2 and 100 (mid-trajectory). Smoothed probability surfaces are normalized to the maximum value for each day and truncated at normalized values of 0.05. Land is indicated by gray areas. Triangles indicate known (simulated) locations (orange), weighted mean pathway locations (pink), and Viterbi pathway locations (blue)

Table 3 Distance mean absolute error (MAE) of estimated locations and depth root mean square error (RMSE) for 4 geolocation treatments of 100 simulated data sets: (1) Standard movement kernel (no mini expansions) and weighted mean most probable track, (2) Standard movement kernel and Viterbi most probable track, (3) Expanding movement kernel (25 mini expansions) and weighted mean most probable track, (4) Expanding movement kernel and Viterbi most probable track

Treatment	Mean distance MAE (km)	Range distance MAE (km)	Mean depth RMSE (m)	Range depth RMSE (m)
Standard/Wt. mean	37.5	18.1–82.3	335	122–883
Standard/Viterbi	39.3	19.1–85.9	45	10–156
Expanding/Wt. mean	37.3	17.2–83.3	337	120–858
Expanding/Viterbi	38.7	17.1–86.4	45	10–157

area [43]. Thus, water clarity conditions may be generally favorable for light-based geolocation in Aleutian waters aside from the spring plankton bloom in May and June.

Reduced accuracy of the stationary tag longitude estimates (RMSE = 1.3 degrees) and more variable longitude estimates from tagged fish (Fig. 4) during the summer months suggest that geolocation results will be less precise for Pacific cod during the summer foraging phase compared to spawning or spring migrations from spawning areas. In addition to possible decreases in water clarity during the summer, increased variability in light-based longitude among tagged fish during summer could also be caused by increased depths occupied during summer foraging or seasonal change in depth at dawn and dusk [5]. However, despite the slight decrease in accuracy of longitude observations, the geolocation model was still able to identify distinct areas that fish were likely to occupy during the summer, such as Seguam Pass.

In addition to longitude, the stationary tag data provided reasonable light-based latitude estimates. We did not include latitude in the data likelihood model for Pacific cod, because the orientation of depth gradients in the study area effectively provides more precise information on latitude, but the information on latitude precision from the stationary tags at depth is potentially valuable for other high-latitude applications. The increase in stationary tag precision in summer compared to winter/early spring could be caused by the well-known difficulties in estimating latitude during spring and fall equinoxes [13, 44, 45]. These results are encouraging for the use of light-based latitude in study areas with similar water quality, where depth gradients do not approximate latitude; however, as for longitude, seasonal changes in the vertical distribution of some fish species are likely to decrease the precision of latitude estimates compared to stationary tags.

Although depth is typically one of the most important geolocation variables for demersal fish in the North Pacific Ocean [19], we found that this was not necessarily the case in the Aleutian Islands study area. As mentioned above, light-based longitude is the primary variable for detecting migrations (east–west movement) in the Aleutian Islands region, while depth provides information about latitude. The steepness of the depth gradients in the area means that even large differences in depth (e.g., 100 m) correspond to very small changes in latitude. It is possible that the assumption of some degree of off-bottom behavior (e.g., the off-bottom coefficient used in the maximum depth likelihood) could bias the location probability toward deeper water if the maximum depth of the tagged fish is actually recorded, while the fish is on the seafloor, but the steepness of the depth gradients acts to minimize the effect of this potential bias on geolocation

estimates in this study. In areas with shallow depth gradients, great care should be taken in the use and selection of the off-bottom coefficient as a potential bias toward deeper waters would affect the geolocation estimates more than in steep depth gradient areas. This points to the need to identify and understand off-bottom behavior from the standpoint of geolocation as well as other purposes (i.e., assessing vulnerability to capture with different gear types).

Another potential source of bias from the model stems from the way the depth likelihood is calculated and the presence of different depth gradients on the north and the south sides of the island chain. Given the same study area grid cell mean, the maximum depth likelihood assigns a higher likelihood to grid cells with a smaller variance for a grid cell depth that is similar to depth measured by the tagged fish. Therefore, grid cells that contain the depth of the fish measured by the PSAT but have steep depth gradients receive much lower likelihood values than grid cells with very shallow gradients at the same depth. Because depth gradients are shallower on the south side of the chain (i.e., the variance of depth in each grid cell is smaller), this could drive the increased location probability to the south side of the island chain (where most of the location probability for 3 of the 4 geolocated fish was observed during migration). If fish swim along steep gradients along the north side of the island chain during migration, this might not be reflected in the data likelihood model. For one tagged cod (tag 178690) that migrated west to Petrel Bank, movement on the south side of the island chain would be energetically beneficial, since the prevailing Alaska Coastal Current and Alaskan Stream flow from east to west [46]. However, the rapid eastward movement south of the island chain as estimated for two other cod would have been hindered by the prevailing current. An eastward migration on the north of the islands would be energetically more efficient as the Aleutian North Slope current runs east to west on the north side of the island chain beginning at Amchitka Pass [46].

Adding water column temperature data to the data likelihood model [26] may allow additional insights into which side of the island chains are used during migration. In general, water temperatures on the Bering Sea side of the chain will be colder than the Gulf of Alaska side. This effect is observed with changing of water temperature by at least 1 degree C between high and low tides in the data sets from the three cod that occupied Seguam Pass during summer foraging. We investigated using two temperature–depth profile maps that could potentially be included in the data likelihood model (Additional file 1). The HYCOM model [47] is a global ocean circulation model that has

been used for temperature–depth geolocation likelihoods for other fish species for which geolocation is challenging, including Atlantic cod [48], basking shark [49], swordfish [50], and Mediterranean spearfish [51]. Based on the data from the three stationary tags, the HYCOM matches observed temperature-at-depth during winter and early spring, when migration occurs. However, the spatial accuracy and resolution of the global HYCOM model depth bins were inadequate for the nearshore conditions in our study area. For example, HYCOM depth bins in nearshore areas tended to be much shallower than actual depths due to the coarse bathymetry used to generate them, whereas the finer-scale bathymetry used for the depth-based likelihood contained deeper depths that matched depths measured by the fish. Because likelihoods for each geolocation variable are multiplied to obtain the overall likelihood, this resulted in a negation of high likelihood values in the nearshore region.

We also compared estimated temperatures from the Bering Sea 10 K ROMS model [52], which primarily covers the Bering Sea region, to the stationary tag data. Although the spatial accuracy of depth bins for this model was better than HYCOM in nearshore areas, the estimated temperatures were much colder than observed temperatures year-round (Additional file 1, Fig. S1.1). The Aleutian Islands are close to the edge of the Bering Sea ROMS 10 K model and are thus not likely to be accurate for the study area (Kelly Kearney, NOAA, pers. comm.). Blended HYCOM models [53] or other hydrographic models (Seth Danielson, University of Alaska Fairbanks, pers. comm.) may be available in the future that could provide better accuracy in nearshore conditions within our study area.

This study has revealed that cod in the Aleutian Islands have a large range of movement in the water column during all migration phases. This behavior means a temperature–depth profile likelihood could be a valuable addition for the data likelihood model if accurate maps become available. However, temperatures in major passes, such as Seguam, where temperatures can differ by more than 1 °C based on tide stage, will likely be challenging to map accurately. Currently, the temperature–depth profile data are primarily used for determining residence in release areas (where temperature–depth profiles from tagged fish match temperature from stationary tags and tagging mortalities in the release area) and broad indications of movement (i.e., rapidly changing temperature-at-depth profiles or distinct changes in temperature–depth profiles over time can indicate movement and thus aid the assignment of movement states in the HMM).

Adjustments for island chain topography

We made two adjustments to HMM geolocation methods to improve performance in nearshore areas. The first adjustment, the expanding kernel method used in the forward filter/backward smoothing, improves the quality of the posterior probability surfaces resulting from the HMM in nearshore areas. When narrow landmasses such as islands or peninsulas are present in the study area, the diffusion kernel used to represent the movement of the tagged fish may cross them when the standard method (convolution of the prior with one update by the movement kernel followed by the data likelihood update) is used. This issue was addressed by Liu et al. [54] by calculating the diffusion kernel using distance between grid cells adjacent only to water. We used an alternative approach that is likely simpler to implement: a series of incremental expansions of the movement model followed by eliminating probability on land after each incremental expansion [17, Uffe Thygesen, DTU, pers. comm.]. This method does not appreciably increase processing time to run the HMM; however, it effectively prevents movement probability from crossing land barriers.

Although the simulation study did not identify any quantitative improvements in distance between known and estimated locations (e.g., reduced distance MAE) across all simulated pathways using the method, posterior probability surfaces from some nearshore pathways were significantly improved. The lack of significant differences between the two methods is likely due to issues with the simulated pathways. First, approximately half of the simulated pathways were mostly offshore and thus would not be improved by the expanding kernel method. Second, geolocation error added to longitude in the simulated data sets could have a much larger effect on model estimates than refinements in the movement kernel.

For the trajectories that were improved, model accuracy increased (e.g., distance MAE was reduced) with increasing numbers of mini expansions up to a certain point, after which accuracy decreased with increasing numbers of mini expansions (Additional file 2). In terms of guidelines for choosing the optimal number of mini expansions, the simulations revealed a potential upper limit for the number of expansions that can be used (described in detail in Additional file 2). However, for specific applications, researchers should endeavor to determine the smallest number of mini expansions needed for the study area (e.g., consider the thickness of land barriers relative to cell size and movement speed of the animal).

The second adjustment, the Viterbi method for determining the most probable track improves the ability to estimate daily point locations from posterior probability surfaces. The most probable track may be used

to represent reconstructed pathways in graphics, to test the fit of models by comparing observed depth or temperature data to conditions at estimated locations, or to conduct movement analyses based on point locations. For study areas, where the posterior probability surface is confined to a single high-probability area and is unobstructed by land (e.g., in pelagic systems), the weighted mean method works well to produce a point estimate corresponding to the highest probability from the posterior probability surfaces. In addition, our simulations also revealed slightly lower distance MAE for weighted mean compared to the Viterbi method, likely due to the continuous nature of the weighted mean latitude and longitude estimates compared to the grid cell center provided by the Viterbi method. Therefore, in such situations the use of the weighted mean method may be preferable to the Viterbi method.

However, in study areas with multi-modal probability surfaces (such as probability on either side of an island chain), the weighted mean method tends to place the point estimate in low-probability areas that lie between multiple high-probability areas (such as land).

When multiple high-probability areas exist for a given time, we found that the Viterbi method, which finds the most probable sequence of grid cells occupied, was far more likely to provide point estimates consistent with high-probability areas from the posterior probability surface compared to the weighted mean method. This was observed for reconstructed Pacific cod trajectories (Fig. 8) and pathways reconstructed with simulated data (Fig. 9). The significant reduction in depth RMSE for locations estimated by the HMM compared to true (simulated) locations also indicates that the Viterbi method results in more realistic point location estimates for tagged fish.

The key to the Viterbi approach described here is to use the posterior probability estimates that result from the forward/backward filter as the likelihood input in the algorithm instead of the data likelihood values used in the forward/backward filter. In this way, the Viterbi algorithm controls which paths are admissible and globally decodes the most probable path through the posterior probability estimates [41]. This modification results in point locations that are always placed in high-probability areas and correspond directly to error estimates generated by the posterior probability surfaces.

Although substituting the posterior probability from the HMM for the data likelihood model saved some processing time (Additional file 3), the amount of processing time was still prohibitive (on the order of hours). We found that truncating the posterior probability surface to the top 95% of probability values each day greatly decreased processing time (on the order of seconds

to minutes) and made the method feasible for use with model diagnostics and other products associated with the geolocation process.

The adjustments for island chain topography are complementary when used together in near-shore applications, resulting in better probability surfaces and more realistic pathways that do not cross land. However, the Viterbi improvements will also be valuable for finding the most probable path for any application, where multiple areas of high probability are present. R code for both the expanding kernel and Viterbi algorithm will be available in a future version of the open source geolocation software HMMocean [26].

Conclusions

This pilot study has demonstrated the value of deploying PSATs that provide archival data for understanding seasonal movement of Pacific cod in the Aleutian Islands. We were able to use the archived data to reconstruct migration pathways between winter spawning and summer foraging areas. In addition to providing preliminary insights into the movement and behavior of Pacific cod in the Aleutian Islands, this work provides an example of adapting a geolocation model for a specific application that provides valuable advances for the broader community of fisheries researchers. Analysis methods developed here will benefit future research on the seasonal movement and behavior of Pacific cod and other demersal fish species in the region and any fish species occupying nearshore areas with convoluted coastlines or island chain topography.

Abbreviations

D	Diffusion
HMM	Hidden Markov model
MAE	Mean absolute error
PSAT	Pop-up satellite archival transmitting tag
RMSE	Root mean square error

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40317-023-00340-3>.

Additional file 1: Figure S1.1. Comparison of stationary tag temperatures to Bering 10Km ROMS model and HYCOM model temperature estimates in the Aleutian Islands.

Additional file 2: Detailed information about the expanding kernel method and guidelines for its use. Figure S2.1. The movement model update (movement kernel after mini-expansions) for Day 2 using different numbers of mini-expansions ("exp"). Release location (Day 1) is indicated by orange triangle. Color coded surface indicates the top 99% of quantile probability for each treatment. Land is indicated by grey grid cells. **Figure S2.2.** Example of HMM summary output (residence distributions) using the standard method (no mini-expansions) for 4 simulated pathways. Color-coded surfaces indicate location probability summarized across the entire simulated data set (top 99% of the quantile probability). Orange lines indicate simulated pathways. Paths 41 and 15 (top panels) have

simulated locations close to islands; thus, location estimation may be improved using the expanding kernel method. Paths 7 and 12 (bottom panels) are located mostly offshore and thus are unlikely to be improved by the expanding kernel method. **Figure S2.3.** Mean absolute error (MAE) between estimated and known (simulated) locations for each example pathway shown in Figure 2.2. Weighted mean pathways are solid lines; dashed lines are Viterbi pathways. Treatments consist of 1 (standard method), 10, 25, 50, 100, and 200 updates. **Figure S2.4.** Example of residence distributions (e.g., summary of the entire pathway) obtained using different numbers of mini-expansions. Simulated path (#41) is indicated by orange lines in the top left panel. Color-coded surfaces indicate the top 99% of quantile probability for each expansion treatment.

Additional file 3: Additional information about the Viterbi algorithm modification and processing time comparisons. Figure S3.1. Study area cropped to the vicinity of the residency distribution for tag 178690 (see Table 2 in main text for information on geolocation parameters).

Figure S3.2. Example of three different types of input into the Viterbi algorithm for day 46 of the 92-day trajectory: A) unmodified (data likelihood model), B) modified to use posterior probability instead of data likelihood model, and C) truncated at highest 95% of posterior probability (probability values are normalized by the maximum value). Grid cells on land are indicated in black, and grid cells with values of zero are indicated in white.

Acknowledgements

In addition to the funding groups described in the funding section, we thank John Gauvin for organizing the contributions from the fishing companies and for general support of the project. We thank David Fraser (Adak Community Development Corporation), Jerry Downing (B&N Fisheries Company), Captain Dan Carney (FV *Ocean Explorer*) and Captain Todd Hoppe (FV *Deliverance*) and crews for contributing invaluable knowledge and advice, providing tagging platforms, and assisting with Pacific cod tagging and release. MCA was supported by the Postdoctoral Scholar Program at Woods Hole Oceanographic Institution with funding provided by the Dr. George D. Grice Postdoctoral Scholarship Fund. CDB was supported by the Independent Research and Development Program at Woods Hole Oceanographic Institution. We thank Katy Echave, Kevin Siwicke, Ned Laman, Jim Lee, and Michael Martin from the NOAA NMFS Alaska Fisheries Science Center for reviewing the manuscript and providing valuable feedback.

Author contributions

SM, DB, KR, and JN participated in the design of the tagging study for Pacific cod. SM and DB conducted fieldwork and tagging. JN performed geolocation of Pacific cod, analysis of simulated pathways, and produced an initial draft of the manuscript. JN, MA, BG, and CB contributed to the modifications of the HMM introduced in the manuscript. All authors assisted in writing, review, and editing and have read and approved the final manuscript.

Funding

Funding for this research was provided by the North Pacific Research Board (Project # 1810), which played no role in study design, data collection, analyses, or interpretation of results. A group of eight fishing companies (Aleutian Spray Fisheries, Adak Community Development Corporation, United States Seafoods, O'Hara Corporation, Ocean Peace Inc., B&N Fisheries Company, Golden Harvest Alaska Seafoods, American Seafood Corp) also contributed funds for purchasing satellite tags. The fishing companies provided assistance with study design and logistical support for fieldwork but did not participate in data collection, analyses, or interpretation of results.

Availability of data and materials

The data sets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

NMFS Animal Care and Use Policy (04-112) is currently limited to research on free-living marine mammals, seabirds, and sea turtles and does not cover

research on captive or wild fish. However, every effort was made to follow accepted standards and ensure the ethical treatment of captured fish including guidelines from the U.S. Government Principles for the Utilization and Care of Vertebrate Animals Used in Testing, Research and Training (<https://olaw.nih.gov/policies-laws/gov-principles.htm>) and the American Fisheries Society Guidelines for the Use of Fishes in Research (https://fisheries.org/docs/policy_useoffishes.pdf; Chapter V).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Kingfisher Marine Research, LLC, 1102 Wee Burn Dr, Juneau, AK 99801, USA. ²National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, 7600 Sand Point Way Northeast, Seattle, WA 98115, USA. ³Under Contract to National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center, Lynker, 202 Church Street SE #536, Leesburg, VA 20175, USA. ⁴Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. ⁵National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Greater Atlantic Regional Office, 55 Great Republic Dr, Gloucester, MA 01930, USA.

Received: 9 February 2023 Accepted: 12 July 2023

Published online: 26 July 2023

References

- Barbeaux S, Ferriss B, Laurel B, Litzow M, McDermott S, Nielsen J, et al. Assessment of the Pacific cod stock in the Gulf of Alaska. In: Plan team for groundfish fisheries of the Gulf of Alaska (compiler), Stock Assessment and Fishery Evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fishery Management Council; 2021.
- Stevenson DE, Lauth RR. Bottom trawl surveys in the northern Bering Sea indicate recent shifts in the distribution of marine species. *Polar Biol.* 2018;42:407–21.
- Shimada A, Kimura D. Seasonal movements of Pacific cod, *Gadus macrocephalus*, in the eastern Bering Sea and adjacent waters based on tag-recapture data. *Fish Bull.* 1994;92:800–16.
- Rand KM, Munro P, Neidetcher SK, Nichol DG. Observations of seasonal movement from a single tag release group of Pacific Cod in the eastern Bering Sea. *Mar Coast Fish.* 2014;6(1):287–96.
- Nichol DG, Kotwicki S, Zimmermann M. Diel vertical migration of adult Pacific cod *Gadus macrocephalus* in Alaska. *J Fish Biol.* 2013;83:170–89.
- Bryan DR, McDermott SF, Nielsen JK, Fraser D, Rand KM. Seasonal migratory patterns of Pacific cod (*Gadus macrocephalus*) in the Aleutian Islands. *Anim Biotelem.* 2021. <https://doi.org/10.1186/s40317-021-00250-2>.
- Galuardi B, Lutcavage M. Dispersal routes and habitat utilization of juvenile Atlantic bluefin tuna, *Thunnus thynnus*, tracked with mini PSAT and archival tags. *PLoS ONE.* 2012;7(5): e37829.
- Schaefer KM, Fuller DW. Methodologies for investigating oceanodromous fish movements: archival and pop-up satellite archival tags. In: Morais P, Daverat F, editors. An introduction to fish migration. Boca Raton, FL, USA: CRC Press; 2016. p. 251–89.
- Luo J, Ault J, Ungar B, Smith S, Larkin M, Davidson T, et al. Migrations and movements of Atlantic tarpon revealed by two decades of satellite tagging. *Fish Fish.* 2020;21(2):290–318.
- Le Bris A, Fisher JAD, Murphy HM, Galbraith PS, Castonguay M, Loher T, et al. Migration patterns and putative spawning habitats of Atlantic halibut (*Hippoglossus hippoglossus*) in the Gulf of St Lawrence revealed by geolocation of pop-up satellite archival tags. *ICES J Mar Sci.* 2017;75(1):135–47.
- Lam CH, Nielsen A, Sibert JR. Improving light and temperature based geolocation by unscented Kalman filtering. *Fish Res.* 2008;9(1):15–25.
- Nielsen A, Sibert JR. State-space model for light-based tracking of marine animals. *Can J Fish Aquat Sci.* 2007;64(8):1055–68.

13. Sibert JR, Musyl MK, Brill RW. Horizontal movements of bigeye tuna (*Thunnus obesus*) near Hawaii determined by Kalman filter analysis of archival tagging data. *Fish Oceanogr.* 2003;12(3):141–51.
14. Pedersen MW, Patterson TA, Thygesen UH, Madsen H. Estimating animal behavior and residency from movement data. *Oikos.* 2011;120(9):1281–90.
15. Patterson TA, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J. State-space models of individual animal movement. *Trends Ecol Evol.* 2008;23(2):87–94.
16. Gatti P, Fisher JAD, Cyr F, Galbraith PS, Robert D, Le Bris A. A review and tests of validation and sensitivity of geolocation models for marine fish tracking. *Fish Fish.* 2021;22(5):1041–66.
17. Pedersen MW, Righton D, Thygesen UH, Andersen KH, Madsen H. Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and behavioural switching. *Can J Fish Aquat Sci.* 2008;65(11):2367–77.
18. Thygesen U, Pedersen M, Madsen H. Geolocating Fish Using Hidden Markov Models and Data Storage Tags. In: Nielsen J, Arrizabalaga H, Fragoso N, Hobday A, Lutcavage M, Sibert J, editors. *Tagging and Tracking of Marine Animals with Electronic Devices*. Springer, Netherlands: Reviews Methods and Technologies in Fish Biology and Fisheries; 2009. p. 277–93.
19. Nielsen JK, Mueter F, Adkison M, McDermott S, Loher T, Seitz AC. Effect of study area bathymetric heterogeneity on parameterization and performance of a depth-based geolocation model for demersal fish. *Ecol Model.* 2019;402:18–34.
20. Spies I, Barbeaux S, Ianelli JN, Ortiz I, Palsson W, Rand K, et al. Assessment of the Pacific cod stock in the Aleutian Islands. In: Plan team for ground-fish fisheries of the Bering Sea and Aleutian Islands (compiler), Stock Assessment and Fishery Evaluation report for the groundfish resources of the Bering Sea and Aleutian Islands. North Pacific Fishery Management Council; 2022.
21. Courtney MB, Scanlon BS, Rikardsen AH, Seitz AC. Marine behavior and dispersal of an important subsistence fish in Arctic Alaska, the Dolly Varden. *Environ Biol Fishes.* 2016;99(2–3):209–22.
22. Drumhiller KL, Johnson MW, Diamond SL, Reese Robillard MM, Stunz GW. Venting or rapid recompression Increase survival and improve recovery of red snapper with barotrauma. *Mar Coast Fish.* 2014;6(1):190–9.
23. Nielsen JK, Rose CS, Loher T, Drobny P, Seitz AC, Courtney MB, et al. Characterizing activity and assessing bycatch survival of Pacific halibut with accelerometer Pop-up Satellite Archival Tags. *Anim Biotelem.* 2018;6(1):10.
24. R Core Team. A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2022.
25. Thygesen UH, Pedersen M, Madsen H. Geolocating Fish Using Hidden Markov Models and Data Storage Tags. In: Nielsen J, Arrizabalaga H, Fragoso N, Hobday A, Lutcavage M, Sibert J, editors. *Tagging and Tracking of Marine Animals with Electronic Devices*. Springer, Netherlands: Reviews Methods and Technologies in Fish Biology and Fisheries; 2009.
26. Braun CD, Galuardi B, Thorrold SR. HMMocean: An R package for improved geolocation of archival-tagged fishes using a hidden Markov method. *Method Ecol Evol.* 2018. <https://doi.org/10.1111/2041-210X.12959>.
27. Le Bris A, Frechet A, Wroblewski JS. Supplementing electronic tagging with conventional tagging to redesign fishery closed areas. *Fish Res.* 2013;148:106–16.
28. Nichol DG, Chilton EA. Recuperation and behaviour of Pacific cod after barotrauma. *ICES J Mar Sci.* 2006;63:83–94.
29. Rose GA, deYoung B, Colbourne EB. Cod (*Gadus morhua* L.) migration speeds and transport relative to currents on the north-east Newfoundland Shelf. *ICES J Mar Sci.* 1995;52(6):903–13.
30. Ingvaldsen RB, Gjørseter H, Ona E, Michalsen K. Atlantic cod (*Gadus morhua*) feeding over deep water in the high Arctic. *Polar Biol.* 2017;40(10):2105–11.
31. Wallis KF. The two-piece normal, binormal, or double Gaussian distribution: its origin and rediscoveries. *Stat Sci.* 2014;29(1):106–12.
32. Webber DN. Modelling complexity and uncertainty in fisheries stock assessment [Doctor of Philosophy in Statistics]. Wellington: Victoria University of Wellington; 2015.
33. Zimmermann M, Prescott MM, Rooper CN. Smooth sheet bathymetry of the Aleutian Islands. U.S. Dep. Commer., NOAA Tech. Memo; 2013.
34. Hill R. Theory of geolocation by light levels. In: Le Boeuf BJ, Laws RM, editors. *Elephant seals: population ecology, behavior, and physiology*. Berkeley: University of California Press; 1994. p. 227–36.
35. Welch DW, Eveson JP. An assessment of light-based geo-position estimates from archival tags. *Can J Fish Aquat Sci.* 1999;56:1317–27.
36. Seitz AC, Norcross BL, Wilson D, Nielsen JL. An evaluation of light-based geolocation for demersal fish in high latitudes. *Fish Bull.* 2006;104:571–8.
37. Teo SLH, Boustany A, Blackwell S, Walli A, Weng KC, Block BA. Validation of geolocation estimates based on light level and sea surface temperature from electronic tags. *Mar Ecol Prog Ser.* 2004;283:81–98.
38. Arostegui M, Gaube P, Berumen M, DiGiulian A, Jones B, Røstad A, et al. Vertical movements of a pelagic thresher shark (*Alopias pelagicus*): insights into the species' physiological limitations and trophic ecology in the Red Sea. *Endang Species Res.* 2020;43:387–94.
39. Groger JP, Rountree RA, Thygesen UH, Jones D, Martins D, Xu Q, et al. Geolocation of Atlantic cod (*Gadus morhua*) movements in the Gulf of Maine using tidal information. *Fish Oceanogr.* 2007;16(4):317–35.
40. Viterbi A. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans Inf Theory.* 1967;13(2):260–9.
41. Lember J, Koloydenko AA. Bridging Viterbi and posterior decoding: a generalized risk approach to hidden path inference based on hidden Markov models. *J Mach Learn Res.* 2014;15:1–58.
42. Sasaoka K, Chiba S, Saino T. Climatic forcing and phytoplankton phenology over the subarctic North Pacific from 1998 to 2006, as observed from ocean color data. *Geophys Res Lett.* 2011;38(15):L15609. <https://doi.org/10.1029/2011GL048299>.
43. Mordy CW, Stabeno PJ, Ladd C, Zeeman S, Wisegarver DP, Salo SA, et al. Nutrients and primary production along the eastern Aleutian Island Archipelago. *Fish Oceanogr.* 2005;14(s1):55–76.
44. Ekstrom P. Error measures for template-fit geolocation based on light. *Deep-Sea Res II Top Stud Oceanogr.* 2007;54(3–4):392–403.
45. Ekstrom PA. An advance in geolocation by light. *Mem Natl Inst Polar Res, Spec Issue.* 2004;58:210–26.
46. Hunt GL Jr, Stabeno PJ. Oceanography and ecology of the Aleutian Archipelago: spatial and temporal variation. *Fish Oceanogr.* 2005;14:292–306.
47. Chassignet EPHH, Smedstad OM, Halliwell GR, Hogan PJ, Wallcraft AJ, Baraille R. Bleck R The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *J Mar Syst.* 2007;65:60–83.
48. Haase S, Krumme U, Gräwe U, Braun CD, Temming A. Validation approaches of a geolocation framework to reconstruct movements of demersal fish equipped with data storage tags in a stratified environment. *Fish Res.* 2021. <https://doi.org/10.1016/j.fishres.2021.105884>.
49. Braun CD, Skomal GB, Thorrold SR. Integrating archival tag data and a high-resolution oceanographic model to estimate basking shark (*Cetorhinus maximus*) movements in the western Atlantic. *Front Mar Sci.* 2018. <https://doi.org/10.3389/fmars.2018.00025>.
50. Braun CD, Gaube P, Afonso P, Fontes J, Skomal GB, Thorrold SR. Assimilating electronic tagging, oceanographic modelling, and fisheries data to estimate movements and connectivity of swordfish in the North Atlantic. *ICES J Mar Sci.* 2019;76(7):2305–17.
51. Arostegui MC, Braun CD, Gaube P. Movement and thermal niche of the first satellite-tagged Mediterranean spearfish (*Tetrapturus belone*). *Fish Oceanogr.* 2018. <https://doi.org/10.1111/fog.12413>.
52. Kearney K, Hermann AJ, Cheng W, Ortiz I, Aydin K. A coupled pelagic–benthic–sympagic biogeochemical model for the Bering Sea: documentation and validation of the BESTNPZ model (v2019.08.23) within a high-resolution regional ocean model. *Geosci Model Dev.* 2020;13:597–650.
53. Mauch M, Durski SM, Kurapov AL. Connectivity of the Aleutian north slope current and Bering Sea basin waters at the level of the subsurface temperature maximum: a modeling study. *J Geophys Res (C Oceans).* 2018;123(11):8608–23.
54. Liu C, Cowles GW, Zemeckis DR, Cadrin SX, Dean MJ. Validation of a hidden Markov model for the geolocation of Atlantic cod. *Can J Fish Aquat Sci.* 2017;74(11):1862–77.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.