This article is the manuscript version of:

Ahmadian, M., Boone, E. L., Chiu, G. S., & Weng, K. C. (2026). A new perspective to fish trajectory imputation: a spatiotemporal probability model for simulating acoustically tagged fish movement. Journal of Applied Statistics: Environmental Statistics and Data Science, 1–21. https://doi.org/10.1080/29984688.2026.2623273

A New Perspective to Fish Trajectory Imputation: A Spatiotemporal Probability Model for Simulating Acoustically Tagged Fish Movement

Mahshid Ahmadian[a,b], Edward L. Boone[a], Grace S. Chiu[b,a], and Kevin C. Weng[b]

[a]Virginia Commonwealth University, USA; [b]William & Mary's Batten School of Coastal & Marine Sciences, Virginia Institute of Marine Science, USA

ABSTRACT

This paper develops a probability model for understanding, interpolating, and predicting fish movement patterns based on spatiotemporal data recorded by spatially static acoustic receivers. Unlike tracking technology with continuous or regular spatial coverage while emitting signals from the animal's location at regular time intervals, acoustic receivers are irregularly spaced stationary sensors that record signals only when the animal is nearby a sensor and sets it off. Thus, for periods of time, the fish can evade detection but still be within the spatial domain of interest, resulting in the absence of location data. This poses challenges to the understanding of fish movement patterns, and hence, the identification of existing statistical inference frameworks for modeling trajectories that must not cross land barriers or the detection range of any receiver that did not record fish signals. As the foundation for an inferential methodology, in this paper, we derive and implement a model-based simulation strategy that relies on both Markov chain and random-walk principles to enhance our dataset over time. This methodology will be generalizable and applicable to all fish species with similar migration patterns or data with similar structures due to the use of static acoustic receivers.

KEYWORDS

Animal movement, Missing not at random, Prior predictive simulation

1. Introduction

Significant emphasis has been put on the study and modeling of spatiotemporal data over the years. In particular, modeling animal movement over time across a spatial landscape has gained considerable attention. For example, investigating the displacements of marine animals in water is crucial for advancing marine science, particularly understanding the population dynamics of various fish species and the effects on environmental and ecosystem conditions due to fishing practices. The structure of datasets in this type of research highly depends on the data collection method. Indeed, a statistical methodology that can generate useful predictions has been lacking for complex data structures like the one in the present study. This poses challenges that are not readily addressed by common methods, as we explain here and in section 2.

CONTACT Mahshid Ahmadian. Email: ahmadianm@vcu.edu

Previous research has employed diverse approaches for modeling animal movement data, such as agent-based models (*6*, *15*, *30*, *32*, *40*), Brownian bridge movement models (*12*, *31*), Lévy flight models (*14*, *38*), Markov models (*26*, *34*), hidden Markov models (*21*, *27*) and other state-space models (*8*, *19*, *20*, *28*, *29*), generalized additive models (*5*), and functional data analysis (*9*). In addition, there is a wealth of research dedicated to statistical methods specifically for studying animal movement using data that are not simultaneously sparse in space and in time (*11*, *13*). Each approach has strengths and limitations. For example, ABMs have been widely used, especially by ecologists and environmental scientists, to simulate the actions and interactions of individual agents (fish, in our case), and assess their effects on the population. Brownian bridge models involve continuous-time random walks between observed locations. Lévy flight models extend the random walks to incorporate a heavy-tailed distribution for each defined step. State-space models use latent and observed states probabilistically. They are often used in time series forecasting, economic indicators, and tracking and navigation. These models have been used in the study of animal movement because they are very effective in handling incomplete data, where animals are partially observed through receivers or satellites. Moreover, state-space models known as hidden Markov models extend the Markov model by incorporating hidden states that represent unobserved steps in an animal movement process. Generalized additive models are used to investigate complex animal behaviors and movement patterns by offering a flexible approach to modeling non-linear relationships between animal movement and environmental factors. However, many of these models are developed for satellite telemetry with regularly recorded locations (*25*, *35*), and they are less suited to our acoustic telemetry dataset. For our data (described in detail in section 2), it is unclear how effectively any of these approaches can address the particular challenges posed by our underwater acoustic telemetry network.

An underwater acoustic telemetry network is made up of spatially static receivers (each about the size of a tube of three tennis balls); they are fixed at sparse and highly irregularly spaced geographical locations chosen due largely to the convenience for ecologists to attach them to pre-existing structures (e.g., under bridges and docks). The set-up is akin to stationary motion sensor networks on land, in which presence-absence data are observed at the sensor level only, although in our case, we have the additional complication of land barriers that are impenetrable by fish. It is important to note that sparse receiver arrays are a fundamental characteristic in most acoustic telemetry systems across both freshwater and marine environments. The sparsity arises from the high cost of acoustic receivers, logistical constraints in deployment and maintenance, and the vast spatial scales of aquatic habitats relative to receiver detection ranges (*3*). Location data that result from such an irregular spatial network, along with highly irregular detection patterns over time, require us to limit the fish's imputed movement at any given time step to locations away from both land and any receiver that never detected the fish at that time step. Contrast this with many popular methods for imputing unobserved locations in an animal's trajectory, that rely on observed locations in the form of radio signals (e.g., satellite) that (a) are emitted (or pinged) at time points on a regular temporal grid and (b) cover a large swath of continuous space. Thus, it is expected that the animal's location is observable for each ping, and the absence of the detection of a ping indicates, perhaps, equipment malfunction (i.e., missingness is reasonably at random) or that the animal is outside of the entire spatial domain of radio monitoring. However, in our case, no *a priori* temporal grid exists because the detection of pings is not expected when the animal is outside of any receiver's detection range, so that the absence of detection at any given time point simply suggests that the fish is beyond the detection range of

any pre-installed sensor, but still can be anywhere else within the spatial domain of interest. Specifically, the absence of detections is imposed by the nature of the equipment and can be entirely unrelated to malfunction. As such, the occasional detections (each accompanied by non-detections at any other receiver), the possibility for a fish to be located anywhere within the spatial domain of interest in between detections, and land barriers together require models that are tailored to address such complexities.

A few studies have explored the challenges in interpreting and modeling acoustic telemetry data. This includes the investigation of the variation in data sparsity (*29*), the application of an inhomogeneous spatial point process model to determine the activity centers of aquatic animals (*41*), the employment of a generalized linear mixed model to evaluate detection range variations of acoustic telemetry arrays (*2*), and the use of a state-space model tailored to reflect home range behavior exhibited by some coastal fish species (*1*, *8*, *29*). However, our species of interest does not exhibit any home range behavior at our spatial scale of interest.

The purpose of this paper is to propose a new probability modeling framework based on the simple concept of random walk for imputing fish trajectories, while (a) tailoring the random-walk framework particularly to acoustic sensor network data on a fish species that exhibits no home range behavior, and (b) accounting for the ecological understanding that an animal may know where it is heading (*7*), so its first and last positions, and the angle between these, should not be ignored when imputing locations of the traveling fish. For our proposed approach, three key components in the movement model are (1) the simple notion of random-walk movement at each modeler-defined time step, (2) the updated direction of fish movement at each time step, and (3) the remaining distance until the next observed detection. Importantly, feature (2) emphasizes the notion of an "empirically determined movement path" associated with each individual fish (*30*), a crucial aspect of realism for a highly mobile species whose locations are largely unknown outside of empirical records. Note that this paper considers each fish individually to illustrate the trajectory simulated using our approach. Similar to ABMs and mathematical models in general, simulations allow direct input from domain scientists (in our case, fisheries scientists) towards model refinement. Our current paper serves as the foundation for future Bayesian inference at the population level, i.e., a model for multiple fish trajectories using random effects. Therefore, instead of simulating fish trajectories by fixing model parameters, we conduct prior predictive simulations, in which each model parameter is simulated from a prior distribution that would be employed in the context of Bayesian inference.

The organization of this paper is as follows: Section 2 introduces the dataset that we used for this study and it highlights the challenges posed by the dataset structure. Section 3 delves into the model on which our imputation algorithm is based, followed by section 4 which discusses the joint probability distribution associated with our method. In section 5, we present a simulation study and the results of the proposed model. In section 6, we compare the output of our method with the popular Browning bridge movement model, and section 7 discusses a few limitations of our proposed approach. Finally, section 8 outlines the next steps and potential directions for future research.

2. Data

Methods for analyzing environmental data may vary across datasets, influenced not only by the nature of the scientific problem but also by the methods of data collection. Our research is tailored to a specific dataset by (*5*) that provides information on fish

movement along the US East Coast. The data, collected from 2016 to 2020, utilize acoustic tagging of a species known as cobia (*Rachycentron canadum*) in the Atlantic waters and include details such as the fixed coordinates of the spatially static acoustic receivers and the corresponding time stamp of the detection of a tagged fish. Studies on cobia have employed various techniques to understand their movement patterns and habitat utilization. Pop-up satellite tags have been used to assess the post-release survival, net travel distance, and habitat utilization of cobia in Virginia waters, providing valuable data on their movements and behavior (*18*). However, the study of fish movement by acoustic tagging technology has become very common. Compared to pop-up satellite archival tags which are more suited to large spatial movement studies alongside depth and temperature measurements, acoustic tag receivers can offer tracking of fish movements at a finer spatial resolution, making them a better choice for detailed, long-term studies of fish such as cobia, which are expected to swim near certain coastal areas where receivers can be readily deployed. Acoustic tags emit a sound signal that is recognizable by the underwater receivers, helping ecologists track tagged marine animals near the receivers, enabling tracking where satellite methods fail (*16*). However, this technology has its challenges, which include high costs, behavioral impacts on animals, data gaps and missingness caused by reasons such as environmental noise or tag failure (*4*, *10*, *39*), and the sparse placement (often due to high costs) as well as static nature of receivers being used to monitor highly mobile animals. Moreover, studies have shown that the detection power of these receivers will decrease as the tagged fish moves farther away from the receiver, even when it remains within the receiver's purported range of detection (*17*, *33*, *36*, *37*). Research also suggests that the detection capability of acoustic receivers is influenced by environmental conditions (*24*, *36*). However, the present study focuses on a small spatial domain (Figure 1) unaccompanied by explicit data on environmental conditions.

In the case of our cobia dataset, the marine science team caught the fish and implanted each with an acoustic tag. The fish was then released into the water, in which acoustic telemetry receivers had been pre-installed at convenient locations to form an irregularly spaced detection network. In particular, the locations were strategically chosen by different research groups at locations that were convenient for installation but also previously identified by local fishermen as having a high likelihood of fish presence (*42*). Each receiver was assumed to have a detection radius of 500 meters, which is based on many range test studies (*22*) and considered by fisheries scientists to be a reasonable approximation in practice. To further enhance realism, in section 3, we model the detection probability to reflect decreasing detection power as the distance from the receiver increases.

In addition to being highly irregular in space, the dataset is made up of temporally irregular detections recorded at any time of the day, leading to a varying number of daily detections from none, one, or several for any given fish. Temporal irregularity is a consequence of the fact that each receiver is only activated when a fish happens to set it off by moving within its detection range. In this paper, we focus on cobia movements around all 46 acoustic receivers positioned in the Chesapeake Bay, Virginia, with non-overlapping detection radii (Figure 1). By restricting our attention to this geographical region, we can target our model at fish movements over shorter distances between detections. In addition, we propose our model by considering daily time steps. For a day on which a fish was detected more than once by any number of receivers, we take its final detection on that day as its assigned location. The temporal scale with daily time steps is chosen because, for this dataset, the number of consecutive time steps without detections is lower (thus, more manageable) than for other similarly convenient

temporal scales such as hourly and weekly.

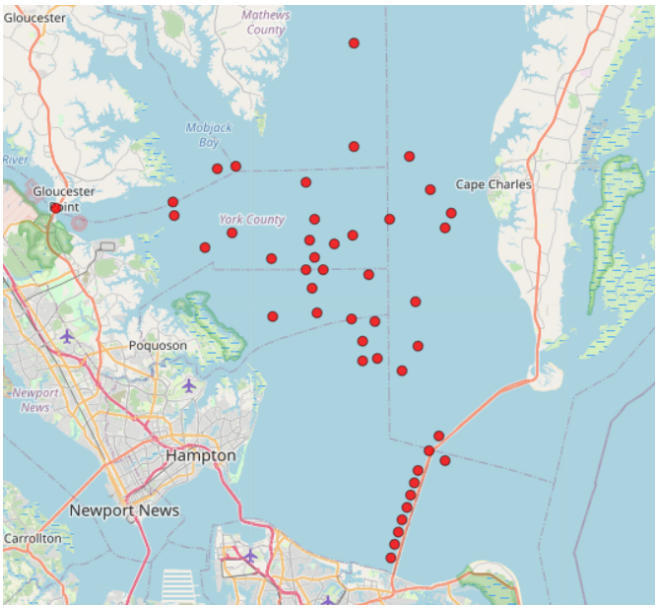


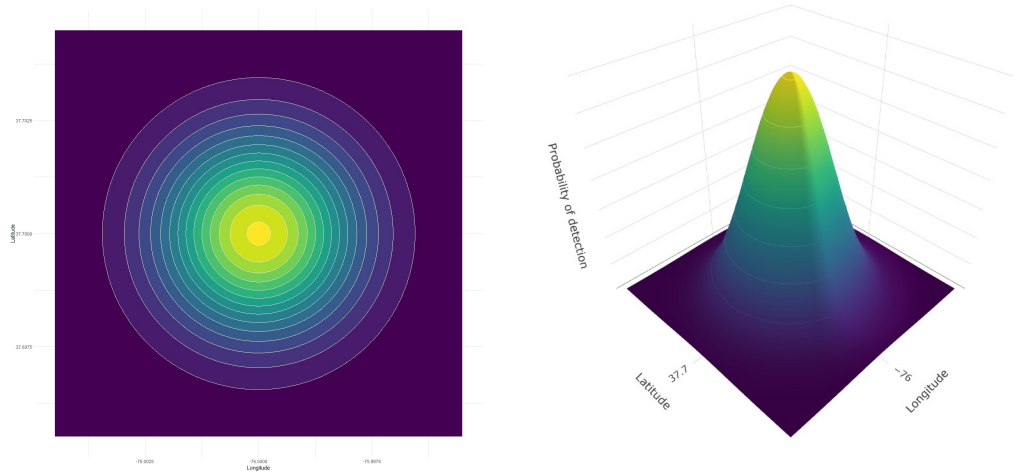
Figure 1. Locations of irregularly spaced receivers, marked by red circles, overlaid on a geographic map of Chesapeake Bay, Virginia. Map data: OpenStreetMap contributors

3. Probability Model

It is important to emphasize that the only spatial information available on each fish is whether or not it was in a receiver's detection range, and hence, instead of the precise location of the individual fish, only the spatial coordinates of the receivers around which the fish was detected are available. Furthermore, this sensing regime implies that no information is available when the fish is swimming outside the detection area of the network of installed receivers. The complexity of the acoustic tagging data, as a result, creates a two-fold problem: the unknown location of any detected fish within the detection radius of the activated receiver, and a high abundance of temporal gaps between detections of the fish. We derive and implement a probability model that incorporates Markov random-walk principles with possible behavioral characteristics (e.g., possible knowledge of destination γ) to enhance the comprehensiveness of our dataset over time. It is also important to note that the following methodology is derived at the level of individual fish to facilitate input from fisheries scientists who are knowledgeable about individual-level movements.

Thus, let us consider the movement trajectory of one fish, and let $\mathbf{X}_{jt}$ be the bivariate vector of longitude and latitude of the location of fish j at time step t, where $j = 1, \cdots, J$, and $t = 1, 2, 3, \cdots$. Further, let τ_{N_j} be the set of time steps in which fish j is observed by receiver i, and τ_{M_j} be the set of time steps when the fish is out of detection range ($\tau_{N_j} \cup \tau_{M_j}$ covers all time steps). Thus, letting $\mathbf{R}_i$ denote the coordinates of receiver i, we have the constraints

$$\|\mathbf{X}_{jt} - \mathbf{R}_i\| \begin{cases} \leq r & \text{at } t \in \tau_{N_j} \text{ for some } i \\ > r & \text{at } t \in \tau_{M_j} \text{ for all } i = 1, \cdots, I \end{cases}, \quad \mathbf{X}_{jt} \neq \text{location on land.} \tag{1}$$



(a) 2D plot (b) 3D plot

Figure 2. Detection probability of an acoustically tagged fish within a rated circular receiver range in Chesapeake Bay, shown as a 2D filled-contour map (left) and its matching 3D surface plot (right), illustrating highest probability at the center where the receiver is located, and tapering off toward the "edge" that is three standard deviations from the center.

Thus, our model imposes constraints that restrict imputed fish locations for $t \in \tau_{N_j}$ to be within the detection radius of the activated receiver, and for $t \in \tau_{M_j}$, to be outside of the detection radii of all receivers. (This differs from (*29*) and similar works in which the fish can enter the detection zone of an unactivated receiver with low probability.) Additionally, for all $t \in \tau_{N_j} \cup \tau_{M_j}$, the fish is prevented from being placed over land. The last condition means the land barriers are recognized by the model to prevent unrealistic terrestrial imputations.

Imputing the location of the fish for times $t \in \tau_{N_j}$ is to pick a random point from a circular cloud of points centered at the receiver's coordinates, where the density of points decreases as the fish moves away from the center. This approach reflects the decreasing detection power of receivers discussed in section 2. Hence, we assume the bivariate normal distribution

$$\mathbf{X}_{jt} \sim \mathbf{N}_2(\mathbf{R}_i, \mathbf{\Sigma}_R), \quad \text{for receiver } i \tag{2}$$

in which the covariance matrix is defined as:

$$\mathbf{\Sigma}_R = \begin{bmatrix} \sigma_R^2 & 0 \\ 0 & \sigma_R^2 \end{bmatrix} \tag{3}$$

where 3σ is taken to be the rated detection radius of $r = 500$ meters. This means that the imputed coordinates of the fish within this detection radius are normally distributed with an equal variance of σ_R^2 in any direction around the receiver (see Figure 2).

However, for the time steps of $t \in \tau_{M_j}$, we use a more complex model that incorporates direction and distance. This approach is tailored to the movement characteristics of each fish and ensures that the fish moves toward the final location in a logical time frame, as follows.

Consider fish j during an unobserved period of time between observations at $t = 1$ and

$t = T_j$. That is, receivers detected fish j at times 1 and T_j, but not at $t = 2, 3, \cdots, T_{j-1}$. By letting θ_{jt} be the angle between the fish's imputed location at time $t-1$ and its final location, say $\mathbf{X}_{jT_j} = \mathbf{X}_j^*$ imputed by equations 2 – 3, the location at each step t for $t = 2, 3, \cdots, T_j - 1$ is modeled as

$$\mathbf{X}_{jt} = \mathbf{X}_{j(t-1)} + \begin{bmatrix} |D_{jt}| \cos(\theta_{jt} + \psi_{jt}) \\ |D_{jt}| \sin(\theta_{jt} + \psi_{jt}) \end{bmatrix}, \tag{4}$$

where

$$D_{jt} \sim \mathcal{N}(\frac{d_{jt}}{T_j - (t-1)}, d_{jt}^{\phi}), \quad \psi_{jt} \sim \mathcal{N}(0, \sigma^2_{\psi_{jt}}), \tag{5}$$

$$d_{jt} = \|\mathbf{X}_j^* - \mathbf{X}_{j(t-1)}\|, \tag{6}$$

$$\sigma^2_{\psi_{jt}} = \begin{cases} \gamma. \exp\{n_j - (t-1)\} & n_j \leq \beta \\ \gamma. U_t & n_j > \beta \end{cases}, \tag{7}$$

and the quantities in equations 4 – 7 are described as follows.

We let D_{jt} be a latent stochastic quantity that determines the distance (measured in degrees) between $\mathbf{X}_{j(t-1)}$ and $\mathbf{X}_{jt}$. To accurately account for the distance traveled at each time step, we model D_{jt} as normally distributed, where the mean is defined to ensure the imputation process considers the remaining distance. Here, d_{jt} represents the total distance that remains to be traveled from time t minus the two receivers' radii. Additionally, the variance of D_{jt} is modeled using a power decay function to capture variability in the remaining distance traveled, where ϕ is a parameter that controls the variance's rate with respect to the remaining distance. (This is a decay function because $d_{jt} < 1$ within the spatial domain and it decreases as t increases.) This formulation ensures that both the mean and the variance of D_{jt} dynamically adjust based on the remaining distance, providing a robust framework for imputing missing location in a spatiotemporal context.

For the angle at which the fish travels from $\mathbf{X}_{j(t-1)}$ to $\mathbf{X}_{jt}$, we incorporate a direction of movement and reflect the fish's determination toward $\mathbf{X}_j^*$ (the end point of its trajectory) through γ in equation 7; it represents the initial value of the variance of the noise term ψ_{jt} associated with θ_{jt}. The threshold parameter β determines the point at which the variance model switches from exponential to a uniform distribution, where $U_t \sim \text{Uniform}(0, \beta)$. This assumption implies that for smaller time gaps ($n_j \leq \beta$, where $n_j = T_j - 2$ is the total number of unobserved time steps) the variance decreases exponentially as the fish moves toward $\mathbf{X}_j^*$. However, for larger time gaps ($n_j > \beta$), we give extra freedom to ψ_{jt} by considering a variance value derived from the uniform distribution to attribute the longer time between detections to the fish's reduced directedness. Our model, as derived above, will be used in section 5 to simulate individual fish trajectories.

4. Likelihood

For any given fish from the dataset, the above probability model can be used to simulate its daily locations between two consecutive detections, forming a single imputed segment of its full trajectory. Connecting all imputed segments gives rise to the full imputed trajectory for the fish. The stochastic nature of the imputed trajectory means that one

simulated trajectory may be more probable than another, even if both trajectories are simulated for the same fish from the same model. In this section, we derive the likelihood for any trajectory that is simulated from the model in section 3. The likelihood will be used to display the 90% most likely paths in the form of a heatmap in section 5, thus allowing us to visualize the geographical areas over which the fish's presence is most likely to be concentrated.

To derive the likelihood, we use a product of conditional probabilities. Hence, let us consider fish j and one segment of this fish trajectory which started at $t = 1$ and ended at $t = T_j$. Thus, we drop the subscript j from the following equations.

According to equation 1, $\forall \quad t \in \tau_N = \{1, T\}$,

$$\mathbf{X}_t = \begin{bmatrix} \mathbf{X}_t^{(1)} \\ \mathbf{X}_t^{(2)} \end{bmatrix} \sim \mathcal{N}_2(\mathbf{R}_t, \Sigma_R)$$

where

$$\mathbf{R}_t = \begin{cases} \mathbf{R}_i & \textit{if receiver } i \textit{ detected the fish at time } t \\ \text{NA} & \textit{otherwise} \end{cases}.$$

Therefore, the joint conditional density of $\mathbf{X}_1$ and $\mathbf{X}_T$ is defined as

$$\mathcal{P}_0(\mathbf{X}_1, \mathbf{X}_T | \mathbf{R}_1, \mathbf{R}_T, \sigma_R^2) = \left(\frac{1}{2\pi\sigma_R^2}\right)^2 \exp\left(-\frac{1}{2\sigma_R^2} \sum_{t=1,T} \left[(\mathbf{X}_t^{(1)} - \mathbf{R}_t^{(1)})^2 + (\mathbf{X}_t^{(2)} - \mathbf{R}_t^{(2)})^2\right]\right).$$

Let t belong to the set $\tau_M = \{2, 3, \cdots, T-1\}$, and $\mathbf{Z}_t = \begin{bmatrix} \cos(\theta_t^*) \\ \sin(\theta_t^*) \end{bmatrix}$, for $\theta_t^* = \theta_t + \psi_t$; hence,

$$\theta_t^* | \theta_t, \sigma_{\psi_t}^2 \sim \mathcal{N}(\theta_t, \sigma_{\psi_t}^2).$$

By using a two-term Taylor expansion of $\mathbf{Z}_t = \mathbf{Z}_t(\theta_t^*)$ around θ_t, we have (see Appendix A)

$$\mathbf{Z}_t | \theta_t, \sigma_{\psi_t}^2 \overset{approx}{\sim} \mathbf{N}_2(\boldsymbol{\mu}_z, \Sigma_z)$$

where

$$\boldsymbol{\mu}_z = \begin{bmatrix} \cos(\theta_t) - \frac{1}{2}\cos(\theta_t)\sigma_{\psi_t}^2 \\ \sin(\theta_t) - \frac{1}{2}\sin(\theta_t)\sigma_{\psi_t}^2 \end{bmatrix},$$

and letting $\mathrm{C}_t = \cos(\theta_t)$ and $\mathrm{S}_t = \sin(\theta_t)$, the covariance matrix is

$$\boldsymbol{\Sigma}_z = \sigma_{\psi_t}^2 \begin{bmatrix} \mathrm{S}_t^2 + \frac{1}{2}\mathrm{C}_t^2\sigma_{\psi_t}^2 & -\mathrm{S}_t\mathrm{C}_t + \frac{1}{2}\mathrm{S}_t\mathrm{C}_t\sigma_{\psi_t}^2 \\ -\mathrm{S}_t\mathrm{C}_t + \frac{1}{2}\mathrm{S}_t\mathrm{C}_t\sigma_{\psi_t}^2 & \mathrm{C}_t^2 + \frac{1}{2}\mathrm{S}_t^2\sigma_{\psi_t}^2 \end{bmatrix}.$$

Next, let $\mathbf{W}_t = |D_t|\ \mathbf{Z}_t$, thus

$$\mathbf{W}_t | D_t, \theta_t, \sigma_{\psi_t}^2 \overset{approx}{\sim} \mathbf{N}_2\left(|D_t|\boldsymbol{\mu}_z, D_t^2\Sigma_z\right).$$

Letting $\mathbf{X}_t = \mathbf{X}_{t-1} + W_t$, we have,

$$\mathbf{X}_t|\mathbf{X}_{t-1}, D_t, \theta_t, \sigma^2_{\psi_t} \overset{approx}{\sim} \mathcal{N}_2\left(\mathbf{X}_{t-1} + |D_t|\boldsymbol{\mu}_z, D_t^2\Sigma_z\right)$$

with density

$$\mathcal{P}_1(\mathbf{X}_t|\mathbf{X}_{t-1}, D_t, \theta_t, \sigma^2_{\psi_t}) = \\ \frac{1}{2\pi D_t^2|\Sigma_z|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{X}_t - \mathbf{X}_{t-1} - |D_t|\boldsymbol{\mu}_z)^T(D_t^2\Sigma_z)^{-1}(\mathbf{X}_t - \mathbf{X}_{t-1} - |D_t|\boldsymbol{\mu}_z)\right).$$

Next, as ϕ is non-stochastic, we have $D_t|d_t, \phi \sim \mathcal{N}(\frac{d_t}{T-(t-1)}, d_t^\phi)$ with density

$$\mathcal{P}_2(D_t|d_t, \phi) = \frac{1}{\sqrt{2\pi d_t^\phi}} \exp\left(-\frac{1}{2d_t^\phi}\left(D_t - \frac{d_t}{T-(t-1)}\right)^2\right).$$

Note that conditioned on $\mathbf{X}_{t-1}$ and $\mathbf{X}^*$, both d_t and θ_t are non-stochastic, and given n_j and β, the variance $\sigma^2_{\psi_t}$ is also non-stochastic. Thus, the joint distribution of stochastic terms relevant to time $t \in \tau_M$ is:

$$f(\mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{X}_1, \mathbf{X}^*, D_t, d_t, \theta_t) = \\ \mathcal{P}_1(\mathbf{X}_t|\mathbf{X}_{t-1}, D_t, \sigma^2_{\psi_t}, \theta_t)\ \mathcal{P}_2(D_t|d_t, \phi)\ f(\mathbf{X}_{t-1}, \mathbf{X}^*).$$

We expand it to all $t \in \tau_N \cup \tau_M$ (see Appendix B) to have

$$f(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_{T-1}, \mathbf{X}^*, D_2, \cdots, D_{T-1}, \theta_2, \cdots, \theta_{T-1}, d_2, \cdots, d_{T-1}) = \\ \mathcal{P}_0(\mathbf{X}_1, \mathbf{X}^*|\mathbf{R}_1, \mathbf{R}_T, \sigma_R^2) \prod_{t=2}^{T-1} \mathcal{P}_1(\mathbf{X}_t|\mathbf{X}_{t-1}, D_t, \theta_t, \sigma^2_{\psi_t})\ \mathcal{P}_2(D_t|d_t, \phi). \tag{8}$$

The function in equation 8 above pertains to one single segment of the full trajectory of the fish. Therefore, with all K segments of the trajectory, the likelihood, i.e., the full joint distribution, is

$$\mathcal{L}(\mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \cdots, \mathbf{X}_{k_{T-1}}, \mathbf{X}_k^*, D_{k_2}, \cdots, D_{k_{T-1}}, \theta_{k_2}, \cdots, \theta_{k_{T-1}}, d_{k_2}, \cdots, d_{k_{T-1}}|\boldsymbol{\Omega}) = \tag{9}$$

$$\prod_{k=1}^{K}\left(\mathcal{P}_0(\mathbf{X}_{k_1}, \mathbf{X}_{k_T}|\mathbf{R}_{k_1}, \mathbf{R}_{k_T}, \sigma_R^2) \prod_{t=2}^{T-1} \mathcal{P}_1(\mathbf{X}_{k_t}|\mathbf{X}_{k_{t-1}}, D_{k_t}, \theta_{k_t}, \sigma^2_{\psi_t})\ \mathcal{P}_2(D_{k_t}|d_{k_t}, \phi)\right) \tag{10}$$

where $\boldsymbol{\Omega}$ denotes the set of model parameters $\{\beta, \gamma, \phi, \sigma_R^2\}$, k_t denotes the t-th time step of the k-th segment of the fish trajectory, and $\mathbf{X}_{k_1} = \mathbf{X}^*_{k-1}$ for all $k > 1$, i.e. the fish's new starting location is its previous end location.

5. Simulation Study

We conduct a simulation study to facilitate the validation of the proposed imputation model by Cobia fisheries scientists. To this end, we apply the method with daily time intervals to two selected fish with acoustic tag numbers 18453 and 18434. One fish

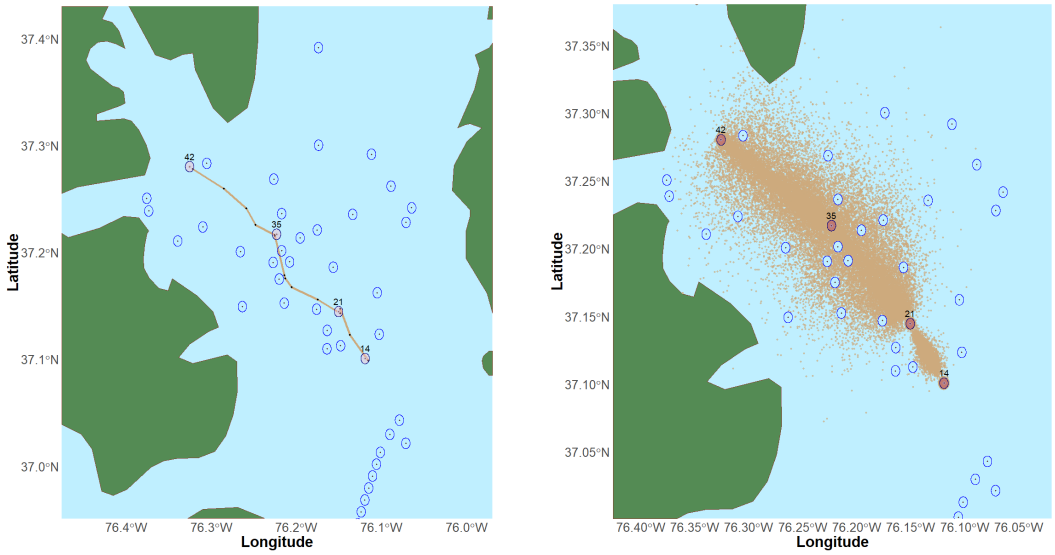
exhibited migratory behavior, while the other fish's recorded movement was almost stationary. The timeline for each fish varies. The first fish's time period is from September 1, 2018, to September 12, 2018, while the second fish's period is from September 23, 2017, to October 10, 2017. Since no environmental covariates or population interactions are considered in this current model, and these two fish trajectories are imputed independently, the difference in timelines has no effect on our imputations. The first fish was detected at time steps $t = 1, 5, 10, 12$, and by receiver numbers $42, 35, 21, 14$. This indicates that the fish was detected at these four receivers over four different time steps, within a span of twelve consecutive time steps, and thus, $K = 3$, $T_{k=1} = 5$, $T_{k=2} = 6$, and $T_{k=3} = 3$. Similarly, the second fish was detected at $t = 1, 7, 18$, and by receivers $36, 39, 22$, and thus, $K = 2$, $T_{k=1} = 7$, and $T_{k=2} = 12$.

We demonstrate the effectiveness of the model for these different fish behaviors using prior predictive sampling along with likelihood filtering, i.e., comparing likelihood values of all generated trajectories to obtain the most credible estimated paths. Prior predictive sampling is employed here to enhance realism, whereby we simulate all non-stochastic quantities in $\mathbf{\Omega}$ from relatively uninformative prior distributions instead of fixing each at a predetermined value. While we do not restrict the prior distributions to take on any specific parameter value, we carefully select hyperparameters that meet our desired criteria. Thus, we select prior distributions for the parameters that reflect the prior knowledge and theoretical considerations regarding the described method. For ϕ, we take $\log \phi \sim \mathcal{N}(0.5, 0.1^2)$ to achieve a positive value with a reasonable range of variability for D_{jt} — for a typical imputed trajectory for either fish ID 18453 or 18434, this prior results in D_{jt} as well as its standard deviation having a range on the order of 10^{-3} to 10^{-1} degrees (geographic coordinate system), which is rather diffuse for Chesapeake Bay. For the parameter γ, we take $\log \gamma \sim \mathcal{N}(-3, 0.1^2)$ distribution, which means $E(\gamma) = 0.05$ and $\text{Var}(\gamma) = 0.005$, with 99% prior probability for $\gamma \in [0.038, 0.064]$. This leads to a variance for the angle, $\sigma^2_{\psi_{jt}}$, that varies depending on the value of β and n_j for each trajectory. For instance, if $\beta = 3$ (its prior mean), then $E(\sigma^2_{\psi_{jt}}|\beta = 3)$ could approximately vary from 0.1 to 1. And more specifically, for $n_j = 3$, and $t = 2$, then $E(\sigma^2_{\psi_{jt}}|\beta = n_j = 3, t = 2) = 0.37$, which gives ψ_{jt} a 99% prior conditional probability of ranging over $[-\pi/2, +\pi/2]$. The prior here for γ ensures stability in $\sigma^2_{\psi_{jt}}$ and minimizes large fluctuations in movement directions while allowing for variability. Next, σ^2_R is modeled as inverse-Gamma$(3, 0.00298)$ to achieve a prior variance value that ensures $3\mathbb{E}(\sigma_R) \approx r$. In other words, σ^2_R accounts for the spatial sparsity of the receiver network and weakening detection power within the rated radius r of any receiver. Lastly, the threshold parameter, β, follows a Poisson distribution, $P(3)$.

We initially focus on the first fish (ID 18453) with relatively short time gaps. The fish's trajectory starts at $t = 1$ near receiver 42, then 3 missing time steps later, near receiver 35 at $t = 5$, then 4 missing time steps later, near receiver 21 at $t = 10$, and then the trajectory continues to the last location at $t = 12$ near receiver 14.

To illustrate the model-based algorithm, we consider the first fish for one iteration (Figure 3a). Black points in this figure show the fish's simulated locations at the missing time points. The khaki lines connecting these points illustrate the trajectory progression from receiver 42 to receiver 14. Thus, it is important to note that this line does not imply a linear path between the simulated points, but is intended to show a visual guide to understand the overall movement. As such, any line segment crossing the detection zone of a receiver that was not activated by the fish (an uncolored bubble in Figure 3a) should not be interpreted as the fish having ever entered the receiver's detection zone.

We repeat the simulation algorithm for 9 additional iterations, each time step within



(a) One simulated trajectory.

(b) 4 500 simulated trajectories.

Figure 3. Comparison of one simulated trajectory (panel (a)) and 4 500 simulated trajectories (panel (b)) for fish ID 18453 over 12-time steps using the proposed algorithm. Receivers 42 ($t = 1$), 35 ($t = 5$), 21 ($t = 10$), and 14 ($t = 12$) detected the fish, and are shown in pink.

each iteration being added to the map, thus forming a movie that shows the simulated progression of the fish's movement from start to finish over ten different simulations. This movie of "spaghetti growth" is available at this link. Lastly, we simulate 5 000 iterations and keep 4 500 of the most probable ones (i.e., with 90% highest likelihood values). Hence, Figure 3b shows 4 500 iterations for fish ID 18453. Additionally, in Figure 4, a heatmap of the same iterations is shown. This heatmap is available in Appendix C with a higher resolution.

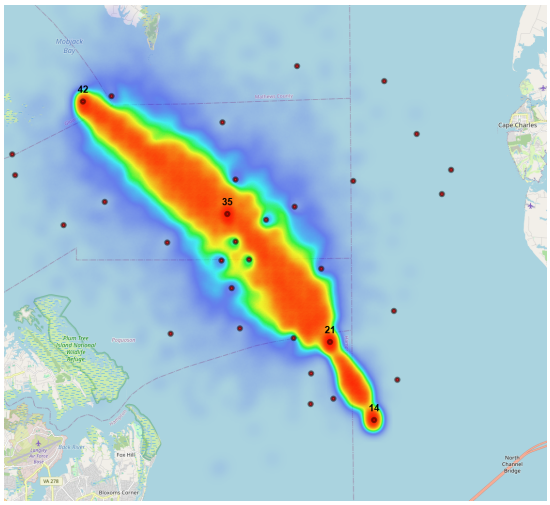
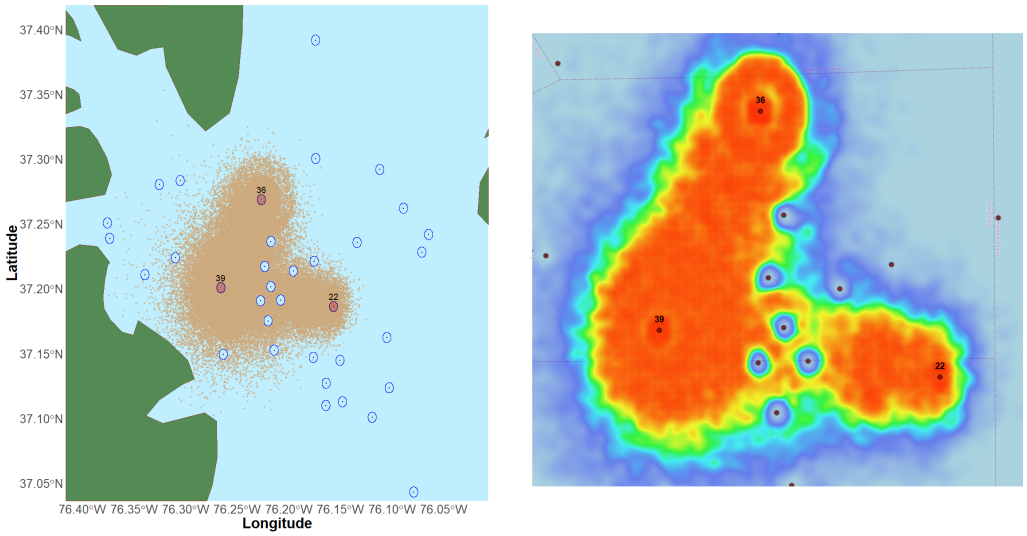


Figure 4. Heatmap corresponding to Figure 3b.

The imputation for the second fish (ID 18434) is shown in Figures 5a and 5b. Figure 5a shows the 90% most probable simulated trajectories out of 5 000, while Figure 5b shows the heatmap of the same trajectories. These two figures demonstrate the potential

realism of the proposed method even when larger time gaps exist, such as the 10-step gap between receivers 39 and 22, for which our model captures what may be considered the movement patterns of a fish that knew where it was heading next.



(a) 4 500 simulated trajectories.

(b) Heatmap corresponding to panel (a).

Figure 5. Simulated trajectories and corresponding heatmap for fish ID 18434 over 18 time steps using the proposed algorithm. Detections were observed by receivers 36 ($t = 1$), 39 ($t = 7$), and 22 ($t = 18$).

6. Comparison with Brownian Bridge Movement Model

To compare our model to one of the most popular fish movement models, we simulate trajectories for the same two fish using the Brownian bridge movement model (BBMM) as a reference. BBMM is one of the most commonly used approaches for data from tracking technologies with continuous spatial coverage, such as GPS (*12*). Note that we use the standard BBMM (*12*) rather than more recent extensions such as the dynamic BBMM (*23*) as a baseline for demonstrating generic properties of existing approaches when applied to the very specific structure of acoustic telemetry data. The comparison of our method against this reference highlights the striking difference between the assumptions of BBMM and similar readily available methods which, unlike ours, do not consider the angle of travel at each time step. Specifically, BBMM and similar methods rely on the premise that detections happen regularly over time, so that each time step with missing detections may be reasonably filled in with a form of "unguided random walk." In contrast, the angle of travel at each time step in our model can dynamically reflect the fish's "empirically determined movement path" (*30*) (as opposed to wandering at random between detections throughout its journey).

As introduced in Section 1, BBMM is a stochastic model that considers the probability of an animal's path between known locations, conditioned on both endpoints. It interpolates plausible trajectories by modeling movement as a Brownian bridge process, with variance increasing until midway between observations, then decreases afterwards, to reflect positional uncertainty as a function of time. A basic BBMM can be expressed as follows.

Let the fish's movement trajectory start at known spatial location $\mathbf{x}_1$ at time $t = 1$ and end at known location $\mathbf{x}_T$ at time $t = T$. Then, the animal's position $\mathbf{X}_t$ at time

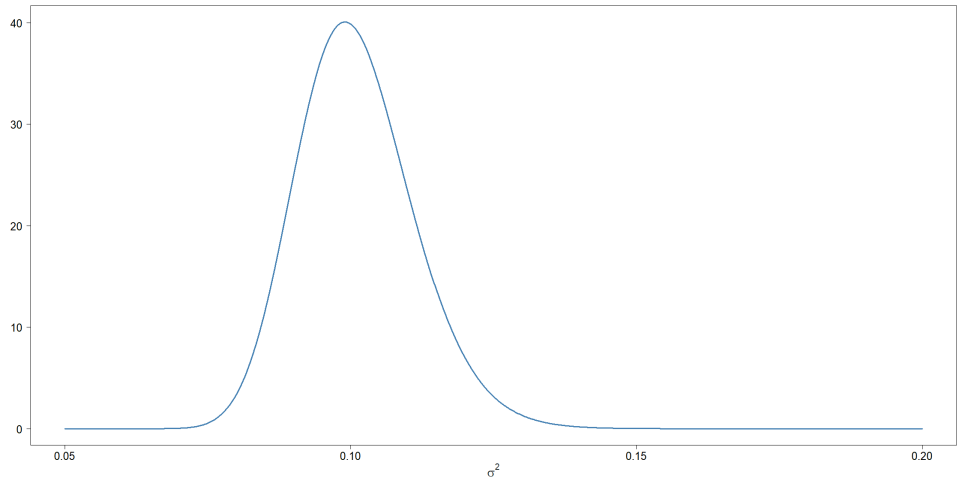


Figure 6. Prior for σ^2 in the Brownian bridge movement model

$t \in (1, T)$ follows a Brownian bridge if

$$\mathbf{X}_t \sim \mathcal{N}\left(\mu_t, \sigma_t^2 \cdot \mathbf{I}\right)$$

where

$$\mu_t = \mathbf{x}_1 + \frac{t-1}{T-1}(\mathbf{x}_T - \mathbf{x}_1), \quad \sigma_t^2 = \frac{(t-1)(T-t)}{T-1}\sigma^2.$$

To apply this BBMM to our highly irregular, sparse detections in both space and time, we first ignore the spatial constraints in equation 1 from our simulation model to remove this upfront advantage of our method over BBMM. Next, for the BBMM only, we take the coordinates of activated acoustic receivers to represent the fish's known locations—despite this being unrealistic for acoustic telemetry (as discussed in Section 2), the BBMM requires the animal's known locations as model inputs. Under this assumption, we impute missing locations for the same two fish (acoustic tag IDs 18453 and 18434) via prior predictive simulations from the BBMM, with the prior distribution $\sigma^2 \sim$ inverse-Gamma$(100, 10)$ for the spread of probable locations. This is a highly informative prior (see Figure 6) that ensures the BBMM is presented in its best-case scenario, avoiding overly broad predictions that would unfairly disadvantage BBMM in this comparison.

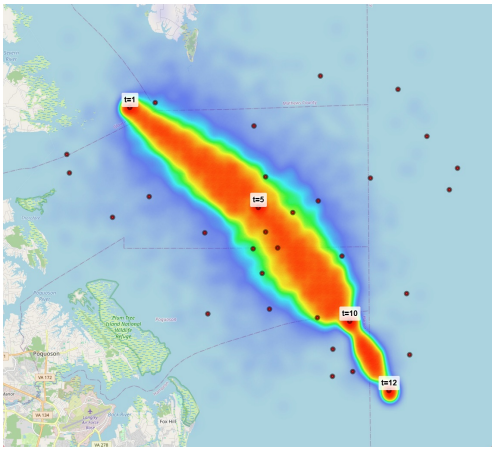
The results are visualized in Figure 7, which includes four sub-graphs: (a) and (b) correspond to 5 000 imputations using our proposed method, and (c) and (d) present the same number of imputations using BBMM. Note that these heatmaps have a consistent zoom scale for both models, with a color gradient in which reds and oranges indicate high-probability zones of fish presence. It is clear that BBMM's predictions cover wider areas with mostly low-probability zones (yellows and greens). This reflects a methodology that is less informed by the relative locations of observed detections. In particular, both BBMM heatmaps (panels (c) and (d)) exhibit diffuse probability density and suggest a lack of direction for the traveling fish toward detected locations, as opposed to the concentrated probability density in our heatmaps (panels (a) and (b)) that quantitatively reveals the fish's directional paths between detections. Our highly probabilistic paths

align with the notion of a fish's "empirically determined movement path" discussed in (*30*). The diffuseness of BBMM's imputations is especially striking when we consider Fish 18434's journey between the three receivers that detected it: as opposed to the wide, rounded density region generated by BBMM (panel (d)) that suggests the fish perhaps stumbled upon the second receiver through its non-directed movement between the first and third (last) receivers, our method generates an L-shaped path (panel (b)) that reflects the fish's empirically determined movement between each pair of receivers it triggered. Furthermore, the regions of higher probability (yellow to red) for BBMM (panel (d)) include more unactivated receivers than for our model (panel (b)).

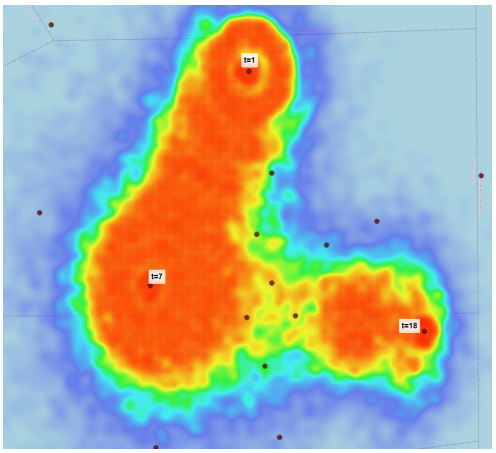
This comparison highlights the potential utility of our model in some crucial aspects, even if one were to remove the necessary spatial constraints imposed by aquatic studies with acoustic telemetry data. First, our proposed model recognizes the fact that spatial coordinates recorded in acoustic telemetry studies represent the receiver's position, not the actual location of the fish. Our imputations account for this feature, and is unlike how BBMM is typically used, which is based on the incorrect premise that observed locations are of the tracked individual. Second, our model accounts for the angle of travel at each time step, thus generating "guided random walk" movements from one detection to the next; this is another critical feature that differs from the BBMM.

7. Some Limitations

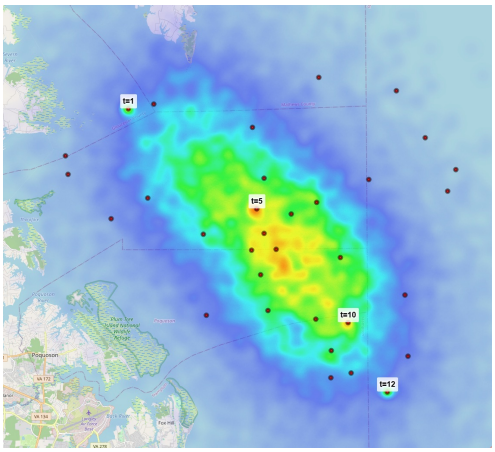
In this paper, we have presented what may be considered a first-order approximation for imputing fish trajectories based on acoustic telemetry data which pose multiple challenges that are not readily addressed by popular animal movement models. While our method generates plausible trajectories under the constraints of the irregular data, it relies on rather simplistic assumptions as a first attempt to mitigate the challenges. One obvious limitation is our use of latitude and longitude as coordinates while transforming meters to degrees for software compatibility and visualization. However, our current focus is on interpolating fish locations within the area of Chesapeake Bay, where the spatial scale is relatively small. Thus, we expect any bias due to treating latitude and longitude as Euclidean coordinates to be minimal, and therefore should not affect the imputation process for our purposes. Moreover, we assume a diagonal covariance matrix for the fish's location vector when it is within the detection radius of any receiver. Although likely unrealistic, this independence assumption simplifies the computations while still accounting for the receiver's decreasing detection ability away from its location. Another obvious limitation is that our model does not account for shoreline variations due to environmental factors (e.g., tides, sea level rise, erosion, coastal development), although we do constrain the imputed fish locations to be within the aquatic boundaries. Lastly, accounting for land barriers in our model by utilizing a shoreline shapefile is an important methodological development. However, the accuracy of the imputations heavily relies on the quality of the coastlines defined within the shapefile. Therefore, greater shapefile precision would result in improved imputations. Lastly, our model regards a non-detection as absence within the receiver's detection range, as opposed to the receiver's potential failure to detect presence (e.g., *29*).



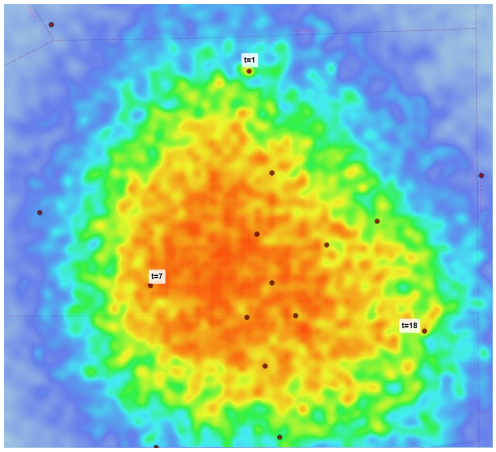
(a) Fish ID 18453: Imputations from our model without spatial restrictions.



(b) Fish ID 18434: Imputations from our model without spatial restrictions.



(c) Fish ID 18453: Imputations from BBMM.



(d) Fish ID 18434: Imputations from BBMM.

Figure 7. Heatmaps comparing 5 000 imputed fish trajectories from our proposed model but without land or receiver restrictions (top panels), and from the Brownian bridge movement model (BBMM, bottom panels). Note the consistency in zoom scale between methods.

8. Discussion

In summary, this paper focuses on the imputation of cobia movement trajectories while accounting for physical limitations due to the tracking technology (acoustic tagging). Our imputation approach relies on simulating from bespoke probability distributions tailored to the data collection technology and animal species of interest. The challenges addressed in this paper are the structure of the data and the methods used for data collection. The data are very sparse in both time and space, the actual fish location is unknown even when presence is detected, and land and detection zones of unactivated receivers are avoided at all times. In fact, these are the complexities of tracking the fish acoustically, which lead to a lack of applicability of many common models for animal movement analysis. In our paper, we developed a new probability modeling framework to handle these challenges for a species that exhibits no homing tendencies but may know its destination as it travels (7), and we demonstrated the potential utility of our

approach by comparing it to Brownian bridge movement modeling (BBMM) that is commonly used to model animal tracking data.

Given that BBMM and our approach generate noticeably different fish movement patterns (quantitatively different according to the respective probability density surfaces), it would be important to further develop a statistical inference framework for our model, so that its inference performance can be rigorously assessed. We intend to apply our method to impute all fish trajectories from the full cobia dataset ($N = 67$ fish) by (*5*). The goal is to develop an inference model that also integrates relevant environmental characteristics along with the imputed data. This will improve understanding and prediction of the behavior of this fish species as well as those animals with similar movement patterns.

The hierarchically structured model introduced in this paper offers a framework for imputing locations during temporal gaps in the trajectory of each fish individually. However, the individuality of fish-specific imputations and the unavailability of the fish's actual locations currently do not provide the context for performing a cross-validation study. Instead, prior predictive simulations allow a visual assessment of the potential utility of our model.

As future work, we will extend this model by incorporating a transition probability matrix for each time step, giving us the probabilities of fish transitioning between different receivers, and therefore, the fish location predictions will allow population-level inference on which k-fold cross-validation may be employed. Additionally, we will improve the fish-specific model's applicability to more extensive time gaps by incorporating multiple layers of mixture modeling to distinguish between "lazy" and "nosy" behavior, and between resident and migratory populations.

Despite the progress in modeling techniques, several challenges remain. Data quality and quantity can be compromised by tag malfunctions, environmental interference, or deep waters, leading to incomplete or distorted datasets (*4*, *10*, *16*, *36*, *39*); these challenges are beyond the scope of our current work. Environmental factors such as water temperature and salinity influence fish movement and add complexity to model predictions. Furthermore, ensuring model accuracy and validation is essential to represent actual movement patterns accurately. Our inference framework (in progress) aims to incorporate these elements into the imputation method presented in this paper. We anticipate our current paper and inference framework (upon completion) to add substantial rigor to this field of research.

Data Availability Statement

All code and data used in this paper are available to the public at this GitHub repository.

Acknowledgement

We gratefully acknowledge the financial support of the Virginia Sea Grant through a graduate fellowship awarded to M.A., which made this research possible. We thank the Editor and reviewers for their helpful comments and input. We also extend our appreciation to Dr. Daniel P. Crear (Inter-American Tropical Tuna Commission) for providing the necessary resources and his guidance during this work. K.C.W. thanks all investigators involved in the FACT Network for acoustic telemetry.

References

(1) J. Alós, M. Palmer, S. Balle, and R. Arlinghaus. Bayesian state-space modelling of conventional acoustic tracking provides accurate descriptors of home range behavior in a small-bodied coastal fish species. *PLoS ONE*, 11(4):e0154089, 2016. DOI: 10.1371/journal.pone.0154089.

(2) J.W. Brownscombe, L.P. Griffin, J.M. Chapman, D. Morley, and A. Acosta. A practical method to account for variation in detection range in acoustic telemetry arrays to accurately quantify the spatial ecology of aquatic animals. *Methods in Ecology and Evolution*, 11(1):82–94, 2020. DOI: 10.1111/2041-210X.13322.

(3) et al. Charles C. Krueger, Christopher M. Holbrook. Acoustic telemetry observation systems: challenges encountered and overcome in the laurentian great lakes. *Journal of Fisheries and Aquatic Sciences*, 75(10):1755–1763, 2018.

(4) S.J. Cooke, S.G. Hinch, M. Wikelski, R.D. Andrews, L.J. Kuchel, T.G. Wolcott, and P.J. Butler. Biotelemetry: a mechanistic approach to ecology. *Trends in Ecology & Evolution*, 19(6):334–343, 2004.

(5) D.P. Crear, B.E. Watkins, V.S. Saba, J.E. Graves, D.R. Jensen, A.J. Hobday, and K.C. Weng. Contemporary and future distributions of cobia, *Rachycentron canadum*. *Diversity and Distributions*, 26(8):1002–1015, June 2020.

(6) M. Dawkins. *Melded Bayesian Inference for Stochastic Theoretical Models with Applications in Agent-Based Modelling*. Honours Thesis, Australian National University, 2017.

(7) Hugh Dingle and V. Alistair Drake. What is migration? *BioScience*, 57:113–121, 2 2007.

(8) Robert M Dorazio and Melissa Price. State-space models to infer movements and behavior of fish detected in a spatial array of acoustic receivers. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(4):543–550, 2019.

(9) Eric Fu and Nancy Heckman. Model-based curve registration via stochastic approximation EM algorithm. *Computational Statistics & Data Analysis*, 131:159–175, 2019.

(10) M.R. Heupel, J.M. Semmens, and A.J. Hobday. Automated acoustic tracking of aquatic animals: scales, design and deployment of listening station arrays. *Marine and Freshwater Research*, 57(1):1–13, 2006.

(11) Mevin B. Hooten, Devin S. Johnson, Brett T. McClintock, and Juan M. Morales. *Animal Movement: Statistical Models for Telemetry Data*. CRC Press, Boca Raton, FL, 2017.

(12) J.S. Horne, E.O. Garton, S.M. Krone, and S.J. Lewis. Analyzing animal movements using Brownian bridges. *Ecology*, 88:2354–2363, 2007.

(13) Nathan J. Hostetter and J. Andrew Royle. Movement-assisted localization from acoustic telemetry data. *Methods in Ecology and Evolution*, 8(15):1–10, 2020.

(14) N.E. Humphries, N. Queiroz, J.R.M. Dyer, N.G. Pade, M.K. Musyl, and K.M. Schaefer et al. Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature*, 465:1066–1069, 2010.

(15) G. Huse, J. Giske, and A.G.V. Salvanes. Modeling individual-based populations: A review and introduction. In *Handbook of Fish and Fisheries*, pages 228–248. Blackwell, Oxford, 2002.

(16) N.E. Hussey, S.T. Kessel, K. Aarestrup, S.J. Cooke, P.D. Cowley, A.T. Fisk, R.G. Harcourt, K.N. Holland, S.J. Iverson, J.F. Kocik, et al. Aquatic animal telemetry: a panoramic window into the underwater world. *Science*, 348(6240):1255642, 2015.

(17) Charlie Huveneers, Colin A. Simpfendorfer, Susan W. Kim, Jayson M. Semmens, Alistair J. Hobday, Hugh Pederson, Thomas C. Stieglitz, Richard Vallee, Dale Mitchell Webber, Michelle R. Heupel, Victor M. Peddemors, and Robert G. Harcourt. The influence of environmental parameters on the performance and detection range of acoustic receivers. *Methods in Ecology and Evolution*, 7, 2016.

(18) D.R. Jensen and J.E. Graves. Movements, habitat utilization, and post-release survival of cobia (rachycentron canadum) that summer in virginia waters assessed using pop-up satellite archival tags. *Animal Biotelemetry*, 8(24), 2020.

(19) D.I. Jonsen, R.A. Myers, and J.M. Flemming. Meta-analysis of animal movement using

state-space models. *Ecology*, 84(11):3055–3063, 2003.
(20) I.D. Jonsen, J.M. Flemming, and R.A. Myers. Robust state-space modeling of animal movement data. *Ecology*, 86(11):2874–2880, 2005.
(21) R. Joo, S. Bertrand, J. Tam, and R. Fablet. Hidden Markov models: the best models for forager movements. *PLoS One*, 8(8):e71246, Aug 23 2013.
(22) S. T. Kessel, S. J. Cooke, M. R. Heupel, N. E. Hussey, C. A. Simpfendorfer, S. Vagle, and A. T. Fisk. A review of detection range testing in aquatic passive acoustic telemetry studies. *Reviews in Fish Biology and Fisheries*, 24:199–218, 3 2014.
(23) R. et al. Kranstauber, B. Kays. A dynamic brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of Animal Ecology*, 81(4):738–746, 2012.
(24) N.H. Mathies, M.B. Ogburn, G. McFall, and S. Fangman. Environmental interference factors affecting detection range in acoustic telemetry studies using fixed receiver arrays. *Marine Ecology Progress Series*, 495:27–38, 2014. DOI: 10.3354/meps10582.
(25) B.T. McClintock. Incorporating telemetry error into hidden markov models of animal movement using multiple imputation. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3):249–269, 2017. DOI: 10.1007/s13253-017-0285-6.
(26) J.R. Norris. *Markov Chains*. Cambridge University Press, 1998.
(27) T.A. Patterson, M. Basson, M.V. Bravington, and J.S. Gunn. Classifying movement behavior in relation to environmental conditions using hidden Markov models. *Journal of Animal Ecology*, 78(6):1113–1123, 2009.
(28) T.A. Patterson, L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. State-space models of individual animal movement. *Trends in Ecology & Evolution*, 23(2):87–94, 2008.
(29) Martin W. Pedersen and Kevin C. Weng. Estimating individual animal movement from observation networks. *Methods in Ecology and Evolution*, 4:920–929, 10 2013.
(30) Thomas W. Pike and Oliver H. P. Burman. Simulating individual movement in fish. *Scientific Reports*, 13(1):14581, September 2023.
(31) V. Pozdnyakov, T. Meyer, Y. B. Wang, and J. Yan. On modeling animal movements using Brownian motion with measurement error. *Ecology*, 95(2):247–253, 2014.
(32) S.F. Railsback. Concepts from complex adaptive systems as a framework for individual-based modeling. *Ecological Modeling*, 139:47–62, 2001.
(33) J. Reubens, P. Verhelst, I. van der Knaap, et al. Environmental factors influence the detection probability in acoustic telemetry in a marine environment: results from a new setup. *Hydrobiologia*, 845:81–94, 2019.
(34) S.M. Ross. *Introduction to Probability Models*. Academic Press, 11th edition, 2014.
(35) H. Scharf, M.B. Hooten, and D.S. Johnson. Imputation approaches for animal movement modeling. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3):270–286, 2017. DOI: 10.1007/s13253-017-0294-5.
(36) Stephen R. Scherrer, Brendan P. Rideout, Giacomo Giorli, Eva-Marie Nosal, and Kevin C. Weng. Depth- and range-dependent variation in the performance of aquatic telemetry systems: understanding and predicting the susceptibility of acoustic tag–receiver pairs to close proximity detection interference. *PeerJ*, 6:e4249, 1 2018.
(37) Thomas H Selby, Kristen M Hart, Ikuko Fujisaki, Brian J Smith, Clayton J Pollock, Zandy Hillis-Starr, Ian Lundgren, and Madan K Oli. Can you hear me now? range-testing a submerged passive acoustic receiver array in a caribbean coral reef habitat. *Ecology and evolution*, 6(14):4823–4835, 2016.
(38) D.W. Sims, E.J. Southall, N.E. Humphries, G.C. Hays, C.J.A. Bradshaw, and J.W. Pitchford et al. Scaling laws of marine predator search behavior. *Nature*, 451:1098–1102, 2008.
(39) E.B. Thorstad, F. Whoriskey, A.H. Rikardsen, and K. Aarestrup. Aquatic animal telemetry. *Marine Ecology Progress Series*, 482:1–3, 2013.
(40) W. Van Winkle, K.A. Rose, and R.C. Chambers. Individual-based approach to fish population dynamics: an overview. *Transactions of the American Fisheries Society*, 122:397–403, 1993.

(41) M.V. Winton, J. Kneebone, D.R. Zemeckis, and G. Fay. A spatial point process model to estimate individual centres of activity from passive acoustic telemetry data. *Methods in Ecology and Evolution*, 9(11):2262–2272, 2018. DOI: 10.1111/2041-210X.13080.

(42) Joy M. Young, Mary E. Bowers, Eric A. Reyier, Danielle Morley, Erick R. Ault, Jonathan D. Pye, Riley M. Gallagher, and Robert D. Ellis. The fact network: Philosophy, evolution, and management of a collaborative coastal tracking network. *Marine and Coastal Fisheries*, 12:258–271, 10 2020.

Appendix A. Second-order Taylor Expansion for Z

For $\mathbf{Z} = \begin{bmatrix} Z^{(1)} = \cos(\theta^*) \\ Z^{(2)} = \sin(\theta^*) \end{bmatrix}$,

$$\begin{aligned} Z^{(1)} &= \cos\theta^* \\ &\approx \cos\theta - (\sin\theta)(\theta^* - \theta) - \frac{1}{2}(\cos\theta)(\theta^* - \theta)^2 \\ &\approx \cos\theta\left(1 - \frac{(\theta^* - \theta)^2}{2}\right) - (\sin\theta)(\theta^* - \theta), \\ \mathbb{E}(Z^{(1)}) &\approx \cos\theta\left(1 - \frac{\sigma_\psi^2}{2}\right), \\ \mathrm{Var}(Z^{(1)}) &\approx (\sin\theta)^2\sigma_\psi^2 + \frac{1}{2}(\cos\theta)^2\sigma_\psi^4, \end{aligned}$$

and

$$\begin{aligned} Z^{(2)} &= \sin\theta^* \\ &\approx \sin\theta + (\cos\theta)(\theta^* - \theta) - \frac{1}{2}(\sin\theta)(\theta^* - \theta)^2 \\ &\approx \cos\theta(\theta^* - \theta) + (\sin\theta)\left(1 - \frac{(\theta^* - \theta)^2}{2}\right), \\ \mathbb{E}(Z^{(2)}) &\approx (\sin\theta)\left(1 - \frac{\sigma_\psi^2}{2}\right), \\ \mathrm{Var}(Z^{(2)}) &\approx (\cos\theta)^2\sigma_\psi^2 + \frac{1}{2}(\sin\theta)^2\sigma_\psi^4, \end{aligned}$$

and by letting $\mathrm{C} = \cos(\theta)$ and $\mathrm{S} = \sin(\theta)$, we have

$$\mathrm{Cov}(Z^{(1)}, Z^{(2)}) \approx -\mathrm{SC}\sigma_\psi^2 + \frac{1}{2}\mathrm{SC}\sigma_\psi^4,$$

$$\begin{bmatrix} Z^{(1)} \\ Z^{(2)} \end{bmatrix} \overset{approx}{\sim} \mathbf{N}_2\left(\begin{bmatrix} \cos(\theta) - \frac{1}{2}\cos(\theta)\sigma_\psi^2 \\ \sin(\theta) - \frac{1}{2}\sin(\theta)\sigma_\psi^2 \end{bmatrix}, \sigma_\psi^2 \begin{bmatrix} \mathrm{S}^2 + \frac{1}{2}\mathrm{C}_t^2\sigma_\psi^2 & -\mathrm{SC} + \frac{1}{2}\mathrm{SC}\sigma_\psi^2 \\ -\mathrm{SC} + \frac{1}{2}\mathrm{SC}\sigma_\psi^2 & \mathrm{C}^2 + \frac{1}{2}\mathrm{S}^2\sigma_\psi^2 \end{bmatrix}\right).$$

Appendix B. Joint Distribution of Stochastic Terms for $T = 4$

Assume the trajectory lacks observations at time steps $t = 2, 3$. Thus, $X_4 = X^*$, and

$$
\begin{aligned}
&\mathcal{P}\left[X_1,\ \ X_2,\ \ X_3,\ \ X_4,\ \ \theta_2,\ \ \theta_3,\ \ D_2,\ \ D_3,\ \ d_2,\ \ d_3\right] \\
=&\mathcal{P}\left[X_3 \,|\ \ X_1,\ \ X_2,\ \ X_4,\ \ \theta_2,\ \ \theta_3,\ \ D_2,\ \ D_3,\ \ d_2,\ \ d_3\right] \times \\
&\mathcal{P}\left[D_3 \,|\ \ X_1,\ \ X_2,\ \ X_4,\ \ \theta_2,\ \ \theta_3,\ \ D_2,\ \ d_2,\ \ d_3\right] \times \\
&\mathcal{P}\left[d_3\,,\ \ \theta_3 \,|\ \ X_1,\ \ X_2,\ \ X_4,\ \ \theta_2,\ \ D_2,\ \ d_2\right] \times \\
&\mathcal{P}\left[X_2 \,|\ \ X_1,\ \ X_4,\ \ \theta_2,\ \ D_2,\ \ d_2\right] \times \\
&\mathcal{P}\left[D_2 \,|\ \ X_1,\ \ X_4,\ \ \theta_2,\ \ d_2\right] \times \\
&\mathcal{P}\left[d_2\,,\ \ \theta_2 \,|\ \ X_1,\ \ X_4\right] \times \\
&\mathcal{P}\left[X_1\right] \mathcal{P}\left[X_4\right] \\
=&\mathcal{P}\left[X_3 \,|\ \ X_2,\ \ D_3,\ \ \theta_3\right] \mathcal{P}\left[D_3 \,|\ \ d_3\right] \times \\
&\mathcal{P}\left[X_2 \,|\ \ X_1,\ \ D_2,\ \ \theta_2\right] \mathcal{P}\left[D_2 \,|\ \ d_2\right] \times \\
&\mathcal{P}\left[X_1\right] \mathcal{P}\left[X_4\right].
\end{aligned}
$$

Note that $\mathcal{P}\left[d_3, \theta_3 \,|X_1, X_2, X_4, \theta_2, D_2, d_2\right] = \mathcal{P}\left[d_3, \theta_3 \,|X_2, X_4\right]$ drops out of the likelihood due to non-stochasticity. The same applies to $\mathcal{P}\left[d_2, \theta_2 \,|X_1, X_4\right]$.

Appendix C. High resolution heat map

Download Figure 4 in high resolution HTML format at this link