Canadian Journal of Fisheries and Aquatic Sciences
Journal canadien des sciences halieutiques et aquatiques

Validation of a hidden Markov model for the geolocation of Atlantic cod

Journal:	*Canadian Journal of Fisheries and Aquatic Sciences*
Manuscript ID	cjfas-2016-0376.R1
Manuscript Type:	Article
Date Submitted by the Author:	06-Dec-2016
Complete List of Authors:	Liu, Chang; University of Massachusetts Dartmouth, School for Marine Science and Technology Cowles, Geoffrey; University of Massachusetts Dartmouth, School for Marine Science and Technology Zemeckis, Douglas; University of Massachusetts Dartmouth, School for Marine Science and Technology Cadrin, Steven X.; University of Massachusetts Dartmouth, School for Marine Science and Technology Dean, Micah; Massachusetts Division of Marine Fisheries, Annisquam River Marine Fisheries Field Station
Keyword:	geolocation, hidden Markov model, fish migration, Atlantic cod, data storage tags

Validation of a hidden Markov model for the geolocation of Atlantic cod

CHANG LIU[1], GEOFFREY W. COWLES[1], DOUGLAS R. ZEMECKIS[1, ⋆], STEVEN X. CADRIN[1], AND MICAH J. DEAN[2]

[1] *Department of Fisheries Oceanography*

School for Marine Science and Technology

University of Massachusetts Dartmouth

706 S Rodney French Blvd, New Bedford, MA 02744, USA

[2] *Annisquam River Marine Fisheries Field Station*

Massachusetts Division of Marine Fisheries

30 Emerson Ave., Gloucester, MA 01930, USA

Corresponding author: Chang Liu, cliu3@umassd.edu

⋆Present address: Department of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, NJ 08901

Abstract

Models developed to geolocate individual fish from data recorded by electronic tags often require significant modification to be applied to new regions, species, or tag types due to variability in oceanographic conditions, fish behavior, and data resolution. We developed a model for geolocating Atlantic cod off New England that builds upon an existing hidden Markov model (HMM) framework and addresses region- and species-specific challenges. The HMM framework contains a likelihood model which compares tag-recorded environmental data (depth, temperature, tidal characteristics) with those derived from an oceanographic model and a behavior model which constrains the horizontal movement of the fish. Validation experiments were performed on stationary tags, double-electronic-tagged fish (archival and acoustic tags), and simulated tracks. Known data, including fish locations and activity metrics, showed good agreement with those estimated by the modified approach, and improvements in performance of the modified method over the original. The modified geolocation approach will be applicable to additional species and regions to obtain valuable movement information that is not typically available for demersal fishes.

Key words: geolocation, hidden Markov model, fish migration, Atlantic cod, Gadus morhua, data storage tags

Introduction

The population structure of many fishery resources is more complex than the homogeneous units that are typically assumed in stock assessments and fishery management (Cadrin and Secor 2009). Recent research has increasingly focused on developing methods for incorporating complex population structures. In order to incorporate these spatial processes into stock assessment models and fishery management plans, it is essential to have a proper understanding of the movement of the species (Cadrin and Secor 2009; Goethel et al. 2011). The most common approach to studying movement of marine fish has been mark-recapture studies with conventional tags (Hall 2014). Conventional tags can provide information on general movements, but are not well suited for understanding behavioral patterns because they do not always reliably inform the trajectory of movement from release to recapture locations. In addition, conventional tagging typically relies on fishery-dependent recaptures, which can be biased by reporting rates and the distribution of fishing effort (Bolle et al. 2005).

To address these limitations, geolocation methods have been developed to utilize electronic tagging data to provide information about fish movements, distribution and behavior by estimating daily positions while fish are at liberty. Geolocation estimates are based on comparison of environmental data acquired from electronic tags (*e.g.*, temperature, pressure) with regional environmental databases (Evans and Arnold 2009). Geolocation methods have primarily utilized environmental data from recovered archival data storage tags (DSTs), including temperature, salinity, pressure (depth), and tidal data (amplitude/phase, tidal range/time of high water) (Arnold and Dewar 2001; Galuardi and Lam 2014), and these methods have been applied to demersal groundfish. Alternative approaches based on light as well as satellite-based geolocation have been used for pelagic fishes and marine mammals (Arnold and Dewar 2001; Block et al. 2011; Pedersen et al. 2011a), but are not applicable to benthic species due to attenuation of these signals in the water column.

Prior work in the geolocation of demersal fish can be categorized into two fundamental approaches: algorithmic methods and State Space Models (SSMs). In the algorithmic class of schemes (*e.g.* Hunter et al. 2003; Gröger et al. 2007; Neuenfeldt et al. 2007), positions at each time step (*e.g.* daily) are determined using a direct comparison of the environmental data recorded by the DST with data derived from regional observations or an oceanographic model. Algorithmic approaches lack the intrinsic ability to quantify uncertainty, which is a significant drawback given the potential for location errors to arise from noisy observations and environmental data (Patterson et al. 2008; Thygesen et al. 2009). In addition, a robust behavior model is often absent in algorithmic methods and conservative assumptions such as swimming speed constraints are instead applied. In contrast, state space models are statistical frameworks that can infer a series of state variables that are not directly measured, based on a series of observations that are conditioned on these unknown states. In the context of marine fish geolocation, the unknown states represent geographical locations of marine fish and the observation series is data recorded by DSTs (Patterson et al. 2008; Jonsen et al. 2013). Approaches based on state space models are largely able to overcome the drawbacks of algorithmic methods, because the uncertainty associated with the geolocations can be estimated, and a movement model describing the fish movement processes can be fit with observed data (Jonsen et al. 2013; Winship et al. 2012).

An important geolocation methodology based on the state space model framework is the hidden Markov model (HMM)(Pedersen et al. 2008, 2011a). The HMM is a form of state space model that deals with discrete states. In HMM, the estimation of the geographical location $\boldsymbol{x}$ is explicitly represented by a probability density function $\phi(\boldsymbol{x}, t)$. In each time step, the observation is dependent on the corresponding hidden state. Such dependency can be described by a likelihood model, represented by probability density functions constructed by comparing environmental data recorded by the tag with those from a model (*e.g.*, twilight light level model for light-based methods, oceanographic model for tidal- or depth/temperature-based methods). The hidden state sequence is a Markov chain bearing the assumption that the state at each time is dependent on the state at the previous time. Such dependency can be described by the behavior model. The output of an HMM is the estimated hidden time series of geographical locations and the associated posterior probability distribution functions.

The HMM method has been applied to the geolocation of Atlantic cod (*Gadus morhua*) in multiple regions (*e.g.*, North Sea (Pedersen et al. 2008; Thygesen et al. 2009), Gulf of St. Lawrence (Le Bris et al. 2013a,b), Iceland (Thorsteinsson et al. 2012)), as well as European seabass (*Dicentrarchus labrax*) along the west coast of France (Woillez et al. 2016). These efforts all used an open source MATLAB-based HMM geolocation toolbox developed by Pedersen (2008) (hereafter referred to as HGT), which is an implementation of a full HMM geolocation model. The kernel of HGT uses Bayes' theorem to calculate the normalized conditional probability distribution ϕ by performing a "time update" and an "observation update" during each timestep (Thygesen et al. 2009). Construction of $\phi(\boldsymbol{x}, t)$ enables the calculation of the most probable track (MPT). All Bayesian calculations in HGT are conducted on a regular orthogonal grid in a geographic coordinate system with a fixed spatial resolution.

A key challenge in the development of toolboxes such as HGT stems from the difficulty of generalizing the approach. For region- and species-specific applications of HMM geolocation, such models need careful calibration with available datasets. Environmental variables with the greatest spatial heterogeneity are most effective for geolocation. Therefore, the variables that are most useful for geolocation frequently vary by region. For example, previous groundfish geolocation efforts utilized different environmental variables such as tidal data in the North Sea (Metcalfe and Arnold 1997; Hunter et al. 2003, 2004; Wright et al. 2006; Thorsteinsson et al. 2012), depth and salinity in the Baltic Sea (Neuenfeldt et al. 2007), and depth and temperature in Gulf of St. Lawrence (Le Bris et al. 2013a,b) to help distinguish between horizontal locations.

Assessing the quality of position estimates is a key component to the development of new geolocation techniques. Previous studies have assessed the accuracy of DST-based geolocation using various approaches. One straightforward method is to compare the environmental parameters (*e.g.*, temperature, depth) measured by the tag with those estimated from the geolocated track (Neuenfeldt et al. 2007). However, a track whose corresponding environmental data matches the tag-measured values is not always biologically realistic (Brickman and Thorsteinsson 2008). Another approach to quantifying the accuracy of the track is to compare the estimated and true recapture location (Hunter et al. 2003). However, the premise of this method is the exclusion of the known recapture location from use in the geolocation process. Such exclusion may compromise the quality of the geolocation results, because the recapture location is a critical piece of information, especially for state space model-based methodologies with backward smoothing steps that propagate the recapture location information back to the whole time series. Other previous validation methods include geolocating DSTs moored on the bottom at fixed locations using tidal data (Hunter et al. 2003; Thorsteinsson et al. 2012), double-tagging the free swimming fish with two different type of electronic tags (Teo et al. 2004; Winship et al. 2012), and generating known movement tracks of virtual fish using simulation (Righton and Mills 2008). None of these approaches has been applied to state space model-based geolocation methodologies using depth and temperature data recorded by DSTs.

In the present work, we focus on the geolocation of Atlantic cod tagged with DSTs off New England, USA. Atlantic cod are an economically-important groundfish species for New England fisheries and many prior conventional tagging studies have been conducted (Hunt et al. 1999; Howell et al. 2008; Tallack 2011; Loehrke 2013). However, uncertainties remain with respect to cod behavior, movements, and stock structure, including the connectivity among subpopulations(Zemeckis et al. 2014b). In order to utilize HGT for the geolocation, several modifications are necessary. Firstly, due to inadequate spatial contrast in tidal characteristics in the western Gulf of Maine, the full tidal-based likelihood model in HGT must be modified to use other environmental variables. Secondly, as identified by Pedersen (2007), the land treatment in the HGT behavior model simply masks out cells that represent land, which potentially allows a fish to cross land. This is especially problematic in our region of interest due to the presence Cape Cod, a narrow and elongated land feature (Fig. 1). Modifications of the HMM methods in HGT were aimed at improving its performance for the current application, with consideration of also making it better suited for geolocating other groundfish species in the Gulf of Maine as well as other geographical areas. To achieve this objective, we made methodological contributions to the HMM geolocation package including incorporation of a depth- and temperature-based likelihood model with tidal-based exclusion in the HMM framework, and employed quantitative error assessment of the geolocation results using multiple approaches, including stationary mooring tags, double-electronic-tagged fish, and simulated tracks.

Materials and Methods

Archival tagging

As part of an interdisciplinary study, Atlantic cod were tagged with DSTs from 2010 through 2012 in the Spring Cod Conservation Zone (SCCZ, Fig. 1) (Dean et al. 2014; Zemeckis et al. 2014a; Zemeckis 2016), which is a seasonal spawning closure in northern Massachusetts Bay in the western Gulf of Maine (Armstrong et al. 2013). The DSTs deployed on a total of 266 Atlantic cod were Star-ODDI milli-L tags (39.4 mm × 13 mm, depth range 1–250 m; Star-ODDI Ltd., Reykjavik, Iceland). From these studies, a total of 49 DSTs were recovered from recaptured fish with data suitable for geolocation. The resolution and accuracy of pressure (depth) measurements was 0.03% and ± 0.8% of the calibrated depth range (1-250 m), respectively. The resolution of temperature measurements was 0.032 °C and the accuracy was ± 0.1 °C. The DSTs were programmed to record pressure and temperature measurements every 15 min and 2 h 45 min, respectively. To be consistent with depth data, temperature data were later interpolated to 15 min intervals using cubic spline interpolation (Trauth et al. 2007). Locations of release and recapture of tagged fish were also recorded. Each recapture location was assigned an uncertainty level of low (15 km) or moderate (30 km) based on the type of fishing gear (i.e. fixed or mobile) used to capture the tagged fish and the reliability of the positions based on the reported format (GPS coordinates, LORAN coordinates, or descriptive locations with reference to landmarks). Uncertainty was greater (moderate) for fish caught in mobile trawl gear due to the average tow distance by trawlers targeting cod in the Gulf of Maine (15.8 $\pm$ 9.3 km) and for reported recaptures that were not in GPS format and therefore less precise.

To provide an independent set of location estimates of better accuracy as a means of validating geolocation results, the DST recaptures included ten fish that also had a surgically-implanted Vemco V16P-6H coded acoustic transmitter (Vemco Division, AMIRIX Systems, Inc., Nova Scotia, Canada) (Zemeckis et al. 2014a). These double-electronic-tagged cod were in spawning condition when released (Dean et al. 2014). Between 2010–2014, acoustic receiver arrays were deployed to monitor cod spawning activity, including a Vemco Positioning System (VPS) in the cod conservation zone (see Fig. 2 in Dean et al. 2014) and acoustic receivers on both Eagle Ridge in Massachusetts Bay ($\sim$15 km south of the cod conservation zone) and Whaleback in Ipswich Bay ($\sim$45 km north of the cod conservation zone) (Zemeckis 2016). The positioning system in the cod conservation zone covered 9.5 km^2 and was able to determine horizontal positions with $<$10 m of error (Dean et al. 2014). In addition, acoustic receivers were deployed in Massachusetts Bay and off Cape Ann to monitor the movements of striped bass (*Morone saxatillis*) with the maximum detection range estimated at $\sim$1 km (see Fig. 1 in Kneebone et al. 2014).

Oceanographic model environmental data

We used bottom water temperature and bathymetry data from the Northeast Coastal Ocean Forecasting System (Beardsley et al. 2013; NECOFS 2013), which is based on the unstructured grid Finite-Volume Community Ocean Model (FVCOM) (Chen et al. 2006; Cowles et al. 2008). The NECOFS domain includes the entirety of the Gulf of Maine, Georges Bank, and the New England Shelf (Fig. 1), which covers all locations where cod from the western Gulf of Maine would be expected to be found based on observations from previous conventional tagging studies. The model mesh contains 90,415 elements in the horizontal grid and 45 vertical layers. The horizontal resolution ranges from 5 km near the open boundary to 500 m along the coast and tidal mixing fronts. The model is forced with hydrography and sea surface height at the open boundary, buoyancy flux from the major regional rivers, and wind stress and heat flux derived from regional hindcasts of the Weather Research and Forecasting (WRF) model. Observed data from moored arrays and sea surface temperature are assimilated into the hindcasts. Model bathymetry is based on the regional USGS 3-arcsec data product (Twomey and Signell 2013). NECOFS was hindcast for the period 1978–present and hydrographic data, velocity, and sea surface height were archived at hourly intervals. For tidal information the eight primary regional constituents (M_2, N_2, S_2, O_1, K_1, K_2, P_1, and Q_1) were derived using harmonic analysis from a barotropic setup of NECOFS used to simulate regional tides. In comparison with data from 98 sea surface gauges, the standard deviation for the model-data difference of the M_2 tidal constituent is 3.21 cm (Chen et al. 2011).

The NECOFS bottom water temperature is a critical component of the present geolocation effort. To assess the skill, model-computed bottom temperatures were compared with *in situ* measurements collected during multiple field surveys carried out between 2003 and 2015 (Table 1). A total of 29,501 data points of measurements that are within the NECOFS model domain cover the Gulf of Maine, Georges Bank, Southern New England and Mid Atlantic Bight, and have not been assimilated to NECOFS. The overall mean of the model-observation difference was $-0.04\,°\mathrm{C}$ and the overall RMSE was $1.61\,°\mathrm{C}$. The model-observation discrepancies did not exhibit significant seasonal or regional variation within the Gulf of Maine. Based on data from NECOFS, a typical range of bottom temperature across the Gulf of Maine and Georges Bank is approximately $7\,°\mathrm{C}$, a variation which is large compared to the NECOFS bottom temperature error. Following Willmott (1981), the NECOFS bottom temperature data was also examined using the non-dimensional metric:

$$W_s = 1 - \frac{\sum |T_{mo} - T_{me}|^2}{\sum(|T_{mo} - \overline{T_{me}}| + |T_{me} - \overline{T_{me}}|)^2}, \tag{1}$$

where T_{me} is the bottom temperature measurements, T_{mo} is the corresponding temperature from NECOFS, and the overbar denotes a mean. As opposed to the more broadly considered R^2, the Willmott score is able to distinguish constant or proportional offset between the two variables (Willmott 1981), and is commonly used in oceanographic model skill assessment studies (*e.g.* Warner et al. 2005; Wilkin 2006; O'Donncha et al. 2015). The skill score W_s has a range of 0–1, with 1 indicating perfect agreement between model and measurement and 0 indicating complete disagreement. For this comparison the skill value was 0.925, demonstrating strong agreement. In conclusion, the NECOFS bottom temperature data is generally appropriate for application to regional geolocation.

Hidden Markov model design

Geolocations for double-electronic-tagged cod were initially estimated using the original HGT which required only minor modification to work with NECOFS bathymetry and tidal data. These tracks were validated by comparison against acoustic telemetry data which provided known positions while the cod were at liberty (Supplementary Material). This study indicated that the accuracy of position estimates for the cod provided by the original HGT were not satisfactory for studying seasonal movement patterns of cod (median error >30 km), due primarily to inadequate spatial contrast in tidal characteristics, fish activity levels, and regional oceanographic conditions. We sought to improve HGT for application in the Gulf of Maine region, and provide a mechanism for enhanced performance in other regions and with other species. Building on previous work that aimed at assigning daily positions to statistical areas based upon DST data (Zemeckis 2016), revisions were made to the likelihood model, behavior model, and the most probable track construction in HGT. The HMM framework from the original HGT was maintained to calculate the posterior daily probability distribution of the fish. The source code of the modified HMM geolocation toolbox (revised HGT) is available at https://github.com/cliu3/hmm_smast. The domain for all HMM calculations presented in this paper ranges from 71°W to 62°W and 40°N to 45°N, including most of the Gulf of Maine and Georges Bank at a resolution of 0.05° which is approximately equal to 4 km.

Likelihood model

Likelihood distributions were derived using a comparison of depth, water temperature, and tidal information extracted from DSTs with the corresponding estimates from the oceanographic model. Daily likelihood distributions $L(\hat{\boldsymbol{x}})$, representing the probability of the observation data given the discrete horizontal geographical location $\hat{\boldsymbol{x}}$, were constructed on the vertices of the unstructured grid of the oceanographic model. The approach considered the influence of temperature and depth separately from that of tides. Limited regional variation of the tidal characteristics in the western Gulf of Maine (Chen et al. 2011) reduces the utility of tides for geolocation. The M_2 amplitude and phase may vary by only 0.25 m and 15°, respectively across a distance of 130 km. Additionally, off-bottom movement of fish can reduce or eliminate the ability to detect tide in the pressure signal. Considering these two factors, a geolocation method based solely on tidal information is not capable of producing sufficient accuracy in the Gulf of Maine for studying seasonal movement patterns of demersal fishes. Nonetheless, useful information may still be extracted from the tide signal. In the present work, an initial likelihood distribution $L_{dt}(\hat{\boldsymbol{x}})$ was constructed using depth and temperature information. Tide, when available, was then used for eliminating unlikely regions in the final $L(\hat{\boldsymbol{x}})$ distribution.

The specific parameterization of the likelihood function depends on the daily activity of each fish, which was categorized as low, medium, or high using pressure data from the DST. We employed the tidal fitting procedure of Pedersen (2007), which calculates the least-square fit of the depth signal with a sinusoidal wave. Days were categorized as low activity when there was a satisfactory fit over a 13 h window, moderate activity days were identified as those with satisfactory fits when using a 5 h window, and high activity days were those during which there were no reliable tidal fits (Fig. 2). This classification is based on the assumption that longer tidal fit represents demersal behavior at a fixed location and depth, and therefore less horizontal movement. The criteria for goodness of fit for detection of tidal signal was strict (root mean square error (RMSE) < 0.35 m, $R^2 > 0.92$, and tidal amplitude between 0.2 m and 2.0 m) to prevent false tidal fits which compromised estimates of tidal phase and therefore geographic position. In contrast, a more relaxed tidal fitting criteria was employed for identifying moderate activity periods ($R^2 > 0.85$), because tidal characteristics were not used for geolocation on moderate activity days.

Assuming that depth and temperature were independent, an initial likelihood distribution $L_{dt}(\hat{\boldsymbol{x}})$ given the observed depth and temperature (z, T) is obtained by forming the product of two integrated normal distributions (modified from Le Bris et al. 2013b):

$$L_{dt}(\hat{\boldsymbol{x}}) \quad = \quad \int_{z-\Delta z}^{z+\Delta z} N\Big(z; \mu_z(\hat{\boldsymbol{x}}), \sigma_z(\hat{\boldsymbol{x}})\Big) \mathrm{d}z \quad \times \quad \int_{T-\Delta T}^{T+\Delta T} N\Big(T; \mu_T(\hat{\boldsymbol{x}}), \sigma_T(\hat{\boldsymbol{x}})\Big) \mathrm{d}T, \quad (2)$$

where Δz and ΔT are the tag measurement error for depth and temperature, respectively, $N(\mu, \sigma^2)$ is a normal distribution function of mean μ and standard deviation σ, and μ_z and μ_T are NECOFS depth and temperature. The standard deviations of bathymetry $\sigma_z(\hat{\boldsymbol{x}})$ and temperature $\sigma_T(\hat{\boldsymbol{x}})$ were determined using the NECOFS depth and temperature values from the neighboring vertices of $\hat{\boldsymbol{x}}$ on the unstructured grid. During low and moderate activity periods, z and T were established using the mean depth and temperature over the satisfactory tidal fit. Taking an average over the depth signal removes the sinusoidal tidal variation and represents better the bathymetry of the fish's location, whereas the mean temperature is an appropriate choice for comparison with the NECOFS daily-averaged bottom temperature data. During high activity periods, the depth-based likelihood factor is replaced by a bathymetry uncertainty, after Pedersen (2007):

$$L_{dt}(\hat{\boldsymbol{x}}) = \Phi\left(\frac{z-\mu_z(\hat{\boldsymbol{x}})}{\sigma_z(\hat{\boldsymbol{x}})}\right) \Big/ \Phi\left(\frac{-\mu_z(\hat{\boldsymbol{x}})}{\sigma_z(\hat{\boldsymbol{x}})}\right) \times \int_{T-\Delta T}^{T+\Delta T} N\Big(T;\mu_T(\hat{\boldsymbol{x}}),\sigma_T(\hat{\boldsymbol{x}})\Big)\mathrm{d}T, \tag{3}$$

where Φ is the cumulative density function of a standard Gaussian distribution, z and T were set using the depth and temperature when the fish was at its maximum depth during the daily interval. This treatment is based on the constraint that the depth of the fish is always less than the local bathymetry and accounts for bathymetry uncertainty.

When available, tidal information derived from tag data was used to eliminate unlikely locations from the initial likelihood distribution. During low activity periods, the tag tidal signal (η) was compared with tidal signals for the same period from the oceanographic model ($\hat{\eta}(\hat{\boldsymbol{x}})$) using the root-mean-square deviation (RMSD) of the two time series at each NECOFS grid point $\hat{\boldsymbol{x}}$:

$$RMSD(\hat{\boldsymbol{x}}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\eta}_i(\hat{\boldsymbol{x}})-\eta_i)^2}, \tag{4}$$

where n is the number of measurements in the 13-hour time series of the tide signal on a given day. The initial likelihood distribution $L_{dt}(\hat{\boldsymbol{x}})$ was then preserved at grid points where two conditions were met: 1) the semi-diurnal amplitude of the tag signal $A(\eta)$ is bounded by the amplitude of M_2 minus that of the sum of the other seven tidal constituents $A_{\mathrm{M}_2-\Sigma7}(\hat{\boldsymbol{x}})$ and the sum of all eight principal tidal constituents $A_{\Sigma8}(\hat{\boldsymbol{x}})$; and 2) the RMSD was smaller than a threshold value Θ which was the 30th percentile of the RMSD calculated for the remaining grid points. Implementation of the first condition avoids the computation effort for reconstructing tidal signals ($\hat{\eta}$) on grid points where the semi-diurnal amplitude clearly do not match that of the tag signal. In the second condition, the value of Θ was established using performance testing which found that it was able to eliminate obviously spurious position assignments. In addition, it also preserved $L(\hat{\boldsymbol{x}})$ within a fairly broad horizontal scale so that potential true positions do not get excluded. This scale was determined based on the observed error of the double-electronic-tagged cod using the original HGT. For grid points not meeting these two criteria, the likelihood was assigned a zero value (Fig. 3). In summary, the final likelihood distribution $L(\hat{\boldsymbol{x}})$ with tidal exclusion can be expressed as:

$$L(\hat{\boldsymbol{x}}) = L_{dt}(\hat{\boldsymbol{x}})H(\hat{\boldsymbol{x}}), \tag{5}$$

where

$$H(\hat{\boldsymbol{x}}) = \begin{cases} 1, & RMSD(\hat{\boldsymbol{x}}) \leq \Theta \\ & \text{and } A(\hat{\eta}) \in [A_{\mathrm{M}_2-7}(\hat{\boldsymbol{x}}), A_8(\hat{\boldsymbol{x}})] \\ \\ 0, & \text{all other positions} \end{cases} . \tag{6}$$

For days when tidal information was insufficient or absent from the tag data (i.e. during moderate or high activity), tidal exclusion was not employed:

$$L(\hat{\boldsymbol{x}}) = L_{dt}(\hat{\boldsymbol{x}}). \tag{7}$$

Behavior model

The behavior model describes the time evolution of the state variable, which is the daily movement of the fish. The horizontal movement of fish can be represented as a random walk (Sibert et al. 1999) which can be mathematically described using the Fokker-Planck diffusion equation:

$$\frac{\partial \phi}{\partial t} = D\nabla^2\phi, \tag{8}$$

where ϕ is the probability density of the fish's location and D is a constant diffusivity coefficient, which is related to the swimming speed of the fish. The discretization scheme of the diffusion process was previously implemented in HGT following Thygesen et al. (2009), using a transition probability matrix representing an isotropic Gaussian kernel corresponding to the solution of Eq. 8. In this approach, the matrix is defined as $\mathbf{H} = (\lambda_{ij})$, where element (i, j) represents a spatial location, and λ_{ij} represents the probability that the fish moves from the center element of $\mathbf{H}$ to element (i, j). The isotropic approach handles dry land by simply setting transition probabilities in these elements to zero (Thygesen et al. 2009; Pedersen et al. 2011a), allowing artificial crossing of fish from one side of a peninsula or other small scale land features to the other within a single time step. To prevent such infeasible results, the generation of the transition probability matrix was modified in the revised HGT. The transition probability matrix $\mathbf{H}$ was first initialized as an empty matrix, with elements representing land masked out. A breadth-first searching algorithm was then used to generate a distance field $\mathbf{S} = (s_{ij})$ of the same size as the transition probability matrix, with values equal to the shortest apparent distance from each element to the center element of the matrix around any masked-out obstacles. The values of the transition probability matrix λ_{ij}

were then reassigned by evaluating the original Gaussian function at values of the apparent distance field **S**. The effect of this treatment near land is equivalent to that of a reflecting boundary condition.

The behavior switching scheme described in Pedersen et al. (2008) which makes use of the activity level classification (Fig. 2) was also used in this work. A lower value of the diffusivity coefficient D was used for low and moderate activity days and a higher D for high activity days. The values of D can be specified as constant values or estimated using maximum likelihood estimation (MLE) (Pedersen et al. 2008). For simplicity and inclusiveness, in this study D was assigned constant values of 10 km^2/d as the lower value and 100 km^2/d as the higher value. This decision was based on the estimation of D from fish swimming speed presented by Pedersen (2007) considering the typical swimming speed of cod (Fernö et al. 2011) and allowing for broader ranges of horizontal movement.

Most probable track

In the original HGT, the most probable track is one that maximizes the overall probability score of the whole sequence of locations using the Viterbi algorithm (Pedersen 2007; Thygesen et al. 2009), and the end point of the most probable track was set to be the grid cell where the value of the probability distribution ϕ on recapture day is the greatest. We modified the approach to make sure the end point of the estimated MPT is close to the reported recapture location. The final point of the tag deployment was set to be the grid cell with the maximum ϕ value among the cells that are within the uncertainty radius of the reported recapture location. This modification effectively nudges the estimated location on the day of recapture to be within the uncertainty radius of the reported recapture location.

In summary, the original HGT consists of a tidal-based likelihood model, a spatially discretized Gaussian behavior model with simple land treatment, and an MPT search scheme based on the Viterbi algorithm. Modifications made in the revised HGT include the utilization of tag-recorded depth and temperature and the exclusion of unlikely locations based on tidal characteristics for the likelihood model, the activity classification based on length of tidal signal detection, improved land treatment in the behavior model, and a method to constrain the end point of the most probable track to be near the reported recapture location.

Validation experiments

To examine the performance of the revised HGT, the method was applied to two classes of DST datasets (including depth and temperature) with known locations. The first, bottom-mooring tags, challenge the model to maintain a fixed position over time. The second class of dataset consists of double-electronic-tagged fish that provide known locations that enable direct quantification of model skill when they pass through acoustic receiver arrays. This second class is useful for providing confidence in the geolocation, because the data is obtained from the tagged fish. To examine whether the revised HGT improves geolocation performance, the performance of the original HGT was also assessed using these two classes of DST datasets for comparison.

Another approach for validating the geolocation methodology is to assess the model's ability to replicate simulated tracks. Data for these fish were generated by interpolating pressure and temperature from the oceanographic model onto artificially constructed tracks.

In this study, simulated fish tracks were generated to examine the effect of season, region, and time at liberty on the accuracy of the geolocation results. The release positions were informed by the time and location of cod presence within the western Gulf of Maine inferred by recapture positions from conventional tag studies (Zemeckis 2016; Zemeckis et al. 2017). Movement tracks were simulated to occupy different regions (Gulf of Maine and Georges Bank) during two seasons (summer and winter) across a range of days at liberty (40 d, 120 d, and 360 d)(Fig. 4). Daily locations for each track were generated using a random walk with the following equation:

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t + R\sqrt{2D\delta t}, \tag{9}$$

where $\boldsymbol{X}_{t+1}$ and $\boldsymbol{X}_t$ are locations in the simulated track on day $t+1$ and t, respectively, R is a random factor producing a standard normal distribution (zero mean and unit variance), D is the diffusivity having a value of 10 km^2/d or 100 km^2/d, and $\delta t = 1$ d is the time interval. Simulated individuals were constrained to remain in the model domain. If an individual moved across land or open-ocean boundary during a time step $t+1$, it was restored to its last position (from the previous time step $\boldsymbol{X}_t$). This boundary treatment method was chosen because of the ease of implementation within the unstructured mesh framework of NECOFS FVCOM. After the simulated track was generated, the corresponding depth and temperature time series were constructed at 15 min intervals using the tidal and bottom temperature data derived from the oceanographic model in order to create a simulated tag. No noise was added to the simulated depth and temperature signals. Ten simulation sets consisting of five runs each were performed. Each set was based on a unique combination of season, region, and time at liberty (Table 2). When performing geolocation using the simulated data, release locations were used without uncertainty, while recapture location uncertainty was 15 km.

Results

Geolocation Model Validation

To validate the activity characterization approach of the likelihood model, we compared the size of the daily 95% utilization distribution derived from VPS detection reported in (Dean et al. 2014) with the daily activity levels determined by the likelihood model. The median areas of the daily 95% utilization distribution were 0.038 km^2 for the low activity days, 0.11 km^2 for the moderate activity days, and 0.26 km^2 for the high activity days (Fig. 5). The relation between these two metrics shows a trend in which days classified as lower levels of activity based on vertical movements are those during which the fish utilized less space horizontally.

A total of 14 Star-ODDI DSTs were moored to different fixed locations on cod spawning sites in Massachusetts Bay and Ipswich Bay between 2010–2012 and Jeffreys Ledge between 2014–2015 in order to test the performance of the DSTs and validate the geolocation methodology. Geolocation using the revised HGT were performed on tag-recorded data from these deployments, in which release and recapture locations were used without uncertainty. Daily location estimations in the most probable track were compared with the known mooring locations. The most probable track estimations for the 14 mooring DST deployments were close to their deployment locations. The RMSE of the daily location estimation from all mooring tags was 11.07 km and the error range was 0.14–25.51 km (Table 3a). The median geolocation error for all mooring tags was 4.93 km. This represents a significant improvement over the error of 33.94 km found using the original HGT (Table 3a, 3b). Tag #73 was the best performing deployment (Fig. 6a) with a median daily location error of 1.86 km, whereas tag no. 71 (off Provincetown, Cape Cod) was the worst performing deployment (Fig. 6c) with a median daily location error of 23.10 km. Tag no. 87, for which the median error was 4.79 km, was representative of the overall mooring tag deployments (median 4.93 km) (Fig. 6b). To assess the accuracy of the constructed probability density functions, mean normalized probability at known locations were calculated for each track to give a value between 0 and 1, where 1 indicates that the probability density function most accurately estimates the known locations, and 0 indicates that the probability density function is unable to correctly estimate the known locations. The overall mean normalized probability at known locations for all mooring tags ranged from 0.30 – 1, with an average of 0.69. Compared with the same metric derived from the original HGT (0.06), this represents a significant improvement (Table 3a, 3b).

High resolution positions of the double-electronic-tagged cod determined by acoustic receivers were compared to the same-day position estimates from the most probable track constructed by the revised HGT. To assess whether the revision to HGT improved geolocation results, acoustically detected location were also compared with position estimates using the original HGT with minimum changes only to enable the input of NECOFS bathymetry and tidal data. Most (217 out of 223, 97.3%) of the daily locations of the most probable track estimated by the revised HGT were within 42 km of the acoustically-detected locations (Fig. 7). The median geolocation error for the revised HGT was 6.45 km, which is an improvement over the value of 34.80 km found using the original HGT (Table 3c, 3d, Supplementary Material Table S1). This reduction in error is essential for studying seasonal movement of cod in the Gulf of Maine, because all the double-electronic-tagged fish were recaptured within 82 km of their release location. The average normalized probability at the acoustically-detected locations was 0.47 for the revised HGT, much higher than that of the original HGT, 0.06 (Table 3c, 3d). Although the median geolocation error was less in the modified model, in rare cases (6 out of 223 estimates, <3%) errors in such estimates were found to be between 33–62 km greater than that of the original HGT. These six estimates also had the greatest error and were all from fish no. 22 which had the longest duration (212 d) (Table 4).

In the simulated track experiments, the most probable track output was compared with the simulated tracks. The mean and median location estimation error for the simulated tracks were 92.40 km and 69.46 km, respectively. The mean normalized probability at known locations was 0.39. A breakdown of the daily location errors for all simulated tracks indicated variation of location errors among seasons, geographical regions, and numbers of days between release and recapture (Fig. 8). Across all seasons, the median error increased when fish were at liberty for a longer period. This finding is consistent with results from the double-electronic-tagging experiments which found that geolocation errors for cod were greater for cod that spent longer time in the water. For simulated runs with duration of 40 d and 120 d, the median error during winter was greater than during summer. Estimated location errors of the Gulf of Maine tracks were slightly greater than those of the Georges Bank tracks in general, with the 120 d tracks released in winter as exceptions.

Geolocation of the double-electronic-tagged cod

The revised HGT was applied to the double-electronic-tagged fish (n=10). All ten cod were recaptured in the Gulf of Maine and within 82 km of their release position in the cod conservation zone (Table 4, Fig. 9), with the average number of days at large being 79.5 days. The distance between the reported and estimated recapture locations were all within the uncertainty radius around the reported recapture locations except fish no. 22, which exceeded its uncertainty radius of 30 km by 4.3 km. Five fish (nos. 7, 8, 11, 12, and 13) moved east towards Stellwagen Bank, with two (nos. 12 and 13) exhibiting a stationary period in southern Massachusetts Bay classified as mostly low activity days (Fig. 10). Geolocation results demonstrated that cod moved offshore after spawning. Most cod remained within the western Gulf of Maine. However, two fish (nos. 18 and 22) moved to the southeast towards the Great South Channel and Georges Bank before migrating north and being recaptured in the Gulf of Maine. These movements represent migrations across the current boundary between the Gulf of Maine and Georges Bank management units (see NEFSC 2013).

Cod no. 16 generally stayed in the cod conservation zone throughout its 27 days at liberty, corroborated by acoustic receiver detections being received on each day when it was at large with the exception of 21 June 2010. No. 17 traveled north towards Ipswich Bay, which is a major cod spawning ground during the spring. No. 24 moved to Stellwagen Bank and was later recaptured on southern Jeffreys Ledge.

Discussion

Geolocation methods

The geolocation method presented in this paper is a direct development from the HMM geolocation method presented by Pedersen et al. (2008) and implemented in HGT. New elements developed in the present geolocation method and implemented into the revised HGT have improved model performance for our application. These include the exclusion of unlikely locations based on tidal characteristics, the utilization of depth and temperature and the tidal-based activity classification for the likelihood model, improved land treatment in the behavior model, and a method to constrain the end point of the most probable track to be near the reported recapture location. The introduction of the moderate activity enhances the utility of vertical behavioral information. Validation in activity classification using the VPS occupancy utilization data links the horizontal and vertical movement of the fish. Although Hobson et al. (2009) concludes that there is no decisive connection from vertical behavior pattern of cod to its horizontal migration or residence behavior, our validation results indicate a pattern that cod tend to utilize larger areas when greater vertical activity is observed, which justifies the use of multiple values of the diffusivity coefficient D corresponding to different activity levels in the behavior model. One caveat of this validation is that such justification is based on data collected from a specific behavior period because the double-electronic-tagged cod were all in spawning condition, which may be a period when cod are more sedentary than they are at other times of the year. Also worth noting is that our behavior classifications are based on available behavioral observations and relevant to Gulf of Maine cod, whereas cod in other regions may exhibit different behavior. Secondly, the exclusion of unlikely locations based on tidal characteristics was inspired by fully tidal-based methods (*e.g.* Hunter et al. 2003, 2004; Gröger et al. 2007; Pedersen et al. 2008), which do not perform well in regions where tidal variation is small. Exploratory experiments in which tidal characteristics were incorporated in the joint likelihood distribution in a similar way with depth and temperature indicated that such inclusion misleads the location estimates in the western Gulf of Maine. By excluding unlikely locations, the accuracy of the likelihood model and the computational efficiency were improved. Therefore, this tidal exclusion scheme is the primary reason that the revised HGT demonstrated better performance over the original HGT in the mooring and double-tagging validation experiments. In the original HGT, the land treatment in the behavior model allowed unrealistic crossing of peninsulas and other promontories. Pedersen et al. (2011b) employed a finite element method to solve the nonlinear Bayesian fish tracking problem on domains with irregular geometry, which is an ideal method for land avoidance in terms of accuracy, but at the expense of computational efficiency. In our modification to the HGT we focused on using an approach that was straightforward to implement to improve the land treatment scheme without significantly increasing the computational load. Our modification eliminates the possibility of fish crossing over land. Lastly, confining the estimated recapture location of the most probable track near the reported recapture location resulted in a track that is more realistic.

Accuracy of geolocation estimations

This validation study is a comprehensive effort for DST-based geolocation methods applied to demersal fishes. Model validation experiments using fixed mooring tags and double-electronic-tagged cod indicated that the revised HGT produces more accurate results than previous tidal- or light-based methods using archival tags. The estimated error using revised HGT for mooring tags at fixed locations was between 0.14 and 25.51 km, with a mean value of 11.07 km (Table 3a, Supplementary Material Table S1). Hunter et al. (2003) and Thorsteinsson et al. (2012) used mooring tags fixed at known locations to validate their tidal-based method and their reported average error was 15.7 $\pm$ 3.5 km and 18.91 km, respectively. The root mean square error (RMSE) of our method for double-electronic-tagged fish was 21.87 km (Table 3b). Double-tagging studies of sharks (Teo et al. 2004; Winship et al. 2012) found errors $> 0.5°$ (approximately equal to 55 km), but the error is likely greater for sharks since they tend to have higher horizontal speeds and travel more frequently than groundfish. Righton and Mills (2008) reported that the average error for their DST-based method using five 50-d simulated tracks determined by the most likely path using a highest total score approach was between 37 and 69 km. The median error of our 40-d simulated track runs, which was determined by the most probable track using similar criteria maximizing the overall score, was 29.16 km.

Comparison of the geolocation results of the ten double-electronic-tagged cod using revised HGT with the statistical area assignment for the same cod (based on the common numbering listed in the "DMF Fish ID" column in Table 4) presented in Zemeckis (2016) and Zemeckis et al. (2017) indicated that the revised HGT was capable of providing superior geolocation estimates compared to a coarse scale algorithmic geolocation method. Although the two methods share the same likelihood model, by introducing HMM in the geolocation method, drawbacks in the previous algorithmic method that lead to occasionally erroneous position assignments were overcome in the revised HGT.

Geolocation of stationary tags indicated that the current method is able to provide highly accurate location estimates for fixed-location objects. Errors in archival tag measurements and depth and temperature data derived from the oceanographic model are potential sources of error in geolocation estimates of the fixed-location tags. In comparison, location estimation error was nearly doubled for the double-tagging experiment of free-swimming cod (Table 3). Such comparison indicates that the current behavior model may be another significant source of location error in addition to that induced by tag data and the oceanographic model errors; the current behavior model is likely the barrier to achieving highly accurate location estimates for free-swimming fish. A behavior model that more accurately describes the spatial movements of the fish species in question is expected to improve the accuracy of geolocation estimates. We assumed fish movement could be modeled with a random walk. The use of alternative schemes such as Brownian motion or Lévy flight have been shown to have a negligible effect on geolocation when compared with the random walk (Thygesen and Nielsen 2009). Moreover, the underlying behavior state time series of a fish can be estimated more accurately using a separate or extended state space model framework (Patterson et al. 2009, 2016). Pedersen et al. (2011a) present a similar HMM framework which estimates behavior and movement at the same time. (Pedersen et al. 2011a) also includes a model selection scheme for the behavior model with a candidate set of models with different set of parameters including advection, which is not considered in the current method. However, the implementation of such behavior state schemes will increase the mathematical complexity and the computational intensity of the geolocation model. When considering alternative behavior models in future efforts, both the computational efficiency and the accuracy of the geolocation should be considered.

Geolocation results of stationary tags (Fig. 6) also suggest that spatially-varying systematic biases may exist in geolocation estimates. Such biases may be caused by local bathymetry and oceanographic conditions that result in similar temperature and depth over a broader area. Similar phenomenon was reported for other telemetry techniques for estimating fish locations and can be potentially corrected by deploying stationary tags throughout the study area (Charles et al. 2016). To better understand the effect of systematic biases in geolocation estimates, fixed-location mooring deployments are recommended for future geolocation tagging projects.

Simulated track experiment results suggested that geolocation estimates using revised HGT were more accurate for fish at liberty for fewer days, tagged during summer when spatial variation of bottom temperature is relatively large, and released in regions where bathymetric variation is large. The seasonality of geolocation accuracy was similar to the conclusions made by Righton and Mills (2008). These findings may provide guidance for future geolocation tagging to help achieve more accurate location estimates.

Exploratory analyses showed that geolocation estimates of the simulated tracks are more accurate using the original HGT compared with those of the present work. This finding is intuitive given the inherent differences between the two approaches. In these simulated tracks, the tidal signal is derived directly from the NECOFS database and thus the tidal model is effectively without error. In contrast to the revised HGT which employs the tidal signal for the purpose of exclusion, the original HGT incorporates the spatial variation of the tidal signal in the geolocation process and thus is able to take advantage of the perfect fit between the model and tag data in the simulations. With real tag data and an imperfect tidal database, attempts to incorporate directly the tidal information can have an adverse effect on the geolocation accuracy (Le Bris et al. 2013b), as demonstrated in the aforementioned double-electronic-tagged experiments. Nonetheless, the original HGT may show good performance in areas where the variation in the spatial tidal characteristics is significant compared to errors associated with tag measurement and tidal database, such as the North Sea.

Applications

Results of this work may have implications for the regional fishery management of cod. The residency exhibited in geolocation estimates of eight double-electronic-tagged cod (nos. 7, 8, 11, 12, 13, 16, 17, and 24) is similar to findings from previous conventional tagging studies (Hunt et al. 1999; Tallack 2011; Loehrke 2013) which classified cod in the Gulf of Maine as sedentary (Howell et al. 2008). However, such agreement may be a result of limited DST durations (<3 months) and limitations of conventional tagging comparing only release and recapture locations, both limitations tend to underestimate the horizontal activity of cod. Moreover, geolocation estimates of the other two double-electronic-tagged cod (nos. 18 and 22) indicate movements across the current management unit boundary between the Gulf of Maine and Georges Bank management units, similar to the results of Gröger et al. (2007). Such movements would not have been observed with conventional tagging methods because these cod were released and recaptured in the same management unit. Results from further application of the geolocation method to available DST tag data of cod off New England may have important implications for future stock identification regarding the delineation of management unit boundaries.

The HMM-based geolocation method presented in this work is expected to be applicable to other demersal groundfish species. For example, within the northeast U.S. region alone, DSTs have been used to study multiple demersal species (*e.g.*, yellowtail flounder, Cadrin and Westwood 2004; monkfish, Grabowski et al. 2013; summer flounder, Henderson and Fabrizio 2014; winter flounder, Coleman 2015; black sea bass, Moser and Shepherd 2009; Atlantic halibut, Kanwit et al. 2008). The lack of access to validated geolocation methods creates barriers to the process of deriving reliable movement information from the tag data. The current study provides a geolocation method that would be applicable to these other datasets, thereby breaking some of these barriers.

Global or regional oceanographic data that are relevant to the current HMM geolocation method, such as temperature, tides, and bathymetry, are readily available, which enables the applicability of the current HMM geolocation method to other regions. The Oregon State University Tidal Inversion Software (OTIS) and the associated MATLAB Tidal Model Driver toolbox (Egbert and Erofeeva 2002) are capable of providing global tidal harmonics data. Databases of ocean general circulation model (OGCM) output typically contain 4-dimensional sea water temperature. A review of some regional and global data products, including model descriptions and how to obtain model outputs, was given by Potemra (2012). For better accuracy of the geolocation estimates, the spatial resolution of such environmental data needs to be higher than the estimated location error scale.

Conclusions

We implemented an HMM-based geolocation model for Atlantic cod in the Gulf of Maine. The model framework utilizes temperature and depth data from DSTs for location estimation, and tidal data for exclusion of unlikely locations. A tidal-based daily activity level classification scheme was implemented to improve the accuracy of the likelihood distribution and determine the behavior states. Comprehensive validation experiments were performed on stationary mooring tags, double-electronic-tagged fish, and simulated tracks. Validation results suggest good performance of the revised geolocation model and improvements in performance over the original approach. This method could be applied to other demersal groundfish species, and is relevant to future stock identification and fishery management.

Acknowledgements

The authors would like to thank the three reviewers for their valuable comments and suggestions that did much to improve the manuscript. Cod tagging research in the Spring Cod Conservation Zones was supported by the United States Fish and Wildlife Service through the Sportfish Restoration Act and the Massachusetts Marine Fisheries Institute. Funding for the research conducted as part of this manuscript was provided by NOAA Saltonstall-Kennedy Grant award NA15NMF4270267 and the University of Massachusetts Intercampus Marine Science (IMS) Graduate Program. Part of the HMM geolocation computations were made on the University of Massachusetts Dartmouth's GPGPU high performance computing cluster, which was acquired with support from NSF award CNS-0959382. The authors would like to thank William Hoffman, Michael P. Armstrong, and David Martins for assisting with fish tagging and data collection, Gregory DeCelles and N. David Bethoney for providing survey data with bottom temperature measurements, Arnault Le Bris for providing feedback that has improved the manuscript, and Martin W. Pedersen for granting permission to reuse the HMM Geolocation Toolbox and providing technical advice.

References

Armstrong, M.P., Dean, M.J., Hoffman, W.S., Zemeckis, D.R., Nies, T.A., Pierce, D.E., Diodati, P.J., and McKiernan, D.J. 2013. The application of small scale fishery closures to protect Atlantic cod spawning aggregations in the inshore Gulf of Maine. Fish. Res. **141**: 62–69. doi:10.1016/j.fishres.2012.09.009.

Arnold, G. and Dewar, H. 2001. Electronic tags in marine fisheries research: A 30-year perspective. In Electronic Tagging and Tracking in Marine Fisheries, edited by J.R. Sibert and J.L. Nielsen, Springer Netherlands, number 1 in Reviews: Methods and Technologies in Fish Biology and Fisheries, pp. 7–64. Doi: 10.1007/978-94-017-1402-0_2.

Beardsley, R.C., Chen, C., and Xu, Q. 2013. Coastal flooding in Scituate (MA): A FVCOM study of the 27 December 2010 nor'easter. J. Geophys. Res.-Oceans **118**(11): 6030–6045. doi:10.1002/2013JC008862.

Block, B.A., Jonsen, I.D., Jorgensen, S.J., Winship, A.J., Shaffer, S.A., Bograd, S.J., Hazen, E.L., Foley, D.G., Breed, G.A., Harrison, A.L., Ganong, J.E., Swithenbank, A., Castleton, M., Dewar, H., Mate, B.R., Shillinger, G.L., Schaefer, K.M., Benson, S.R., Weise, M.J., Henry, R.W., and Costa, D.P. 2011. Tracking apex marine predator movements in a dynamic ocean. Nature **475**(7354): 86–90. doi:10.1038/nature10082.

Bolle, L.J., Hunter, E., Rijnsdorp, A.D., Pastoors, M.A., Metcalfe, J.D., and Reynolds, J.D. 2005. Do tagging experiments tell the truth? Using electronic tags to evaluate conventional tagging data. ICES J. Mar. Sci. **62**(2): 236–246. doi:10.1016/j.icesjms.2004.11.010.

Brickman, D. and Thorsteinsson, V. 2008. Geolocation of Icelandic cod from DST data using a modified particle filter method. ICES CM 2008/P.09 .

Cadrin, S.X. and Secor, D.H. 2009. Accounting for spatial population structure in stock assessment: Past, present, and future. In The Future of Fisheries Science in North America, edited by R.J. Beamish and B.J. Rothschild, Springer Netherlands, number 31 in Fish & Fisheries Series, pp. 405–426. Doi: 10.1007/978-1-4020-9210-7_22.

Cadrin, S.X. and Westwood, A.D. 2004. The use of electronic tags to study fish movement: a case study with yellowtail flounder off New England. ICES CM 2004/K **81**.

Charles, C., Gillis, D.M., Hrenchuk, L.E., and Blanchfield, P.J. 2016. A method of spatial correction for acoustic positioning biotelemetry. Animal Biotelemetry **4**: 5. doi:10.1186/s40317-016-0098-3.

Chen, C., Beardsley, R.C., and Cowles, G. 2006. An unstructured grid, finite-volume coastal ocean model (FVCOM) system. Oceanography **19**(1): 78.

Chen, C., Huang, H., Beardsley, R.C., Xu, Q., Limeburner, R., Cowles, G.W., Sun, Y., Qi, J., and Lin, H. 2011. Tidal dynamics in the Gulf of Maine and New England Shelf: An application of FVCOM. J. Geophys. Res.-Oceans **116**(C12). doi:10.1029/2011JC007054.

Coleman, K.E. 2015. Understanding the winter flounder (*Pseudopleuronectes americanus*) southern New England/Mid-Atlantic stock through historical trawl surveys and monitoring cross continental shelf movement. MS thesis, Rutgers University, New Brunswick, NJ.

Cowles, G.W., Lentz, S.J., Chen, C., Xu, Q., and Beardsley, R.C. 2008. Comparison of observed and model-computed low frequency circulation and hydrography on the New England Shelf. J. Geophys. Res.-Oceans **113**(C9): C09015. doi:10.1029/2007JC004394.

Dean, M.J., Hoffman, W.S., Zemeckis, D.R., and Armstrong, M.P. 2014. Fine-scale diel and gender-based patterns in behaviour of Atlantic cod (*Gadus morhua*) on a spawning ground in the Western Gulf of Maine. ICES J. Mar. Sci. **71**(6): 1474–1489. doi:10.1093/icesjms/fsu040.

Egbert, G.D. and Erofeeva, S.Y. 2002. Efficient inverse modeling of barotropic ocean tides. J. Atmos. Oceanic Technol. **19**(2): 183–204. doi:10.1175/1520-0426(2002)019⟨0183:EIMOBO⟩2.0.CO;2.

Evans, K. and Arnold, G. 2009. Summary report of a workshop on geolocation methods for marine animals. In Tagging and tracking of marine animals with electronic devices, edited by J.L. Nielsen, H. Arrizabalaga, N. Fragoso, A. Hobday, M. Lutcavage, and J. Sibert, Springer Netherlands, number 9 in Reviews: Methods and Technologies in Fish Biology and Fisheries, pp. 343–363.

Fernö, A., Jørgensen, T., Løkkeborg, S., and Winger, P.D. 2011. Variable swimming speeds in individual Atlantic cod (*Gadus morhua* L.) determined by high-resolution acoustic tracking. Mar. Biol. Res. **7**(3): 310–313. doi:10.1080/17451000.2010.492223.

Galuardi, B. and Lam, C.H.T. 2014. Telemetry analysis of highly migratory species. In Stock Identification Methods (Second Edition), edited by S.X. Cadrin, L.A. Kerr, and S. Mariani, Academic Press, San Diego, pp. 447–476.

Goethel, D.R., Quinn, T.J., and Cadrin, S.X. 2011. Incorporating spatial structure in stock assessment: movement modeling in marine fish population dynamics. Rev. Fish. Sci. **19**(2): 119–136. doi:10.1080/10641262.2011.557451.

Grabowski, J., Sherwood, G.D., and Bank, C. 2013. Northeast regional monkfish tagging program: Additional archival tagging and otolith analyses to assess monkfish movements and age. Technical report. Final Report to the 2010 Monkfish Research Set Aside Program.

Gröger, J.P., Rountree, R.A., Thygesen, U.H., Jones, D., Martins, D., Xu, Q., and Rothschild, B.J. 2007. Geolocation of Atlantic cod (*Gadus morhua*) movements in the Gulf of Maine using tidal information. Fish. Oceanogr. **16**(4): 317–335. doi: 10.1111/j.1365-2419.2007.00433.x.

Hall, D.A. 2014. Conventional and radio frequency identification (RFID) tags. In Stock Identification Methods (Second Edition), edited by S.X. Cadrin, L.A. Kerr, and S. Mariani, Academic Press, San Diego, pp. 365–395.

Henderson, M.J. and Fabrizio, M.C. 2014. Small-scale vertical movements of summer flounder relative to diurnal, tidal, and temperature changes. Mar. Coast. Fish. **6**(1): 108–118. doi: 10.1080/19425120.2014.893468.

Hobson, V., Righton, D., Metcalfe, J., and Hays, G. 2009. Link between vertical and horizontal movement patterns of cod in the North Sea. Aquat. Biol. **5**: 133–142. doi: 10.3354/ab00144.

Howell, W.H., Morin, M., Rennels, N., and Goethel, D. 2008. Residency of adult Atlantic cod (*Gadus morhua*) in the western Gulf of Maine. Fish. Res. **91**(23): 123–132. doi: 10.1016/j.fishres.2007.11.021.

Hunt, J.J., Stobo, W.T., and Almeida, F. 1999. Movement of Atlantic cod, *Gadus morhua*, tagged in the Gulf of Maine area. Fish. Bull. **97**(4): 842–860.

Hunter, E., Aldridge, J.N., Metcalfe, J.D., and Arnold, G.P. 2003. Geolocation of free-ranging fish on the European continental shelf as determined from environmental variables - I. Tidal location method. Mar. Biol. **142**(3): 601–609. doi:10.1007/s00227-0984-5.

Hunter, E., Metcalfe, J.D., Holford, B.H., and Arnold, G.P. 2004. Geolocation of free-ranging fish on the European continental shelf as determined from environmental variables - II. Reconstruction of plaice ground tracks. Mar. Biol. **144**(4): 787–798. doi:10.1007/s00227-003-1242-1.

Jonsen, I., Basson, M., Bestley, S., Bravington, M., Patterson, T., Pedersen, M., Thomson, R., Thygesen, U., and Wotherspoon, S. 2013. State-space models for bio-loggers: A methodological road map. Deep Sea Res. Part II **88-89**: 34–46. doi:10.1016/j.dsr2.2012.07.008.

Kanwit, K., De Graaf, T., and Bartlett, C. 2008. Biological sampling, behavior and migration study of Atlantic halibut (*hippoglossus hippoglossus*) and cusk (*brosme brosme*) in the Gulf of Maine, year 2. Final report submitted to the Northeast Cooperative Research Partners Program, Maine Department of Marine Resources. Available at https://www1.maine.gov/dmr/science-research/species/documents/08halibutcusk.pdf. Accessed: 2016-8-12.

Kneebone, J., Hoffman, W.S., Dean, M.J., and Armstrong, M.P. 2014. Movements of striped bass between the exclusive economic zone and Massachusetts state waters. N. Am. J. Fish. Manage. **34**(3): 524–534. doi:10.1080/02755947.2014.892550.

Le Bris, A., Frechet, A., Galbraith, P.S., and Wroblewski, J.S. 2013a. Evidence for alternative migratory behaviours in the northern Gulf of St Lawrence population of Atlantic cod (*Gadus morhua L.*). ICES J. Mar. Sci. **70**(4): 793–804. doi:10.1093/icesjms/fst068.

Le Bris, A., Fréchet, A., and Wroblewski, J.S. 2013b. Supplementing electronic tagging with conventional tagging to redesign fishery closed areas. Fish. Res. **148**: 106–116. doi:10.1016/j.fishres.2013.08.013.

Loehrke, J.L. 2013. Movement patterns of Atlantic cod (*Gadus morhua*) spawning groups off New England. MS thesis, University of Massachusetts Dartmouth, Dartmouth, MA.

Metcalfe, J.D. and Arnold, G.P. 1997. Tracking fish with electronic tags. Nature **387**(6634): 665–666. doi:10.1038/42622.

Moser, J. and Shepherd, G.R. 2009. Seasonal distribution and movement of black sea bass (*Centropristis* striata) in the northwest Atlantic as determined from a mark-recapture experiment. J. Northwest Atl. Fish. Sci. **40**: 17–28.

NECOFS 2013. Northeast Coastal Ocean Forecasting System (NECOFS) Main Portal http://fvcom.smast.umassd.edu/necofs/. Accessed: 2016-05-20.

NEFSC 2013. 55th Northeast Regional Stock Assessment Workshop (55th SAW) Assessment Report. Technical report, Northeast Fisheries Science Center. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 13-11; 845 p. Available from: National Marine Fisheries Service, 166 Water Street, Woods Hole, MA 02543-1026, or online at http://www.nefsc.noaa.gov/nefsc/publications/.

Neuenfeldt, S., Hinrichsen, H.H., Nielsen, A., and Andersen, K. 2007. Reconstructing migrations of individual cod (*Gadus morhua L.*) in the Baltic Sea by using electronic data storage tags. Fish. Oceanogr. **16**(6): 526–535.

O'Donncha, F., Hartnett, M., Nash, S., Ren, L., and Ragnoli, E. 2015. Characterizing observed circulation patterns within a bay using HF radar and numerical model simulations. Journal of Marine Systems **142**: 96–110. doi:10.1016/j.jmarsys.2014.10.004.

Patterson, T., Thomas, L., Wilcox, C., Ovaskainen, O., and Matthiopoulos, J. 2008. State–space models of individual animal movement. Trends Ecol. Evol. **23**(2): 87–94. doi:10.1016/j.tree.2007.10.009.

Patterson, T.A., Basson, M., Bravington, M.V., and Gunn, J.S. 2009. Classifying movement behaviour in relation to environmental conditions using hidden Markov models. J. Anim. Ecol. **78**(6): 1113–1123. doi:10.1111/j.1365-2656.2009.01583.x.

Patterson, T.A., Parton, A., Langrock, R., Blackwell, P.G., Thomas, L., and King, R. 2016.

Statistical modelling of animal movement: a myopic review and a discussion of good practice. pre-print ArXiv:1603.07511 [q-bio, stat].

Pedersen, M.W. 2008. HMM Geolocation Toolbox homepage http://mwpedersen.dk/tracking.html. Accessed: 2016-05-20.

Pedersen, M.W., Patterson, T.A., Thygesen, U.H., and Madsen, H. 2011a. Estimating animal behavior and residency from movement data. Oikos **120**(9): 1281–1290. doi:10.1111/j.1600-0706.2011.19044.x.

Pedersen, M.W., Righton, D., Thygesen, U.H., Andersen, K.H., and Madsen, H. 2008. Geolocation of North Sea cod (*Gadus morhua*) using hidden Markov models and behavioural switching. Can. J. Fish. Aquat. Sci. **65**(11): 2367–2377.

Pedersen, M.W. 2007. Hidden Markov models for geolocation of fish. Ph.D. thesis, Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark.

Pedersen, M., Thygesen, U., and Madsen, H. 2011b. Nonlinear tracking in a diffusion process with a Bayesian filter and the finite element method. Computational Statistics & Data Analysis **55**(1): 280–290. doi:10.1016/j.csda.2010.04.018.

Potemra, J.T. 2012. Numerical modeling with application to tracking marine debris. Mar. Pollut. Bull. **65**(13): 42–50. doi:10.1016/j.marpolbul.2011.06.026.

Righton, D. and Mills, C. 2008. Reconstructing the movements of free-ranging demersal fish in the North Sea: a data-matching and simulation method. Mar. Biol. **153**(4): 507–521. doi:10.1007/s00227-007-0818-6.

Sibert, J.R., Hampton, J., Fournier, D.A., and Bills, P.J. 1999. An advection–diffusion–reaction model for the estimation of fish movement parameters from tagging data, with application to skipjack tuna (*Katsuwonus pelamis*). Can. J. Fish. Aquat. Sci. **56**(6): 925–938. doi:10.1139/f99-017.

Tallack, S.M. 2011. Stock identification applications of conventional tagging data for Atlantic cod in the Gulf of Maine. In American Fisheries Society Symposium, volume 76. volume 76, pp. 1–15.

Teo, S.L., Boustany, A., Blackwell, S., Walli, A., Weng, K.C., and Block, B.A. 2004. Validation of geolocation estimates based on light level and sea surface temperature from electronic tags. Mar. Ecol. Prog. Ser. **283**: 81–98.

Thorsteinsson, V., Pálsson, O.K., Tómasson, G.G., Jónsdóttir, I.G., and Pampoulie, C. 2012. Consistency in the behaviour types of the Atlantic cod: repeatability, timing of migration and geo-location. Mar. Ecol. Prog. Ser. **462**: 251–260. doi:10.3354/meps09852.

Thygesen, U.H. and Nielsen, A. 2009. Lessons from a Prototype Geolocation Problem. In Tagging and Tracking of Marine Animals with Electronic Devices, edited by J.L. Nielsen, H. Arrizabalaga, N. Fragoso, A. Hobday, M. Lutcavage, and J. Sibert, Springer Netherlands, number 9 in Reviews: Methods and Technologies in Fish Biology and Fisheries, pp. 257–276.

Thygesen, U.H., Pedersen, M.W., and Madsen, H. 2009. Geolocating fish using hidden Markov models and data storage tags. In Tagging and Tracking of Marine Animals with Electronic Devices, edited by J.L. Nielsen, H. Arrizabalaga, N. Fragoso, A. Hobday, M. Lutcavage, and J. Sibert, Springer Netherlands, number 9 in Reviews: Methods and Technologies in Fish Biology and Fisheries, pp. 277–293.

Trauth, M., Gebbers, R., Sillmann, E., and Marwan, N. 2007. MATLAB® Recipes for Earth Sciences. Springer Berlin Heidelberg.

Twomey, E. and Signell, R. 2013. Construction of a 3-arcsecond digital elevation model for the Gulf of Maine. Technical Open-File Report 2011-1127, U.S. Geological Survey. http://pubs.usgs.gov/of/2011/1127/.

Warner, J.C., Geyer, W.R., and Lerczak, J.A. 2005. Numerical modeling of an estuary: A comprehensive skill assessment. J. Geophys. Res. **110**(C5): C05001. doi:10.1029/2004JC002691.

Wilkin, J.L. 2006. The summertime heat budget and circulation of southeast New England shelf waters. Journal of Physical Oceanography **36**(11): 1997–2011. doi:10.1175/JPO2968.1.

Willmott, C.J. 1981. On the validation of models. Physical Geography **2**(2): 184–194. doi:10.1080/02723646.1981.10642213.

Winship, A.J., Jorgensen, S.J., Shaffer, S.A., Jonsen, I.D., Robinson, P.W., Costa, D.P., and Block, B.A. 2012. State-space framework for estimating measurement error from double-tagging telemetry experiments. Methods Ecol. Evol. **3**(2): 291–302. doi:10.1111/j.2041-210X.2011.00161.x.

Woillez, M., Fablet, R., Ngo, T.T., Lalire, M., Lazure, P., and de Pontual, H. 2016. A HMM-based model to geolocate pelagic fish from high-resolution individual temperature and depth histories: European sea bass as a case study. Ecol. Model. **321**: 10–22. doi: 10.1016/j.ecolmodel.2015.10.024.

Wright, P.J., Neat, F.C., Gibb, F.M., Gibb, I.M., and Thordarson, H. 2006. Evidence for metapopulation structuring in cod from the west of Scotland and North Sea. J. Fish Biol. **69**: 181–199. doi:10.1111/j.1095-8649.2006.01262.x.

Zemeckis, D.R., Hoffman, W.S., Dean, M.J., Armstrong, M.P., and Cadrin, S.X. 2014a. Spawning site fidelity by Atlantic cod (*Gadus morhua*) in the Gulf of Maine: implications for population structure and rebuilding. ICES J. Mar. Sci. doi:10.1093/icesjms/fsu117.

Zemeckis, D.R., Martins, D., Kerr, L.A., and Cadrin, S.X. 2014b. Stock identification of Atlantic cod (*Gadus morhua*) in US waters: an interdisciplinary approach. ICES J. Mar. Sci. **71**(6): 1490–1506. doi:10.1093/icesjms/fsu032.

Zemeckis, D., Liu, C., Cowles, G., Dean, M., Hoffman, W., Martins, D., and Cadrin, S. 2017. Seasonal movements and connectivity of an Atlantic cod *Gadus morhua* spawning component in the western Gulf of Maine. ICES J. Mar. Sci. doi:10.1093/icesjms/fsw190.

Zemeckis, D.R. 2016. Spawning dynamics, seasonal movements, and population structure of Atlantic cod (*Gadus morhua*) in the Gulf of Maine. PhD thesis, University of Massachusetts Dartmouth, Dartmouth, MA.

Figure captions

Figure 1 (a)Model domain, horizontal mesh, and bathymetry (m) of the Northeast Coastal Ocean Forecasting System (NECOFS). (b) Map of western Gulf of Maine, with the acoustic receiver arrays (inset) deployed within the Spring Cod Conservation Zone

Figure 2 Examples of the three activity levels identified in data from the archival data storage tags using the tidal fitting algorithm: a) low activity, b) moderate activity, and c) high activity. The shaded areas represent the 13 h window used to identify low activity periods and the 5 h window used to identify moderate activity periods.

Figure 3 Example of the likelihood functions based on temperature and depth $[L_{dt}(\hat{\boldsymbol{x}})]$ and modified with tidal exclusion $[L(\hat{\boldsymbol{x}})]$ for a given day.

Figure 4 Example of simulated tracks in the Gulf of Maine (GoM) and Georges Bank (GB) with duration of 40 (yellow), 120 (yellow and red), and 360 (yellow, red, and blue) days.

Figure 5 Areas of daily 95% utilization distribution determined from acoustic array detection of the high, moderate, and low activity levels determined by the likelihood model. Box plots show median values (red horizontal line), 25% and 75% percentile values (box outline), and the highest and lowest value within 1.5 times the interquartile range (whiskers).

Figure 6 Actual (star) and estimated (dot) locations of mooring tag deployments for tags a) #73; b) #84; and c) #71, in order of increasing location error.

Figure 7 Locations of the 10 double-electronic-tagged cod detected by the acoustic receivers (blue triangles) and the corresponding same-day estimates constructed by the revised (red dots) and original (open circles) HMM Geolocation Toolbox.

Figure 8 Daily location estimation error for the simulated experiments. Box plots show median values (horizontal line), 25% and 75% percentile values (box outline), outliers (diamonds), and the highest and lowest value within 1.5 times the interquartile range (whiskers).

Figure 9 The most probable track and the associated total posterior distribution for the double-electronic-tagged cod. The Spring Cod Conservation Zone (SCCZ, Fig. 1) is also shown (red rectangle).

Figure 10 Depth (blue line) and temperature (red line) time series recorded by DST and the activity classification (shading color, dark green: low, light green: moderate, white: high) for double-electronic-tagged cod nos. 12 and 13.

Table captions

Table 1	Comparisons of bottom temperature between NECOFS FVCOM predictions and survey measurements. NEFSC: NOAA Northeast Fisheries Science Center, MADMF: Massachusetts Division of Marine Fisheries, SMAST: School for Marine Science and Technology, UMass Dartmouth, IBS: Industry-Based Surveys.
Table 2	Experimental setup for the simulated tracks. GoM=Gulf of Maine; GB=Georges Bank; Summer=Aug 10, 2012; Winter=Jan 12, 2013
Table 3	Validation results for mooring tags and double-electronic-tagged cod.
Table 4	Summary of tagging and geolocation data for 10 double-electronic-tagged Atlantic cod. All tagged cod were released at 42.52° N, 70.70° W. MPT: most probable track

Table 1

Survey	Time	Number of measurements	Model-observation difference (°C)				
			Mean	S.D.	RMSE	Min	Max
NEFSC Bottom Trawl Survey	2009, 2014–2015	1 478	0.13	1.79	1.80	-6.58	7.53
NEFSC Shrimp Survey	2009 – 2013	361	-0.26	0.97	1.01	-4.30	2.05
MADMF Bottom Trawl Survey	2010 – 2015	1 299	-0.21	1.72	1.73	-7.44	4.66
SMAST Study Fleet	2003 – 2007	17 009	0.14	1.37	1.38	-10.73	8.84
SMAST 2010 Winter Flounder IBS	2010	336	0.62	1.57	1.68	-4.69	5.33
SMAST 2011 Winter Flounder IBS	2011	257	0.99	3.08	3.23	-4.77	6.32
SMAST 2012 Winter Flounder IBS	2012	159	-0.99	1.33	1.66	-5.12	0.90
SMAST Cod IBS	2003 – 2007	2 310	-0.43	0.98	1.07	-5.68	2.64
SMAST Video Survey	2013 – 2015	6 292	-0.40	2.09	2.12	-7.02	7.80
Total		29 501	-0.04	1.61	1.61	-10.73	8.84

Table 2

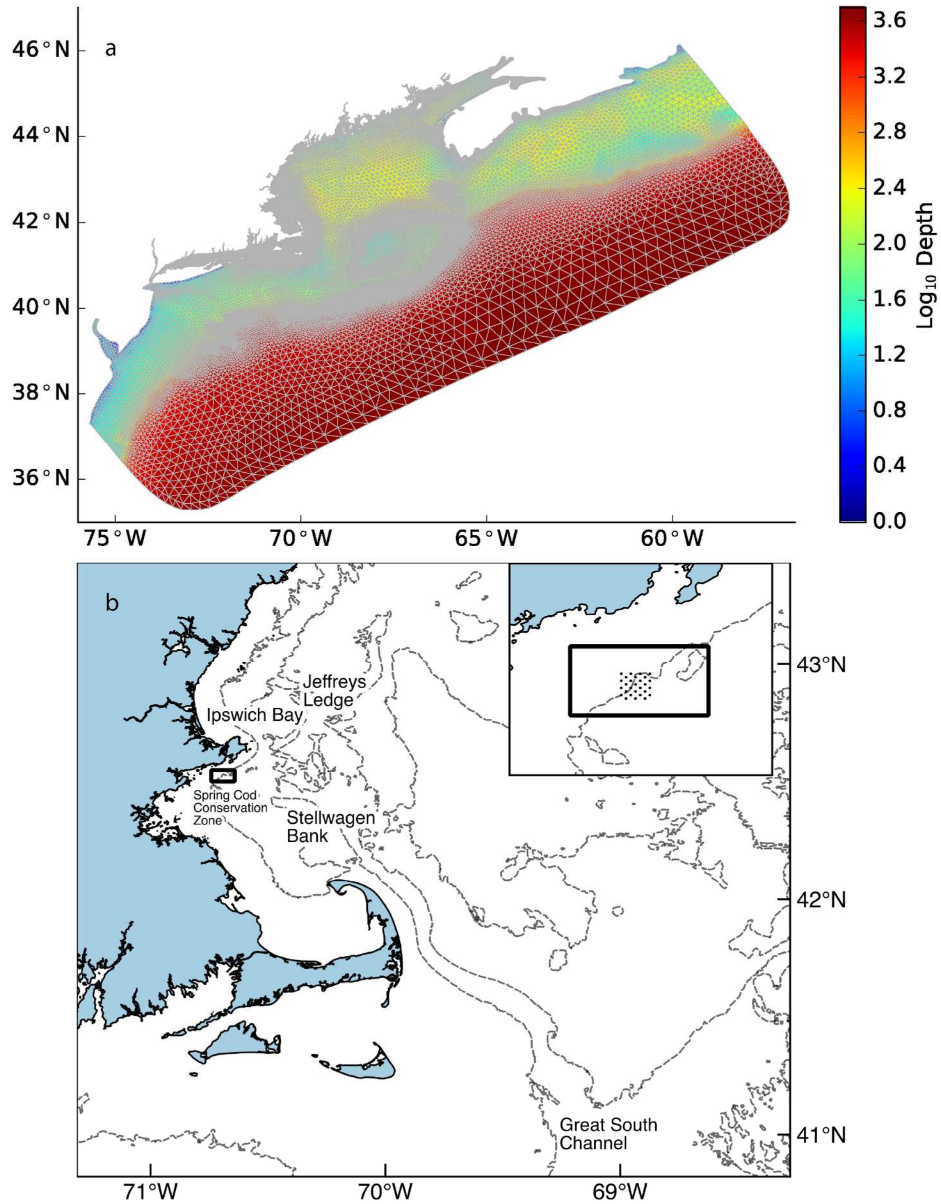
Set	Tag No.	Region	Season of release	Duration in water (d)
1	1–5	GoM	Summer	40
2	6–10	GoM	Summer	120
3	11–15	GoM	Summer	360
4	16–20	GoM	Winter	40
5	21–25	GoM	Winter	120
6	26–30	GB	Summer	40
7	31–35	GB	Summer	120
8	36–40	GB	Summer	360
9	41–45	GB	Winter	40
10	46–50	GB	Winter	120

Table 3

Experiment	Tag/fish No.	Deployment date	Deployment location	Days of data	Error range (km)	RMSE (km)	Median (km)	SD (km)	Mean normalized probability at know location(s)
a) Stationary	63	Jun 18, 2010	42.53° N, 70.70° W	31	3.10–12.12	9.22	9.02	1.81	0.84
	64	Apr 1, 2012	42.87° N, 70.60° W	33	0.76–5.25	3.27	2.88	1.22	1.00
	65	Apr 2, 2012	42.43° N, 70.68° W	17	5.85–11.11	8.94	9.35	1.81	0.84
	66	Apr 3, 2012	42.52° N, 70.69° W	27	0.46–4.63	2.58	2.44	1.81	0.77
	67	Apr 11, 2012	42.69° N, 70.43° W	23	0.73–4.88	2.96	2.34	1.25	0.86
	71	Aug 1, 2014	42.10° N, 70.08° W	126	3.23–25.51	22.86	23.10	2.62	0.80
	72	Apr 6, 2015	42.84° N, 70.27° W	36	1.13–21.38	18.28	18.88	4.23	0.65
	73	Apr 6, 2015	42.80° N, 70.27° W	36	0.58–4.24	2.09	1.86	0.84	1.00
	81	Apr 6, 2015	42.82° N, 70.27° W	113	0.14–6.09	3.42	3.16	1.40	0.52
	82	Apr 6, 2015	42.79° N, 70.32° W	134	3.47–8.20	6.03	5.83	1.19	0.30
	83	Sep 1, 2015	42.81° N, 70.29° W	43	1.65–6.75	4.54	4.79	1.39	0.77
	84	Sep 1, 2015	42.80° N, 70.31° W	43	1.35–6.81	4.85	4.59	1.37	0.94
	85	Sep 1, 2015	42.82° N, 70.27° W	43	0.22–5.97	3.49	3.29	1.39	0.94
	86	Sep 1, 2015	42.81° N, 70.27° W	43	0.28–5.22	2.83	2.64	1.24	0.72
	Total			**748**	**0.14–25.51**	**11.07**	**4.93**	**7.63**	**0.69**
b) Stationary (total, with original HGT)				**748**	**0.06–46.87**	**29.88**	**33.94**	**15.33**	**0.06**
c) Double-electronic-tagged	7	May 7, 2010	42.52° N, 70.70° W	15	1.08–19.27	6.51	3.12	4.89	0.74
	8	May 7, 2010	42.52° N, 70.70° W	17	1.87–25.95	13.89	13.25	5.14	0.26
	11	May 11, 2010	42.52° N, 70.70° W	16	6.52–31.35	18.11	15.65	8.55	0.61
	12	May 11, 2010	42.52° N, 70.70° W	36	6.75–58.20	42.17	44.97	18.51	0.54
	13	May 11, 2010	42.52° N, 70.70° W	34	1.18–57.32	42.37	48.41	22.10	0.47
	16	Jun 18, 2010	42.52° N, 70.70° W	26	0.39–7.41	3.16	2.19	1.68	0.79
	17	Jun 18, 2010	42.52° N, 70.70° W	23	0.55–12.21	8.55	8.33	2.31	0.41
	18	Jun 18, 2010	42.52° N, 70.70° W	14	0.61–4.37	2.40	2.22	1.04	0.39
	22	Jul 7, 2010	42.52° N, 70.70° W	8	18.28–134.38	97.44	91.45	43.53	0.34
	24	May 20, 2011	42.52° N, 70.70° W	35	6.11–12.87	8.95	8.85	1.75	0.83
	Total			**223**	**0.38–97.27**	**21.87**	**6.45**	**16.69**	**0.47**
d) Double-electronic-tagged (total, with original HGT)				**223**	**0.59–51.70**	**32.76**	**34.80**	**15.43**	**0.06**

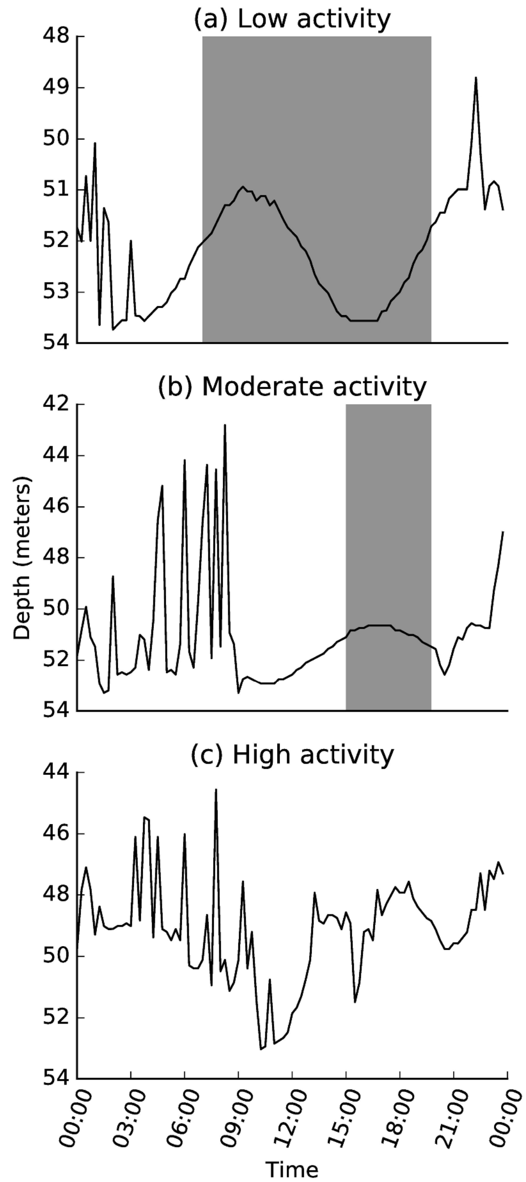
Table 4

Fish No.	DMF Fish ID	Tag No.	Release Date	Recapture				Days at large	Displacement distance (km)	Length of estimated movement (MPT, km)	Error of estimated recapture location (km)	# days of low activity	# days of moderate activity	Average movement rate (km/day)
				Date	Latitude (°)	Longitude (°)	Uncertainty (km)							
7	156	S11951	May 7, 2010	May 26, 2010	42.40 N	70.37 W	15	19	29.28	87.93	8.38	11	7	4.63
8	157	S11938	May 7, 2010	Jun 4, 2010	42.41 N	70.37 W	15	28	29.09	112.15	8.52	18	5	4.01
11	173	S11971	May 11, 2010	Jun 8, 2010	42.40 N	70.38 W	15	28	29.07	111.94	8.09	17	4	4.00
12	172	S11974	May 11, 2010	Jul 6, 2010	42.26 N	70.55 W	30	56	30.97	208.16	28.18	29	12	3.72
13	175	S11976	May 11, 2010	Nov 21, 2010	42.32 N	70.25 W	15	194	42.47	586.87	6.80	152	32	3.03
16	229	S12060	Jun 18, 2010	Jul 15, 2010	42.51 N	70.63 W	15	27	4.97	86.48	2.22	18	9	3.20
17	230	S12061	Jun 18, 2010	Aug 29, 2010	42.95 N	70.37 W	15	72	55.16	276.52	14.25	27	33	3.84
18	231	S12059	Jun 18, 2010	Sep 18, 2010	42.15 N	69.86 W	15	92	80.62	487.52	16.00	44	33	5.30
22	242	S12068	Jul 7, 2010	Feb 4, 2011	42.98 N	69.92 W	30	212	81.07	974.48	34.33	24	42	4.60
24	282	S11845	May 20, 2011	Jul 26, 2011	42.71 N	70.25 W	15	67	41.87	274.96	8.68	45	18	4.10



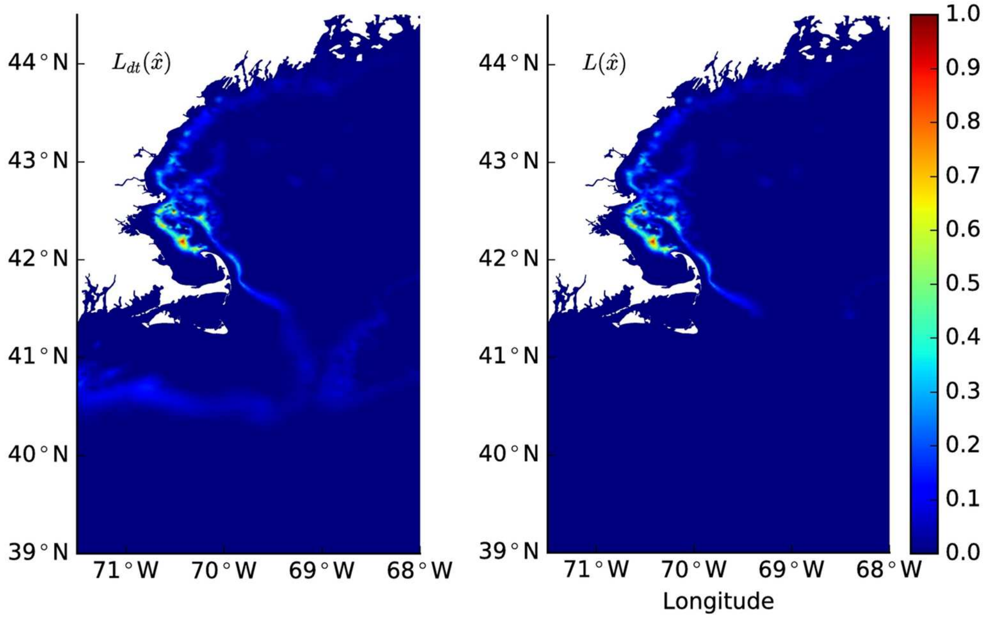
(a)Model domain, horizontal mesh, and bathymetry (m) of the North- east Coastal Ocean Forecasting System (NECOFS). (b) Map of west- ern Gulf of Maine, with the acoustic receiver arrays (inset) deployed within the Spring Cod Conservation Zone

237x306mm (300 x 300 DPI)



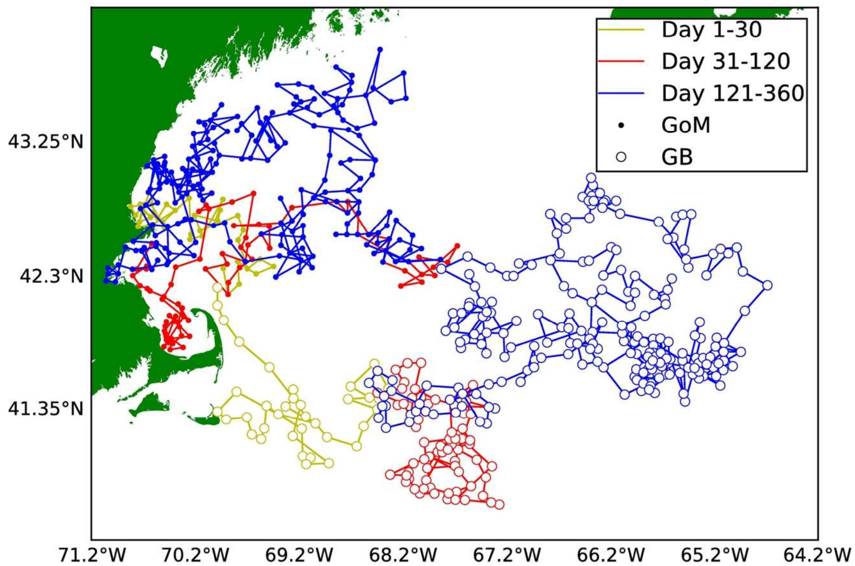
Examples of the three activity levels identified in data from the archival data storage tags using the tidal fitting algorithm: a) low activity, b) moderate activity, and c) high activity. The shaded areas represent the 13 h window used to identify low activity periods and the 5 h window used to identify moderate activity periods.

235x521mm (300 x 300 DPI)



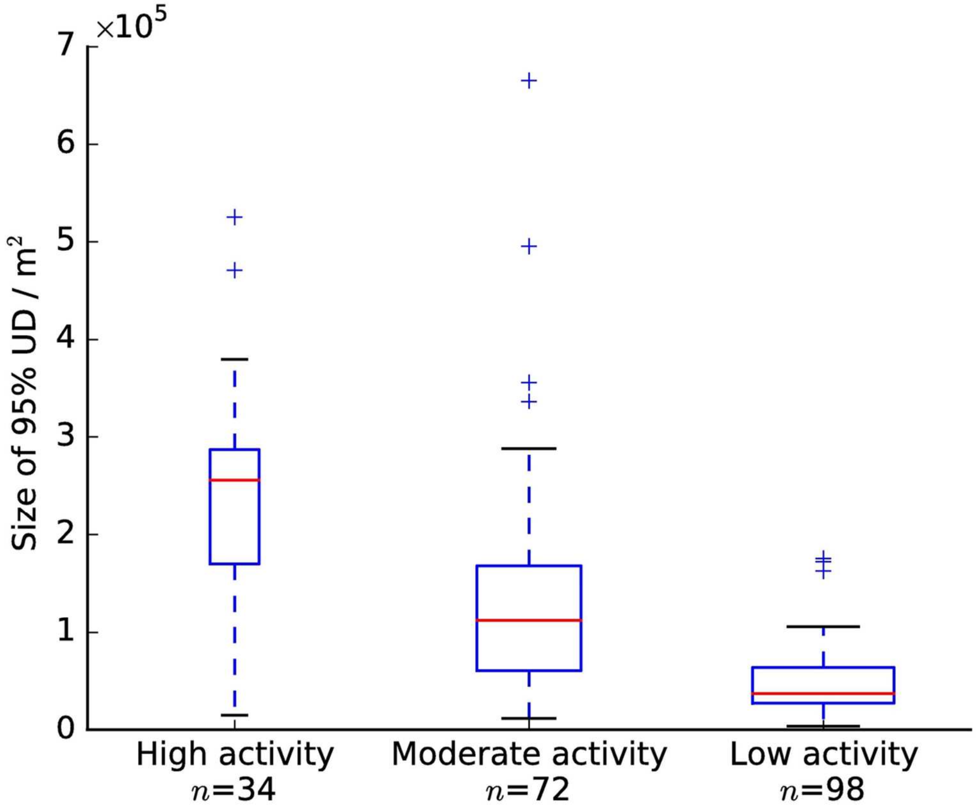
Example of the likelihood functions based on temperature and depth [Ldt(x)] and modified with tidal exclusion [L(x)] for a given day.

119x76mm (300 x 300 DPI)



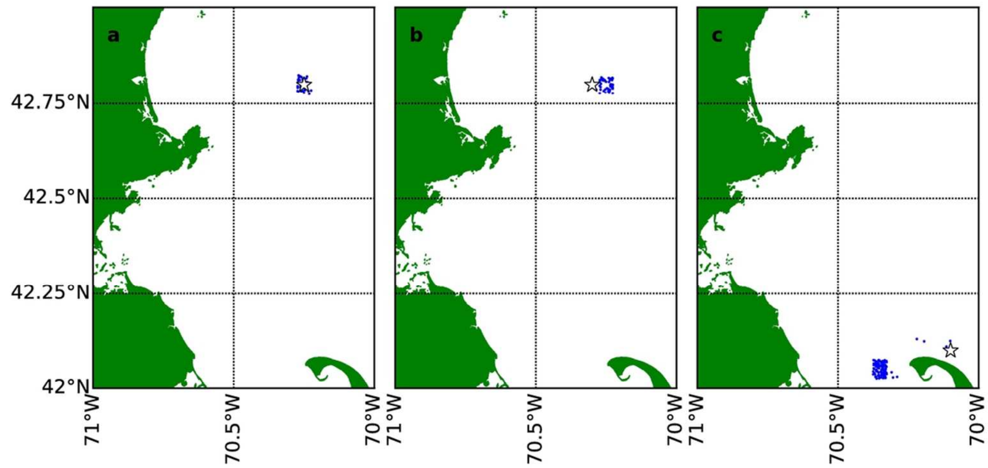
Example of simulated tracks in the Gulf of Maine (GoM) and Georges Bank (GB) with duration of 40 (yellow), 120 (yellow and red), and 360 (yellow, red, and blue) days.

152x101mm (300 x 300 DPI)



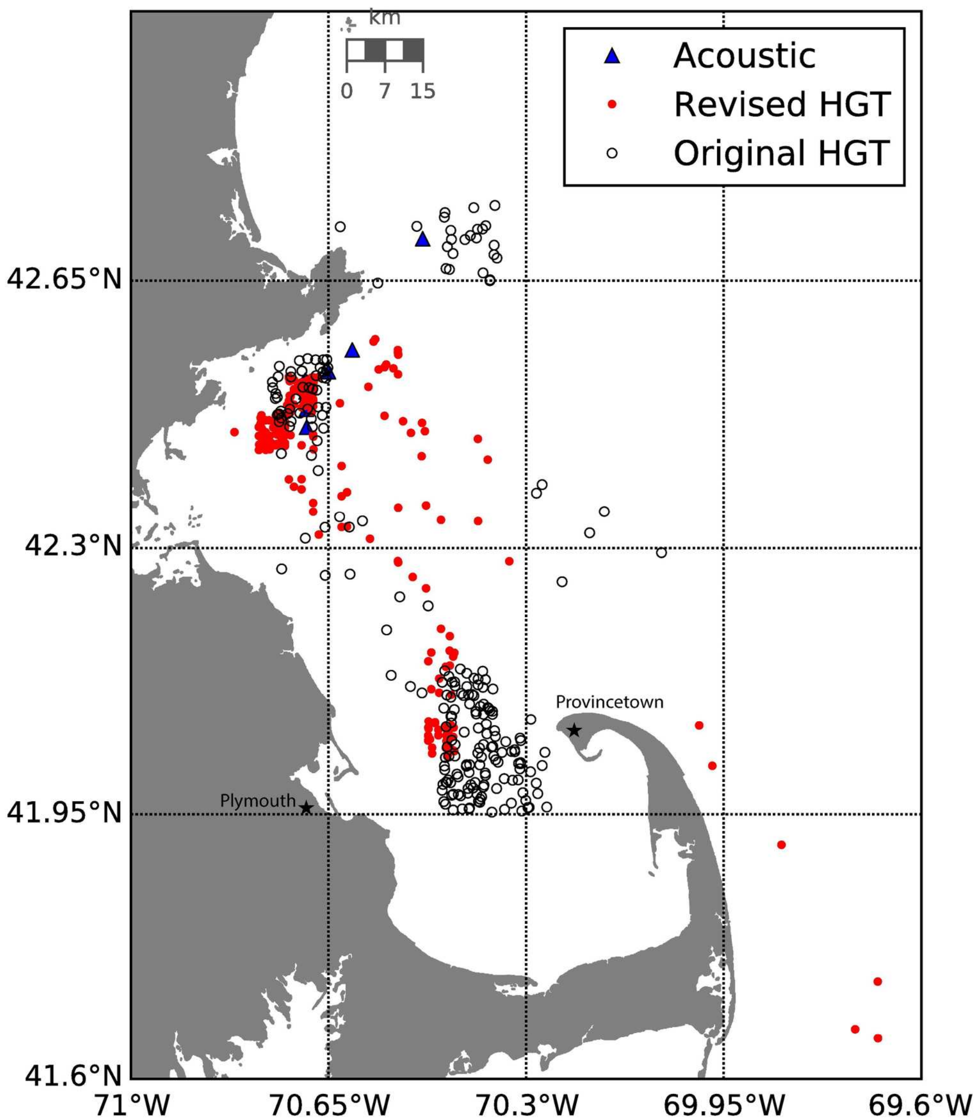
Areas of daily 95% utilization distribution determined from acoustic array detection of the high, moderate, and low activity levels determined by the likelihood model. Box plots show median values (red horizontal line), 25% and 75% percentile values (box outline), and the highest and lowest value within 1.5 times the interquartile range (whiskers).

111x93mm (300 x 300 DPI)



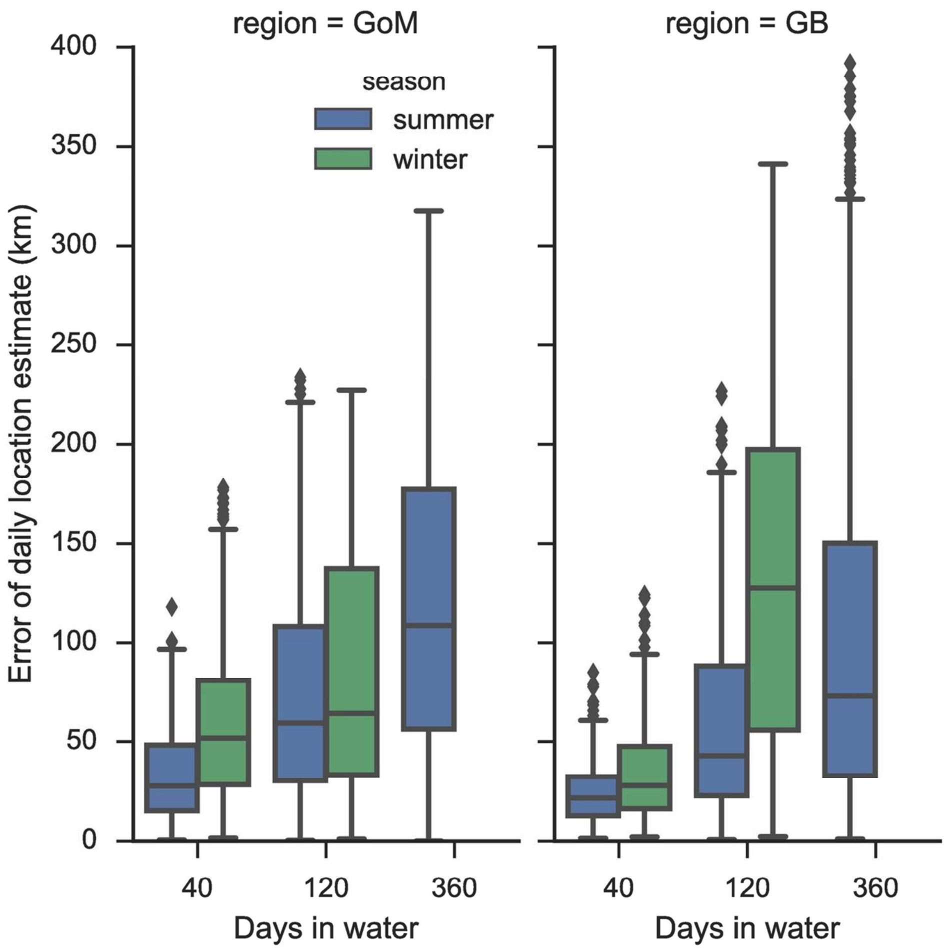
Actual (star) and estimated (dot) locations of mooring tag deployments for tags a) #73; b) #84; and c) #71.

105x50mm (300 x 300 DPI)



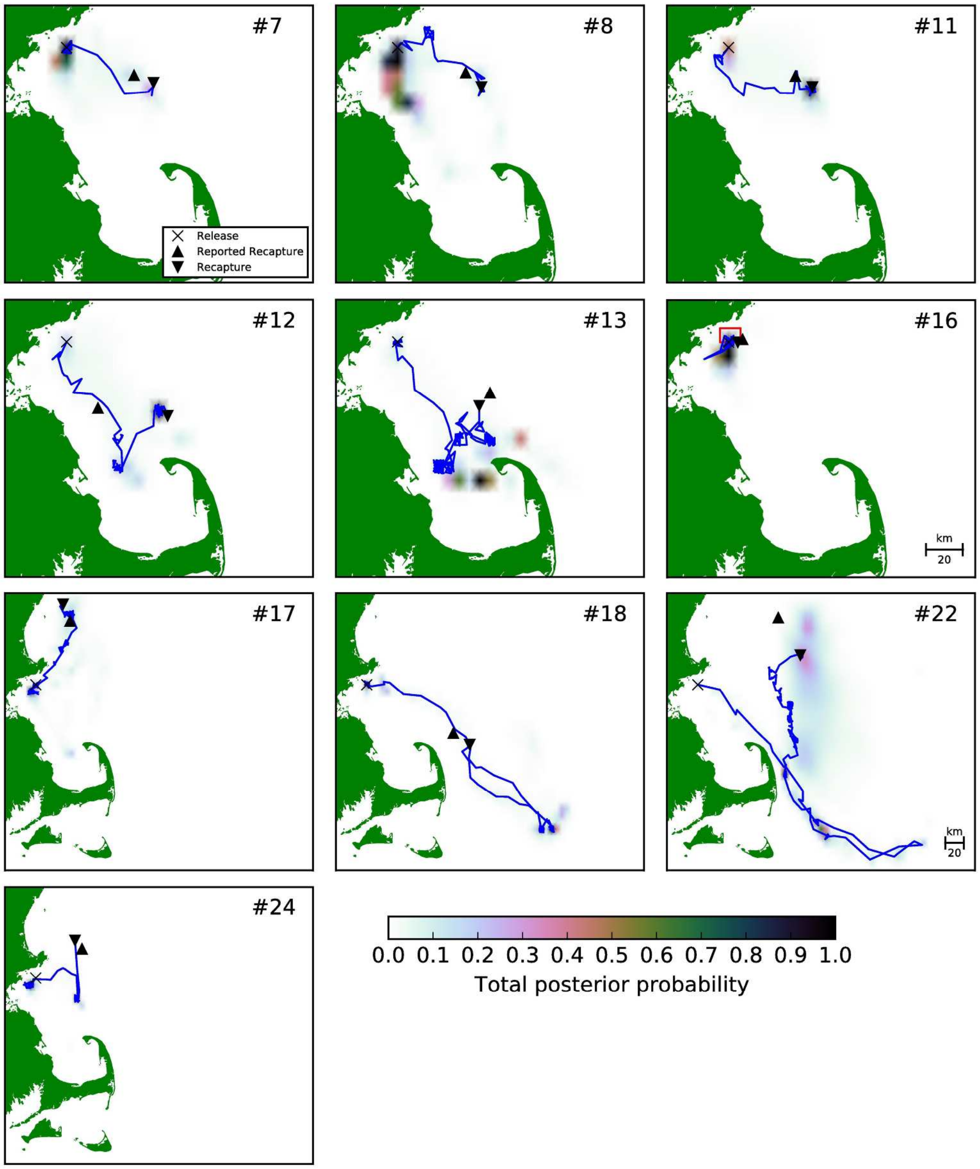
Locations of the 10 double-electronic-tagged cod detected by the acoustic receivers (blue triangles) and the corresponding same-day estimates constructed by the revised (red dots) and original (open circles) HMM Geolocation Toolbox.

171x196mm (300 x 300 DPI)



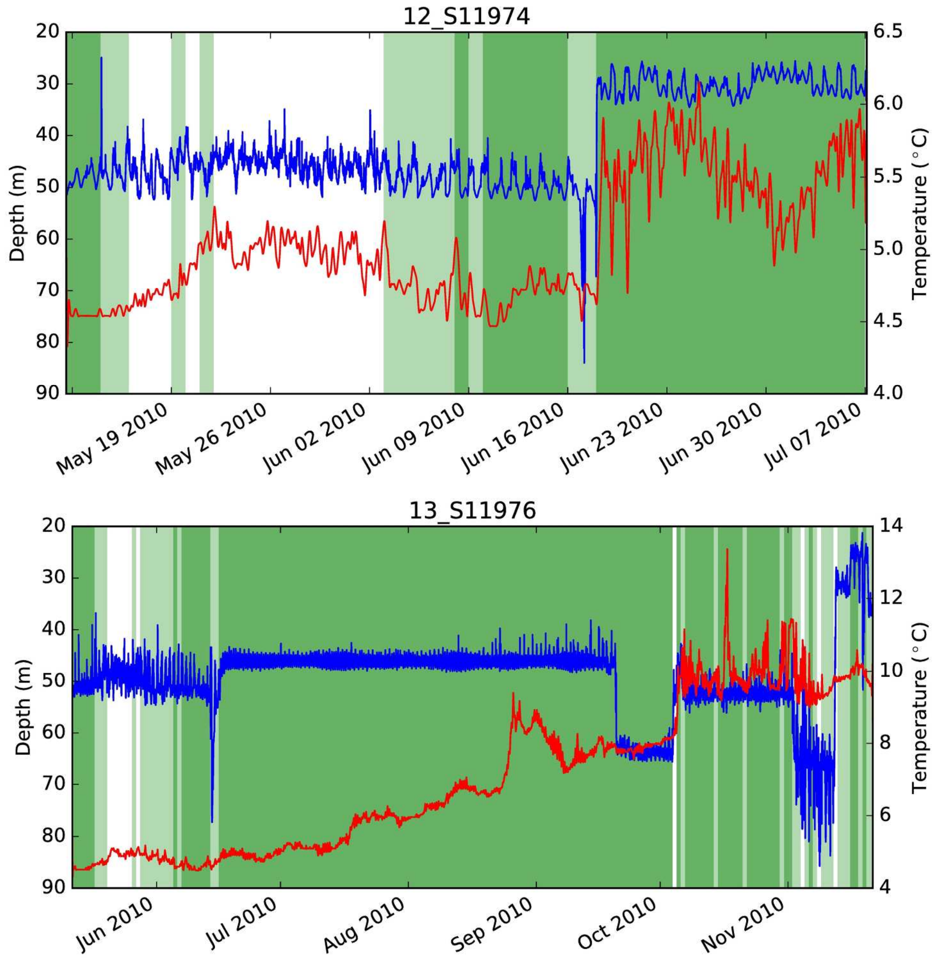
Daily location estimation error for the simulated experiments. Box plots show median values (horizontal line), 25% and 75% percentile values (box outline), outliers (diamonds), and the highest and lowest value within 1.5 times the interquartile range (whiskers).

127x127mm (300 x 300 DPI)



The most probable track and the associated total posterior distribu- tion for the double-electronic-tagged cod. The Spring Cod Conserva- tion Zone (SCCZ, Fig. 1) is also shown (red rectangle).

239x286mm (300 x 300 DPI)



Depth (blue line) and temperature (red line) time series recorded by DST and the activity classification (shading color, dark green: low, light green: moderate, white: high) for double-electronic-tagged cod nos. 12 and 13.

214x208mm (300 x 300 DPI)

Supplementary material for Liu et al. CJFAS

Table S1: Validation results for mooring tags and double-electronic-tagged cod using original HGT.

Experiment	Tag/fish No.	Deployment date	Deployment location	Days of data	Error range (km)	RMSE (km)	Median (km)	SD (km)	Mean normalized probability at know location(s)
	63	Jun 18, 2010	42.53° N, 70.70° W	31	25.27–35.74	30.78	30.39	3.44	0.08
	64	Apr 1, 2012	42.87° N, 70.60° W	33	2.87–14.12	8.88	8.62	3.15	0.03
	65	Apr 2, 2012	42.43° N, 70.68° W	17	0.48–8.55	4.90	5.12	2.39	0.02
	66	Apr 3, 2012	42.52° N, 70.69° W	27	0.91–7.18	3.87	3.06	1.67	0.08
	67	Apr 11, 2012	42.69° N, 70.43° W	23	0.58–6.83	4.61	4.39	1.88	0.08
	71	Aug 1, 2014	42.10° N, 70.08° W	126	18.13–41.87	36.61	36.27	3.07	0.02
	72	Apr 6, 2015	42.84° N, 70.27° W	36	13.62–24.05	19.16	19.01	2.64	0.02
a) Stationary	73	Apr 6, 2015	42.80° N, 70.27° W	36	0.89–8.47	5.17	4.88	1.87	0.02
	81	Apr 6, 2015	42.82° N, 70.27° W	113	13.36–46.87	40.98	41.05	4.11	0.10
	82	Apr 6, 2015	42.79° N, 70.32° W	134	9.45–41.69	36.19	36.14	3.75	0.10
	83	Sep 1, 2015	42.81° N, 70.29° W	43	0.06–6.79	3.94	3.87	1.72	0.03
	84	Sep 1, 2015	42.80° N, 70.31° W	43	0.34–6.28	3.78	3.71	1.44	0.04
	85	Sep 1, 2015	42.82° N, 70.27° W	43	12.76–23.50	18.60	18.28	3.17	0.04
	86	Sep 1, 2015	42.81° N, 70.27° W	43	17.32–46.10	40.38	40.98	4.62	0.10
	Total			**748**	**0.06–46.87**	**29.88**	**33.94**	**15.33**	**0.06**
	7	May 7, 2010	42.52° N, 70.70° W	15	1.07–7.50	5.04	4.34	1.79	0.07
	8	May 7, 2010	42.52° N, 70.70° W	17	28.60–66.93	59.33	61.23	10.10	0.06
	11	May 11, 2010	42.52° N, 70.70° W	16	1.64–55.12	36.04	28.08	22.97	0.08
	12	May 11, 2010	42.52° N, 70.70° W	36	2.80–66.40	52.63	51.49	13.92	0.10
	13	May 11, 2010	42.52° N, 70.70° W	34	2.06–66.46	50.06	54.09	22.33	0.06
c) Double-electronic-tagged	16	Jun 18, 2010	42.52° N, 70.70° W	26	2.86–37.79	30.26	31.07	6.78	0.09
	17	Jun 18, 2010	42.52° N, 70.70° W	23	12.95–68.62	59.48	61.00	11.17	0.06
	18	Jun 18, 2010	42.52° N, 70.70° W	14	1.20–5.83	4.03	4.29	1.49	0.06
	22	Jul 7, 2010	42.52° N, 70.70° W	8	4.13–57.32	40.61	40.78	21.01	0.06
	24	May 20, 2011	42.52° N, 70.70° W	35	21.58–71.15	60.94	63.66	13.69	0.03
	Total			**223**	**0.59–51.70**	**32.76**	**34.80**	**15.43**	**0.06**

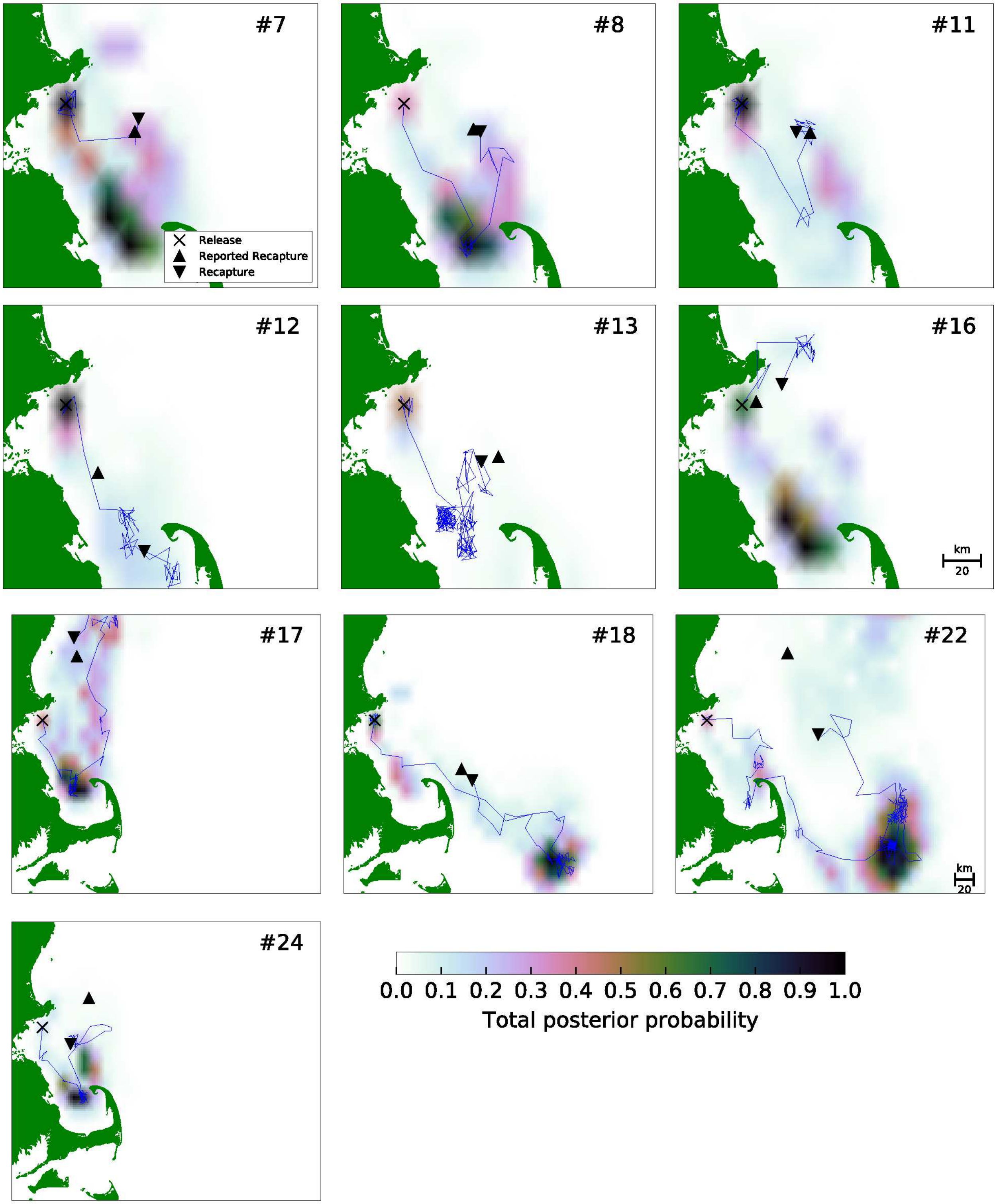


Figure S1: The most probable track and the associated total posterior distribution for the double-electronic-tagged cod, using original HGT.