

# Estimating fine-scale movement rates and habitat preferences using multiple data sources

James T. Thorson<sup>1</sup>  | Steven J. Barbeaux<sup>2</sup>  | Daniel R. Goethel<sup>3</sup>  |  
 Kelly A. Kearney<sup>4,5</sup>  | Edward A. Laman<sup>6</sup>  | Julie K. Nielsen<sup>7</sup>  |  
 Matthew R. Siskey<sup>8</sup>  | Kevin Siwicke<sup>3</sup>  | Grant G. Thompson<sup>2</sup> 

<sup>1</sup>Habitat and Ecological Processes Research Program, Alaska Fisheries Science Center, NOAA, Seattle, WA, USA

<sup>2</sup>Status of Stocks and Multispecies Assessments, Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, NOAA, Seattle, WA, USA

<sup>3</sup>Marine Ecology and Stock Assessment, Auke Bay Laboratories, Alaska Fisheries Science Center, NOAA, Seattle, WA, USA

<sup>4</sup>Cooperative Institute for Climate, Ocean and Ecosystem Studies, University of Washington, Seattle, WA, USA

<sup>5</sup>Resource Ecology and Ecosystem Modeling Program, Resource Ecology and Fisheries Management Division, Alaska Fisheries Science Center, NOAA, Seattle, WA, USA

<sup>6</sup>Groundfish Assessment Program, Resource Assessment and Conservation Division, Alaska Fisheries Science Center, NOAA, Seattle, WA, USA

<sup>7</sup>Kingfisher Marine Research, LLC, Juneau, AK, USA

<sup>8</sup>School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA

## Correspondence

James T. Thorson, Habitat and Ecological Processes Research Program, Alaska Fisheries Science Center, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115, USA.  
 Email: James.Thorson@noaa.gov

## Funding information

Joint Institute for the Study of the Atmosphere and Ocean, Grant/Award Number: NA15OAR4320063

## Abstract

Fisheries scientists and managers must track rapid shifts in fish spatial distribution to mitigate stakeholder conflict and optimize survey designs, and these spatial shifts result in part from animal movement. Information regarding animal movement can be obtained from selection experiments, tagging studies, flux through movement gates (e.g. acoustic arrays), fishery catch-per-unit effort (CPUE), resource surveys and genetic/chemical markers. However, there are few accessible approaches to combine these data types while accounting for spatially correlated residual patterns. We therefore discuss a movement model involving diffusion (random movement), taxis (movement towards preferred habitat) and advection (passive drift following ocean currents). We specifically outline how these movement processes can be fitted to data while discretizing space and time and estimating non-linear habitat preferences using environmental layers as well as spatial process errors. Finally, we introduce an R package, ATM, by fitting the model to bottom trawl survey, longline fishery and tagging data for Pacific cod (*Gadus macrocephalus*, Gadidae) in the Bering Sea during winter/summer seasons from 1982 to 2019. Combining data types predicts an increasing proportion of cod residing in the northern Bering Sea from 2013 to 2019, and estimates are informative in a recent stock assessment model. We fit sensitivity analyses by dropping tag, survey or fishery data, and this analysis shows that tagging data are necessary to identify diffusion rates, while survey data are informative about movement among biogeographic strata. This “hybrid” species distribution model can help explain poleward movement, project distributions under future climate conditions and evaluate alternative tag-deployment scenarios to optimize tagging designs.

## KEY WORDS

advection, diffusion, habitat preference, instantaneous movement, Pacific cod, taxis

## 1 | INTRODUCTION

Understanding animal migrations and distributional shifts is central to science supporting ocean management, including stock, ecosystem, climate, habitat and protected species assessments routinely conducted in the United States, Europe and worldwide. For example, climate-driven shifts in the spatial distribution of humpback whales led to increased entanglements in the valuable Dungeness Crab (*Cancer magister*, *Cancridae*) fishery off California in 2014–2016 (Santora et al., 2020), and decreasing wintertime sea-ice is leading to increased overlap between shipping routes and protected species migrations in the Arctic and adjacent oceans (Hauser et al., 2018). Similarly, understanding movement and resulting distribution shifts is necessary to evaluate the impact of localized habitat protections upon population productivity at geographically distant locations that are linked via ontogenetic habitat shifts. For these and other reasons, improved understanding of movement has been called a “Grand Challenge” for habitat science (Thorson et al., 2021).

The past decade has seen an explosion of new and cost-effective technologies to study individual movements of marine species (Lowerre-Barbieri et al., 2019), resulting in broad deployment of many sampling methods including the following:

1. Conventional tags providing release and recovery location at known release and recovery dates.
2. Electronic (satellite and/or archival) tags providing environmental measurements that can be used to estimate animal location and behaviour between release and recovery (Lam et al., 2008).
3. Point-count data from resource surveys, providing a snapshot of habitat utilization resulting from movement.
4. Point-count data from fishery operations, typically occurring over a wider seasonal interval than surveys, where locational choices and gear deployment are not experimentally controlled and therefore must be corrected for during analysis.
5. Movement gates (e.g. upward facing acoustics, acoustic telemetry arrays, environmental DNA, acoustical moorings and videocamera arrays) providing a high-frequency measurement of aggregate flux (the product of local density and movement rates) and/or average residence time at a given location (Shertzer & Bacherel, 2020).
6. Habitat selection experiments (and similar process-research designs) specifying exogenous treatments in laboratory or field conditions to provide robust inference about the causal relationship between environmental features and movement decisions (e.g. Laurel et al., 2007).
7. Chemical, genetic and biological tracers, providing information on parentage (e.g. based on genetic markers of subpopulation structure), natal origin or lifetime habitat utilization (e.g. based on stable isotope, trace element or parasite markers of estuarine, nearshore and offshore habitat) (e.g. Spies et al., 2020).

1	INTRODUCTION	2
2	METHODS	4
2.1	Diffusion–advection–taxis movement process	4
2.2	Spatio-temporal model for spatial distribution	7
2.3	Metrics of resulting distribution shift	7
2.4	Parameter estimation	7
2.5	Case-study application	8
2.5.1	Investigating performance	8
2.5.2	Sensitivity analyses	9
3	RESULTS	9
4	DISCUSSION	13
	ACKNOWLEDGEMENTS	16
	DATA AVAILABILITY STATEMENT	16
	REFERENCES	16

These data types (Table 1) see varying usage across marine to freshwater systems. For example, conventional tags are often analysed in combination with fishery-dependent data for high-value commercial tunas that are broadly distributed across tropical and temperate oceans (Fournier et al., 1998; Lehodey et al., 2008), while surveys and movement gates are seeing increased use to infer movement in coastal shelf systems for species that cannot reliably be captured and released to deploy tags (O’Leary, Thorson, Ianelli, et al., 2020; Pinsky et al., 2013).

Given rapid declines in cost for new tagging and tracking technologies, researchers have declared 2018–2028 the “Biologging decade” (Lowerre-Barbieri et al., 2019). However, rapid deployment of new tracking technologies also raises analytical challenges, whereby new data types must be integrated with existing knowledge and monitoring programmes to provide a synthetic picture of animal movements (Hays et al., 2019). There is an ongoing effort to integrate all available data in its “rawest” form feasible throughout ecology, including (but not limited to) marine stock assessment (Maunder & Punt, 2013), terrestrial monitoring programmes (Kéry & Schaub, 2021), life-history theory (Thorson, 2020) and evolutionary analysis (Lee & Palci, 2015). Integrated models typically combine multiple data types to rectify deficiencies that arise in individual data sets (e.g. combining spatially extensive occupancy data with replicated counts at key monitoring sites) (Zipkin et al., 2017). Similarly, alternative movement technologies provide complementary information where, for example, movement gates (eDNA autosamplers and acoustics) can provide high temporal frequency at fixed locations, while resource surveys provide spatially distributed information about resulting distribution and density at seasonal or annual scales.

Spatially explicit integrated populations models (IPMs) have increasingly become useful tools for identifying changes in population

**TABLE 1** List of some (but not all) categories of information available to estimate movement, illustrating the potential benefit of a data-integrated movement model that can assimilate a wide variety of data types

Data type	Examples	Benefits	Difficulties
Conventional tags	Floy tags for sablefish	Widely used and low cost per tag	Low return rates; Dependent upon distribution of fishing effort
Survey data	Bottom trawl surveys	Available in many regions worldwide (Maureaud et al., 2021)	Reveals movement indirectly via habitat utilization
Fishery data	Fishery catch-per-unit-effort; Fishery effort	Available in regions without surveys	Reveals movement indirectly via habitat utilization Interpretation depends upon correct assumptions
Archival tags	mPAT; etc.	Provide high-resolution information about environmental conditions in utilized habitat	Relatively high cost per tag Expensive to process output
Movement gates	Upward facing acoustics Weirs	High temporal frequency Directly measures flux (aggregate movement)	Few technologies for use in marine environment Small spatial coverage
Habitat selection experiments	Laboratory selection experiments	Robust understanding of causal relationship between environment and movement	Difficult to “scale-up” results from laboratory to basin scales
Chemical markers	Stable isotope ratios compared with isoscapes; trace element suites	Retrospective study; useful for untaggable fish; cheaper than tags; elucidates behavioural differences	Non-lethal methods reduce temporal resolution (muscle, short-term); otoliths provide increased resolution (lifetime), but lethal
Genetic markers	SNP parentage assignment	Retrospective study; useful for untaggable fish	Requires full life-cycle movement models (including larval advection)
Parasite markers	Tetracotyle metacercarian parasite bio-tag data	Retrospective study; provides time-frame of exposure & location; useful for untaggable fish; cheaper than tags; elucidates behavioural differences	Complete knowledge of parasite biology & ecology required; taxonomic identification required; age of fish often required
Occurrence/density in predator food-habits samples	Sea birds as biological samplers for forage fish densities	Available opportunistically due to protected species sampling programmes; Predators are efficient samples of some prey species	Implicitly depends upon understanding of predator selection and functional responses; likely provides a noisy measurement of target species densities

status and distribution to manage human impacts on natural resources (Berger et al., 2017). Conventional stock assessment models make implicit assumptions about spatial processes, and spatial IPMs can improve these by letting data apply to only the portion of the population from which it was collected. However, spatial IPMs require specifying the fraction of abundance in each stratum that moves to other spatial strata (which we call “movement fractions”); information to specifying these movement fractions remains limited for spatial IPMs, especially for marine species where visual sampling of precise movement tracks can be difficult to attain (Goethel et al., 2021). Because fisheries management often relies on complex spatio-temporal policies (e.g. marine protected areas and area-specific quota recommendations), there is increasing need for the development of stock assessment models that are able to separately track abundance and age/size-structure across multiple spatial strata or interconnected population components, including straying among population or management units (Berger et al., 2017).

Although spatially stratified stock assessment models have been developed that can estimate movement fractions among spatial strata by incorporating typical fishery data as well as conventional, electronic and natural tags (Hampton & Fournier, 2001; Methot & Wetzel, 2013), there have been no previous models that simultaneously account for fine-scale spatial dynamics and spatially correlated process errors when fitting data from tags. Instead, fine-scale spatial IPMs such as SPM (Dunn et al., 2015) and SEAPODYM (Lehodey et al., 2008; Senina et al., 2019) use habitat preference functions to improve precision about local movement but lack spatially correlated process errors, while other fine-scale IPMs include spatial process errors but do not fit to tagging data (Cao et al., 2020; Kristensen et al., 2014). Additionally, no previous study has proposed methods to incorporate the full array of data that are informative about connectivity, movement, migration and habitat preference in marine populations (listed in Table 1). As IPMs continue to integrate insight from the movement ecology paradigm (Nathan et al., 2008), analysts

require methods that can incorporate the full array of available spatial data, incorporate information from habitat and oceanographic variables, and explicitly estimate movement (Bruneel et al., 2018). This synthesis will improve spatial IPMs by informing or directly specifying movement fractions and/or habitat preferences, thus reducing the number of estimated parameters (Goethel et al., 2021).

As an alternative to fitting tags within IPMs, parallel research has developed a growing suite of models (and associated estimation approaches) for individual movement at high spatial and temporal resolution. This includes, for example, a correlated random walk describing the animal location as a latent variable across time (Johnson et al., 2008) or fast approximations that discretize space and track animal residence within these discrete spatial cells while defining movement as a continuous-time Markov chain, CTMC (Wikle, 2003; Hanks et al., 2015). These approaches can then incorporate a habitat preference (a.k.a. resource selection) function to inform movement (Preisler et al., 2013; Michelot et al., 2019), and the stationary distribution of this CTMC is then a prediction of long-term habitat utilization (Wilson et al., 2018). Despite these improvements, there are few examples of CTMC or other tracking models that integrate point-count data from surveys and fisheries, although again noting Lehodey et al. (2008) and Senina et al. (2019) as exceptions.

In this study, we demonstrate the potential to integrate a broad range of information regarding movement in marine species within a single model with high spatial resolution. The approach decomposes an instantaneous movement rate into components representing diffusion, advection and taxis and incorporates climate and habitat information to inform all three processes. It then integrates these movement processes over seasonal and annual intervals to identify path-dependent movement probabilities and uses these probabilities to fit the movement model to survey, fishery and tagging data. Throughout, we specifically discuss how to extend this framework to integrate information from movement gates, archival tags and other alternative technologies, while providing a publicly available R package to demonstrate the analytical approach. In particular, we demonstrate this R package by fitting summer survey, summer and winter fishery catch-and-effort, and conventional tagging data for Pacific cod (*Gadus macrocephalus*, Gadidae) in the Bering Sea, using bottom temperature and bathymetry data to interpret their seasonal habitat preferences. Finally, we demonstrate how results can be integrated into an existing stock assessment model for Pacific cod, which has exhibited rapid poleward distribution shifts in response to warming conditions over the past decade. We also explore the sensitivity of model results to each fitted data set to determine the relative importance of tags, survey and fishery data in this application.

## 2 | METHODS

To fit multiple data types that are informative about movement (Table 1), we discuss a model for instantaneous movement among discrete spatial areas representing diffusion, advection and taxis (a DAT model). We integrate this instantaneous process over discrete time interval  $t$  to

define a movement matrix  $\mathbf{M}(t)$  that describes the fraction of animals  $m(g_2, g_1, t)$  moving from cell  $g_1$  to  $g_2$  during interval  $t$ . We call  $\mathbf{M}(t)$  the “movement matrix” and  $m(g_2, g_1, t)$  a “movement fraction.” We then use this movement matrix  $\mathbf{M}(t)$  within a spatio-temporal model that predicts numeric abundance  $n(g, t)$  for each grid cell  $g$  and time interval (see Appendix S1: Tables S1–S4 for summary of all notation), as well as movement of individual animals and associated tags among cells. Estimates of movement affect our predictions of tag locations, as well as expected changes in survey abundance and fishery catch-and-effort data over time; movement parameters are therefore informed by all data types and this justifies our development of a joint model.

Grid cells  $g \in \{1, 2, \dots, n_g\}$  are evenly spaced within a spatial domain where every grid cell is adjacent to one to four other grid cells (i.e. all cells are connected). We track the sequence of seasons and years by index  $t$ , such that movement in each interval  $t$  can be predicted in part by the season  $u_t$  or year  $y_t$  associated with interval  $t$ ; this then allows for differences in environmental values and habitat preferences among seasons representing, for example, spawning versus foraging preferences (e.g. Lehodey et al., 2008). Although we discretize space and time in the following, future research could fit individual data sets using continuous space and/or time while using the same underlying movement process and parameters, and this could be appropriate for data that are informative about higher-resolution processes (e.g. electronic tags).

### 2.1 | Diffusion-advection-taxis movement process

We define the diffusion-advection-taxis process among discrete cells to predict an instantaneous movement (a.k.a. transition) rate  $m^*(g_2, g_1, t)$  from cell  $g_1$  to cell  $g_2$  at any moment during interval  $t$ . Fish can only move between adjacent cells as time interval  $\Delta t \rightarrow 0$ , so  $m^*(g_2, g_1, t) = 0$  for any two cells that are not adjacent. We use an asterisk to indicate an instantaneous rate, for example  $\mathbf{M}^*(t)$  for the matrix of instantaneous movement rates, and subsequently drop this asterisk, for example  $\mathbf{M}(t)$ , to indicate a process that is integrated over the entire time interval. Instantaneous movement matrix  $\mathbf{M}^*(t)$  is a continuous-time Markov chain CTMC (e.g. Hanks et al., 2015), and deriving movement matrix  $\mathbf{M}(t)$  from this allows for path-dependent movement between non-adjacent locations (i.e. movement between two locations is higher when there is a corridor of preferred habitats between those locations).

We decompose instantaneous movement matrix  $\mathbf{M}^*(t)$  into the following:

1. Diffusion matrix  $\mathbf{D}^*(t)$ , representing the undirected movement of animals away from their present location, treated as a random walk process.
2. Taxis matrix  $\mathbf{Z}^*(t)$  representing the directed movement of animals towards preferred habitats.
3. Advection matrix  $\mathbf{V}^*(t)$  representing the passive transport of animals along one or more pre-specified vector fields, for example ocean currents.

where

$$m^*(g_2, g_1, t) = d^*(g_2, g_1, t) + z^*(g_2, g_1, t) + v^*(g_2, g_1, t) \quad (1)$$

This decomposition distinguishes taxis and advection, although other authors use these terms interchangeably. Previous authors (Sibert et al., 1999; Wikle, 2003) have derived this CTMC among discrete spatial areas from a partial differential equation (PDE) for movement in continuous space, and we simplify our presentation by not repeating a formal derivation. Conceptually, however, a partial differential equation (PDE) for diffusion specifies that density decreases where the second derivative is negative (i.e. density  $d(s)$  is higher at location  $s$  than the average of nearby locations), and a PDE for taxis specifies that densities flow towards areas with higher preference (i.e. densities decrease where preference is lower than nearby locations and correspondingly increase in those nearby preferred locations). The CTMC model replaces this differential equation with analogous "movement rules" among grid cells, where diffusion involves movement away from the current grid cell (i.e. diagonal elements of  $D^*(t)$  are negative), while taxis involves movement towards grid cells with higher preference (i.e.  $v^*(g_2, g_1, t)$  is positive if and only if  $g_2$  has higher preference than  $g_1$  in time  $t$ ).

Previous studies for marine fishes have developed "Markov movement models" (e.g. Heifetz & Fujioka, 1991; Webster et al., 2013; Hanselman et al., 2015) that estimate a pairwise movement rate for each pair of locations. The diffusion-advection-taxis model differs from these previous "unstructured Markov" models in two important ways:

1. Our approach uses environmental data to predict habitat preferences, thereby linking movement to habitat information;
2. As a result of the first, we can define movement at fine spatial scales using a small number of habitat preference and diffusion parameters. This allows us to downscale predictions of movement to much finer scales than unstructured Markov models, without an explosion in the number of pairwise movement rates that must be estimated.

This decomposition has some similarity to that used in the spatial population model (SPM) (Dunn et al., 2015), for example for Antarctic toothfish (Mormede et al., 2014). However, SPM decomposes seasonal movement probabilities into preference and/or diffusion, whereas we decompose instantaneous movement rates explicitly into the three components listed. One consequence of this distinction is that SPM does not account for path dependence (the impact of intervening habitats on movement probabilities for two non-adjacent cells) when integrating paths to calculate seasonal movement probabilities.

Taxis is predicted using environmental data  $x(g, t, p)$  for  $p$  habitat variables, such as those commonly available for species distribution (e.g. bathymetry) and climate modelling (e.g. ocean temperature). In the following, we apply basis expansion to transform these variables  $x(g, t, p)$  into a vector of covariates  $x(g, t, k)$  used to model habitat

preferences, where  $k$  indexes the resulting basis functions. These transformations can involve any standard basis expansion, for example using splines, interactions or spatial kernels, and the analyst can explore alternative formulations during model development (in our software using the formula interface in R).

Specifically, diffusion is the random movement of individuals from any cell  $g_1$  to adjacent cells  $g_2$ . In the following, we specify a constant diffusion rate:

$$d^*(g_2, g_1, t) = \begin{cases} e^{2\beta} & \text{if } g_1 \text{ and } g_2 \text{ are adjacent} \\ - \sum_{g' \neq g_1} d^*(g', g_1, t) & \text{if } g_1 = g_2 \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where  $\beta$  is the log-diffusion rate and  $g'$  (used in the 2nd row of Eq. 2) is an index used to sum across columns such that columns sum to zero. We note that future research could easily extend this notation to specify that diffusion is a function of local environmental conditions (e.g. Wikle, 2003), but do not elaborate here for clarity of presentation.

Similarly, advection represents movement due to a vector field, for example the passive drift of larvae due to ocean currents and tidal flows. We envision the analyst specifying  $n_m$  vector fields indexed by  $m$ :

$$v^*(g_2, g_1, t) = \sum_{m=1}^{n_m} \gamma(m) w^*(g_2, g_1, t, m), \quad (3)$$

where  $w^*(g_2, g_1, t, m)$  is transport rate from  $g_1$  to adjacent cell  $g_2$  during interval  $t$  for vector field  $m$  (e.g. oceanographic currents for passive drift), and  $\gamma(m)$  are estimated parameters relating vector field  $w^*(g_2, g_1, t, m)$  to advection rate  $v^*(g_2, g_1, t)$ . We note that  $w^*(g_2, g_1, t)$  would likely be defined to conserve abundance, that is  $\sum_{g_2=1}^{n_g} w^*(g_2, g_1, t, m) = 0$ .

Finally, taxis represents directional movement towards preferred habitats, where habitat preference  $h(g, t)$  is defined as a function of local environmental conditions:

$$h(g, t) = \sum_{k=1}^{n_k} \alpha(k) x(g, t, k), \quad (4)$$

where  $x(g, t, k)$  is a basis expansion of environmental layers  $x(g, t, p)$ , for example via application of splines and variable interactions, and  $\alpha(k)$  are estimated environmental-preference parameters. Taxis is then defined based on local differences in the habitat preference function:

$$z^*(g_2, g_1, t) = \begin{cases} h(g_2, t) - h(g_1, t) & \text{if } g_1 \text{ and } g_2 \text{ are adjacent} \\ - \sum_{g' \neq g_1} \{h(g_2, t) - h(g_1, t)\} & \text{if } g_1 = g_2 \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

We note that only the difference in preference  $h(g_2, t) - h(g_1, t)$  is used, and therefore, habitat preference covariates  $x(g, t, k)$  should

not include an intercept term. Deriving taxis from a habitat preference function was proposed by Brillinger (2012); it has been adapted in other studies (Preisler et al., 2013), and it appears to perform similarly to alternatives such as Langevin diffusion (Michelot et al., 2019).

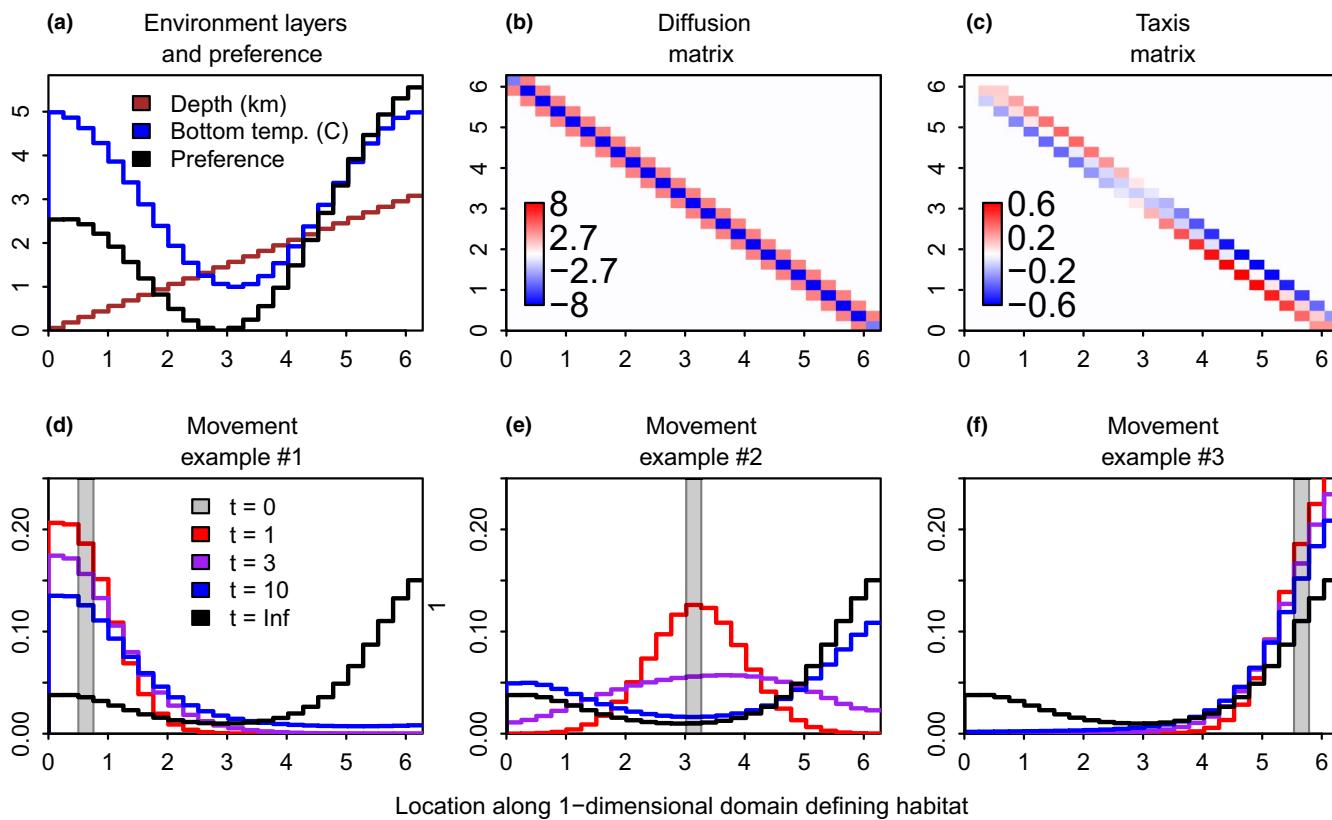
We note that advection and taxis are both directional movement and therefore may seem similar or even redundant. However, taxis is a conservative vector field such that differences in habitat preference  $h(g_1, t)$  and  $h(g_2, t)$  between two locations will always be identical regardless of the path followed: this characteristic allows taxis to be uniquely defined given habitat preference function  $h$ , such that covariates can be used flexibly to define  $h$ . By contrast, advection is not restricted to be a conservative vector field. In particular, advection can represent cyclic or chaotic movement due to passive transport in eddies, streamers and other oceanographic features, which taxis cannot approximate due to its specification as a conservative vector field.

Given this definition for instantaneous movement rate  $\mathbf{M}^*(t)$ , we then integrate movement fraction  $\mathbf{M}(t)$  over the interval starting at time  $t$  as:

$$\mathbf{M}(t) = e^{\mathbf{M}^* \Delta t}, \quad (6)$$

where we specifically use the matrix-exponential operator and  $\Delta t$  is the duration of the interval starting at time  $t$ . Given that  $\mathbf{M}^*$  is a Metzler matrix (i.e.  $m^*(g_2, g_1, t) \geq 0$  for all  $g_2 \neq g_1$ ) and has columns that sum to zero,  $\mathbf{M}(t)$  will be non-negative (i.e.  $m(g_2, g_1, t) \geq 0$ ) and will have columns that sum to one. As a result, transformation  $\mathbf{M}(t)$  conserves abundance and will have a leading eigenvalue of 1.0 such that the leading eigenvector represents the hypothetical long-term habitat utilization if those environmental conditions persisted indefinitely; this property of a continuous-time Markov chain has been noted previously elsewhere (Wilson et al., 2018; Thorson et al., 2021).

We illustrate these calculations in a simplified one-dimensional spatial domain to clarify concepts (Figure 1) and provide an R package on GitHub to facilitate rapid deployment for other stocks (<https://github.com/James-Thorson-NOAA/ATM>) while using release 1.0.0 here. The package uses a “formula” interface to allow users to specify non-linear and interactive covariate responses using covariates provided in widespread geospatial data types (Pebesma, 2019) to facilitate model exploration.



**FIGURE 1** Simplified illustration of diffusion–taxis movement model along a 1-dimensional spatial domain defining habitat,  $0 \leq x \leq 2\pi$ , binned evenly into 25 grid cells  $g \in \{1, 2, \dots, 25\}$ , where depth  $x_1(g) = x/2$ , temperature  $x_2(g) = 3 + 2\cos(x)$ , habitat preference  $h(g) = x_1(g) + x_2(g)$  (see panel A), diffusion rate  $\tau = 4$ , and assuming a reflective boundary (i.e. individuals do not emigrate). The instantaneous movement rate matrix  $\mathbf{M}^*$  from each to every other grid cell is the sum of the diffusion rate matrix  $\mathbf{D}$  (panel B) and the taxis rate matrix  $\mathbf{Z}$  (panel C), where  $d(g_1, g_2) = z(g_1, g_2) = 0$  for any two grid cells that are not adjacent  $g_1 \neq g_2 \pm 1$ , otherwise  $d(g_1, g_2) = \tau$  and  $z(g_1, g_2) = h(g_1) - h(g_2)$ , and where the diagonal  $d(g, g)$  and  $z(g, g)$  is defined such that columns sum to one (i.e. abundance is conserved during movement). For an individual starting at  $t = 0$  in grid cell  $g = 3$  (panel D),  $g = 13$  (panel E) or  $g = 23$  (panel F), we then show movement probabilities after one interval,  $\text{matexp}(\mathbf{M}^*)$ , three intervals,  $\text{matexp}(3\mathbf{M}^*)$ , ten intervals,  $\text{matexp}(10\mathbf{M}^*)$ , or its limit after a long time has passed (i.e. long-term habitat utilization, calculated as the dominant eigenvector of  $\text{matexp}(\mathbf{M}^*)$ ). We note that long-term habitat utilization is identical regardless of the initial location, and it resembles (but is not identical to) the habitat preference function

## 2.2 | Spatio-temporal model for spatial distribution

We next embed this DAT movement model within a spatio-temporal model for numerical abundance that can be used to predict changes in distribution resulting from alternative or future values of environmental layers. In this sense, our spatio-temporal DAT model represents a “hybrid” species distribution model (hybrid SDM), with properties that are intermediate between “correlative” spatio-temporal SDMs and “mechanistic” SDMs that include behavioural information about individual movement (Dormann et al., 2012).

Specifically, we estimate numerical abundance  $n(g, t)$  as a latent variable. This involves specifying a state-space model for  $n(g, t)$  for every year  $t$ :

$$\log(n(t)) = \begin{cases} \delta(t) + \varepsilon(t) & \text{if } t=1 \\ \delta(t) + \mathbf{M}(t-1)\mathbf{n}(t-1) + \varepsilon(t) & \text{if } t>1 \end{cases}, \quad (7)$$

where  $\delta(t)$  represents interannual variation in total abundance and  $\varepsilon(t)$  is a Gaussian Markov random field representing residual spatio-temporal variation in  $n(g, t)$ :

$$\varepsilon(t) \sim \begin{cases} \text{MVN}\left(0, \sigma_0^2 \mathbf{Q}^{-1}\right) & \text{if } t=1 \\ \text{MVN}\left(0, \sigma_\varepsilon^2 \mathbf{Q}^{-1}\right) & \text{if } t>1 \end{cases}, \quad (8)$$

where the initial process error has a different variance ( $\sigma_0^2$ ) than the variance of subsequent process errors ( $\sigma_\varepsilon^2$ ), and  $\mathbf{Q}^{-1}$  is an approximation to the Matérn correlation among grid cells. This Matérn correlation function represents a decline in correlation as a function of distance between any two locations  $s_1$  and  $s_2$ , and we specifically estimate a  $2 \times 2$  transformation matrix  $\mathbf{H}$  such that correlations decline as a function of transformed Euclidean distance  $\mathbf{H}(s_1 - s_2)$ . This linear transformation  $\mathbf{H}$  represents geometric anisotropy and governs the rate at which correlations decline in different cardinal directions (Lindgren et al., 2011; Thorson et al., 2015). We additionally use the SPDE approximation to the correlation  $\mathbf{Q}^{-1}$  among grid cells (Lindgren et al., 2011), where inverse-correlation  $\mathbf{Q}$  in Eq. 8 is specified via sparse matrices calculated using R-INLA (Lindgren, 2012).

## 2.3 | Metrics of resulting distribution shift

Many stock assessment models operate at a coarser spatial resolution than the spatial scale for environmental layers that are feasible here. To demonstrate how results from the DAT model can be adapted for use in a spatially stratified assessment model, we therefore use “change-in-support” methods to coarsen the resolution while calculating abundance and movement among strata. This process involves calculating the abundance-weighted average movement among each pair of grid cells, while summing abundance for all grid cells within each stratum.

First, we calculate abundance for each stratum  $o$  as:

$$n_{\text{coarse}}(o, t) = \sum_{g=1}^{n_g} p(g, o) n(g, t), \quad (9)$$

where  $p(g, o)$  is the proportion of grid cell  $g$  that is within stratum  $o$ . Next, we calculate abundance  $n_{\text{coarse}}(o_2, o_1, t)$  moving from each stratum  $o_1$  to each other stratum  $o_2$ :

$$n_{\text{coarse}}(o_2, o_1, t) = \sum_{g_2=1}^{n_g} \sum_{g_1=1}^{n_g} n(g_1, t) p(g_2, o_2) p(g_1, o_1) m(g_2, g_1, t), \quad (10)$$

where coarsened movement fraction  $m_{\text{coarse}}(o_2, o_1, t)$  from stratum  $o_1$  to each other stratum  $o_2$  are then calculated by normalizing for each column:

$$m_{\text{coarse}}(o_2, o_1, t) = \frac{n_{\text{coarse}}(o_2, o_1, t)}{\sum_{o'=1}^{n_o} n_{\text{coarse}}(o', o_1, t)}, \quad (11)$$

where  $m_{\text{coarse}}(o_2, o_1, t)$  can be used similarly to the movement rates that are calculated from previous Markov models applied to conventional tagging data (Hanselman et al., 2015; Heifetz & Fujioka, 1991).

We also note that either fine-scale or coarsened movement fractions ( $m(g_2, g_1, t)$  and  $m_{\text{coarse}}(o_2, o_1, t)$ , respectively) can be used to calculate the stationary distribution for abundance  $\mathbf{n}(t)$ , representing the equilibrium proportion of abundance in each cell  $g$  or stratum  $o$  that would occur if current conditions persisted indefinitely. Specifically, the stationary distribution is calculated as the dominant eigenvector of  $\mathbf{M}(t)$  or  $\mathbf{M}_{\text{coarse}}(t)$ , normalized to sum to one (noting that the eigenvalue associated with the stationary distribution is 1.0 given that movement is defined to conserve abundance). This stationary distribution can be calculated given conditions in each individual year-season combination or for the product of movement over all seasons in a given year to yield the “annualized” stationary distribution.

## 2.4 | Parameter estimation

We estimate parameters by identifying their values that maximize the log-likelihood of available data while integrating across the probability of random effects. We approximate this integral using Template Model Builder (Kristensen et al., 2016) within the R statistical environment (R Core Team, 2017). We treat  $\log(n(t))$  as a latent process, that is using a state-space parameterization for the Gaussian Markov random field  $\varepsilon_t$  representing spatio-temporal variation in Eqs 7 and 8 and then integrating across  $\log(n(t))$  by treating it as a random effect. The log-likelihood is optimized using a gradient-based Nelder-Mead algorithm, followed by two iterations of a Newton algorithm to decrease the final gradient. We confirm that models are converged by (1) ensuring that the gradient of the log-likelihood with respect to each fixed effect is  $<0.0001$  and (2) that the Hessian matrix is positive definite. We then calculate standard errors for parameters and derived quantities using this Hessian matrix and a generalization of the delta-method (Kass & Steffey, 1989),

or by sampling from the joint precision matrix for fixed and random effects. All parameter estimation is done using release number 1.0.0 of a new R package *ATM* that is publicly available on GitHub (<https://github.com/James-Thorson-NOAA/ATM>), and this paper serves as introduction for that package. *ATM* shares software dependencies (i.e. utility functions in package *FishStatsUtils*) with package *VAST* (Thorson & Barnett, 2017) to facilitate comparison of parameter estimates across these two model platforms.

The maximum-likelihood estimation approach requires specifying a probability distribution for data given parameters. We introduce this likelihood for survey data, fishery catch-per-unit-effort, conventional tags, electronic tags, movement gates and ecogeographical tracers in Appendix S2.

## 2.5 | Case-study application

We apply this model to three types of data for Pacific cod in the eastern and northern Bering Sea. We specify a resolution involving two 6-month seasons per year (summer: April–September and winter: October–March) and 38 years (1982–2019), and using 100 square grid cells each covering a 100 km × 100 km area to discretize the fully connected spatial domain. The base model took under two hours to estimate parameters and calculate associated standard errors on a personal laptop starting from uninformative parameter values.

We fit three types of data to inform movement rates (see Appendix S3 for details):

1. Summer bottom trawl survey from 1982 to 2019 with fewer years sampled in the northern than eastern Bering Sea.
2. Winter and summer longline fishery catch-and-effort from 1996 to 2019, using the total number of hooks as the measure of effort  $a_f$ .
3. Tag release and recapture locations for 2670 conventional tags from 1982 to 2006 and 72 archival tags released from 2002 to 2004 (Nichol et al., 2007), recovered by a mixture of commercial fisheries and resource surveys. We measure fishery effort as the total number of longline fishery hooks in each grid cell for a given season and year and assume that this is proportional to recapture probability conditional on the tag being present in a given cell. We note that conventional and archival tags were also recovered by other fisheries and resource surveys that have a different spatial footprint. We do not have access to high-resolution locational information for these other fisheries and recommend future research to reconstruct this information to better interpret recapture probabilities for Pacific cod conventional tags. However, an alternative hypothesis is that recapture probabilities are spatially constant, and we confirm that results are qualitatively similar when a uniform distribution of fishing effort is specified instead (results not shown).

We fit these data using two environmental layers (see Appendix S4: Figures S1–S3):

- A Seafloor bathymetry, previously developed for use in designating essential fish habitat (see Appendix S4: Figure S1) and compiled from several sources (Zimmermann & Benson, 2013; Zimmermann & Prescott, 2018).
- B Bottom temperature, using hindcasted values derived from the Bering10K model (Hermann et al., 2016; Kearney et al., 2020). We use the simulated bottom temperature values averaged across the bottom 5 m above the seafloor and then average this across months and cells to conform to the seasonal and spatial resolution of the DAT model (Appendix S4: Figures S2 and S3).

We hypothesize that Pacific cod will exhibit a non-linear habitat preference for each environmental layer. We therefore explore estimating a separate cubic basis-spline with three knots (four parameters total) representing the habitat preference for each layer in each season. However, preliminary exploration suggested that it was not feasible to estimate a separate temperature preference for each season, so we instead used three splines (12 covariate response parameters total) representing a non-linear response to summer bathymetry, winter bathymetry and bottom temperatures.

### 2.5.1 | Investigating performance

We investigate model performance in a variety of different ways:

1. We visualize probability integral transform (PIT) residuals (Smith, 1985) calculated from the conditional simulation distribution for fishery and survey point-count data. This is useful to detect departures from the specified distribution for those data.
2. We visualize the predicted and observed count of conventional tags that are detected in each grid cell upon recapture and again calculate PIT residuals from this distribution under the assumption that counts follow a Poisson distribution.
3. We visualize the predicted movement of individuals from a hypothetical release location in the southern middle domain, if released during the beginning of a prolonged cold stanza (summer 2007) or warm stanza (summer 2014), and projecting their subsequent movement given environmental conditions over the following five years.
4. We visualize predicted habitat preference  $h(g, t)$  for selected warm (2002, 2017/2018) and cold years (2012) during summer and winter, and also the predicted log-abundance in those same seasons and years.

We also conduct a “self-test” simulation experiment with 25 replicates to corroborate that the model can estimate parameters given the quantity of data available. In each replicate, we:

- A simulate new response values for survey and fishery catch-and-effort data as well as the location of tag recaptures, conditional upon the model structure, estimated fixed effects, and predicted random effects from the case-study model;

- B refit the original estimation model to these new simulated data and record the estimates of habitat preference  $\hat{h}(g, t)$ ;
- C calculate the centred habitat preference  $h_{\text{centred}}(g, t) = h(g, t) - \left( \sum_{g=1}^{n_g} h(g, t) \right) / n_g$  and  $\hat{h}_{\text{centred}}(g, t) = \hat{h}(g, t) - \left( \sum_{g=1}^{n_g} \hat{h}(g, t) \right) / n_g$  in recognition that only the difference in preference between two locations in a given year is used when calculating movement;
- D compare estimates of centred habitat preference  $h_{\text{centred}}(g, t)$  used to simulate data with estimates of habitat preference  $\hat{h}_{\text{centred}}(g, t, r)$ . We specifically compute the proportion of variation explained:

$$V(t) = 1 - \frac{\sum_{g=1}^{n_g} \{h_{\text{centred}}(g, t)\}^2}{\sum_{g=1}^{n_g} \{h_{\text{centred}}(g, t) - \hat{h}_{\text{centred}}(g, t)\}^2} \quad (12)$$

- E repeat steps A-D for each of 25 replicates, and compile  $V(t)$  across all replicates and years.

A well-performing model will result in  $V(t)$  close to 1.0, while a model that cannot explain variation in preference will result in  $V(t)$  close to 0.

The Pacific cod stock assessment has explored a spatially stratified age- and length-structured stock assessment model implemented using Stock Synthesis (Methot & Wetzel, 2013; Thompson et al., 2020); this model separately tracks abundance at age in the eastern Bering Sea and northern Bering Sea (where each stratum is defined as the spatial footprint of an associated bottom trawl survey programme), while estimating movement between those two strata. We therefore use the DAT model to explore three metrics for the proportion of biomass in the northern Bering Sea:

- A **Abundance:** The proportion of abundance  $n_{\text{coarse}}(o, t) / \sum_{o'=1}^{n_o} n_{\text{coarse}}(o', t)$  in the eastern Bering Sea relative to total abundance (where each is defined as the spatial footprints of the eastern and northern Bering Sea bottom trawl surveys).
- B **Equilibrium:** The stationary distribution resulting from abundance-expanded movement among strata,  $\mathbf{M}_{\text{coarse}}(t)$ , given conditions in time  $t$ .
- C **Forward projection:** The proportion of abundance in each stratum initialized from the stationary distribution for abundance in time  $t = 1$ , but projected forward deterministically for subsequent times using abundance-expanded movement among strata,  $\mathbf{M}_{\text{coarse}}(t)$ .

The first metric depends only upon predicted abundance  $n(g, t)$  and is therefore informed primarily by survey and fishery data. The second and third metrics, by contrast, are calculated entirely from  $\mathbf{M}_{\text{coarse}}(t)$ , which in turn combines information about distribution from survey/fishery data with information about diffusion and taxis. These two metrics differ in whether they assume that present conditions persist indefinitely (*equilibrium*) or using the time series of environmental conditions in each preceding time (*forward projection*).

They therefore correspond to alternative definitions of prevailing environmental conditions and are analogous to "Moving average" and "Dynamic" calculations for biological reference points for use in fisheries management (O'Leary, Thorson, Miller, et al., 2020).

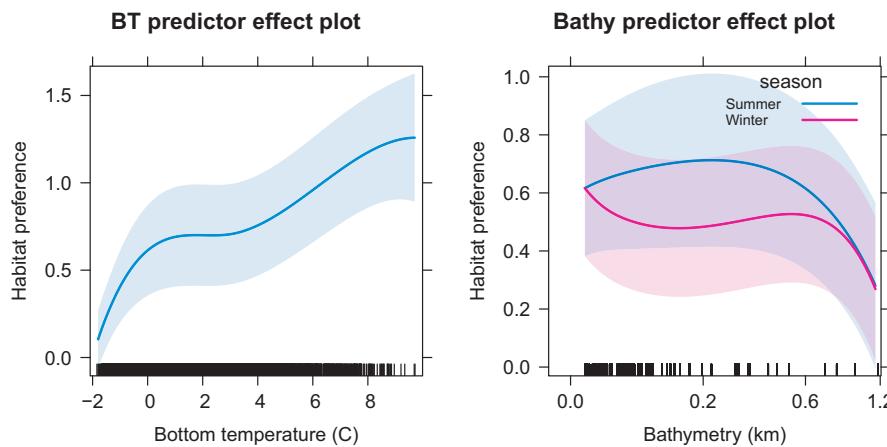
Finally, we also explore the impact of including results from this model in one of the assessment models (named Model 20.7) explored in the 2020 stock assessment for Bering Sea Pacific cod; see Appendix S5 for details regarding this assessment model. To explore the impact of using movement indices in this assessment, we first coarsen the spatial resolution of our movement model to calculate seasonal movement  $\mathbf{M}_{\text{coarse}}(t)$  between these two spatial strata using Eqs 9–11. We next aggregate across both seasons, such that annual movement fractions incorporate movement during summer and the following winter (i.e. summer-to-summer movement). We include this coarsened and annualized movement fraction within the spatially stratified Pacific cod assessment model, and extract resulting estimates of age 1+ abundance in the eastern and northern Bering Sea, converting this to the fraction of total abundance in the eastern Bering Sea, and compare this fraction with the survey index fraction in years with consistent bottom trawl sampling across both strata (2010, 2017, 2018 and 2019). A well-performing assessment model is expected to estimate a similar fraction in the eastern and northern Bering Sea to what was observed in this survey.

## 2.5.2 | Sensitivity analyses

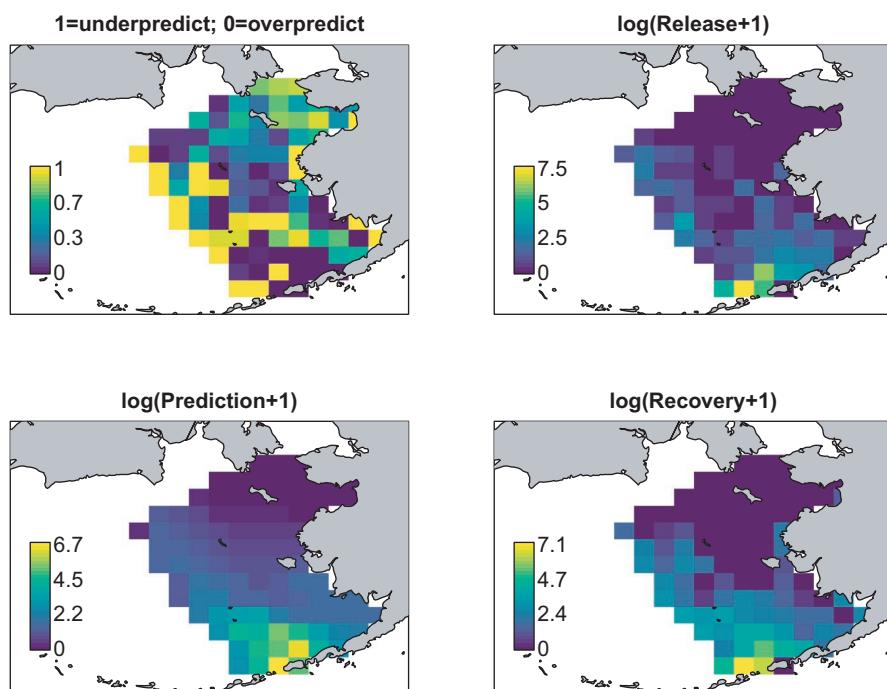
We explore the sensitivity of model results to sequentially excluding one of the three major data types: (1) conventional tags; (2) survey data; and (3) fishery catch-and-effort data. This sensitivity test is designed to illustrate how each data type is influencing results. We acknowledge that the model developed here differs in many ways from conventional methods for analysis of tags, survey data and fishery CPUE data. We therefore distribute the code as a publicly available R package and encourage future research comparing results to alternative implementations using individual data sets and/or alternative model structures.

## 3 | RESULTS

By fitting to fishery catch-and-effort, survey and tagging data (the "base model"), the diffusion-advection-taxis model estimates that adult Pacific cod prefer increasing temperatures (within the ranges encountered here), with a rapid decline in habitat preference for temperatures  $<0^\circ\text{C}$  (Figure 2A). Similarly, adults show a strong preference (high peak) for depths 50–400 m during summer and a broader depth preference during winter (Figure 2B). Diagnostics for model goodness-of-fit indicate little residual spatial pattern in the match between predicted and observed recapture locations (Figure 3), and also survey and fishery residuals have quantile residuals that follow the expected one-to-one line (Appendix S4: Figures S4 and S5). Similarly, the self-test simulation experiment confirms that data are



**FIGURE 2** Habitat preference (y-axis) for covariate values (x-axis) for each modelled covariate, showing mean response (line) and 95% confidence interval (shaded area), visualized using the R package effects (Fox & Weisberg, 2018)



**FIGURE 3** Map of probability integral transform (PIT) residuals for the recovery location of all tags (top-left panel), the release location for tags (top-right column), the predicted recovery location (bottom-left panel) and the observed recovery location (bottom-right panel), when summing all tag releases, recoveries and quantile residuals across all tags, seasons and years. The residuals are provided to illustrate any systematic spatial patterns in the sign of residuals

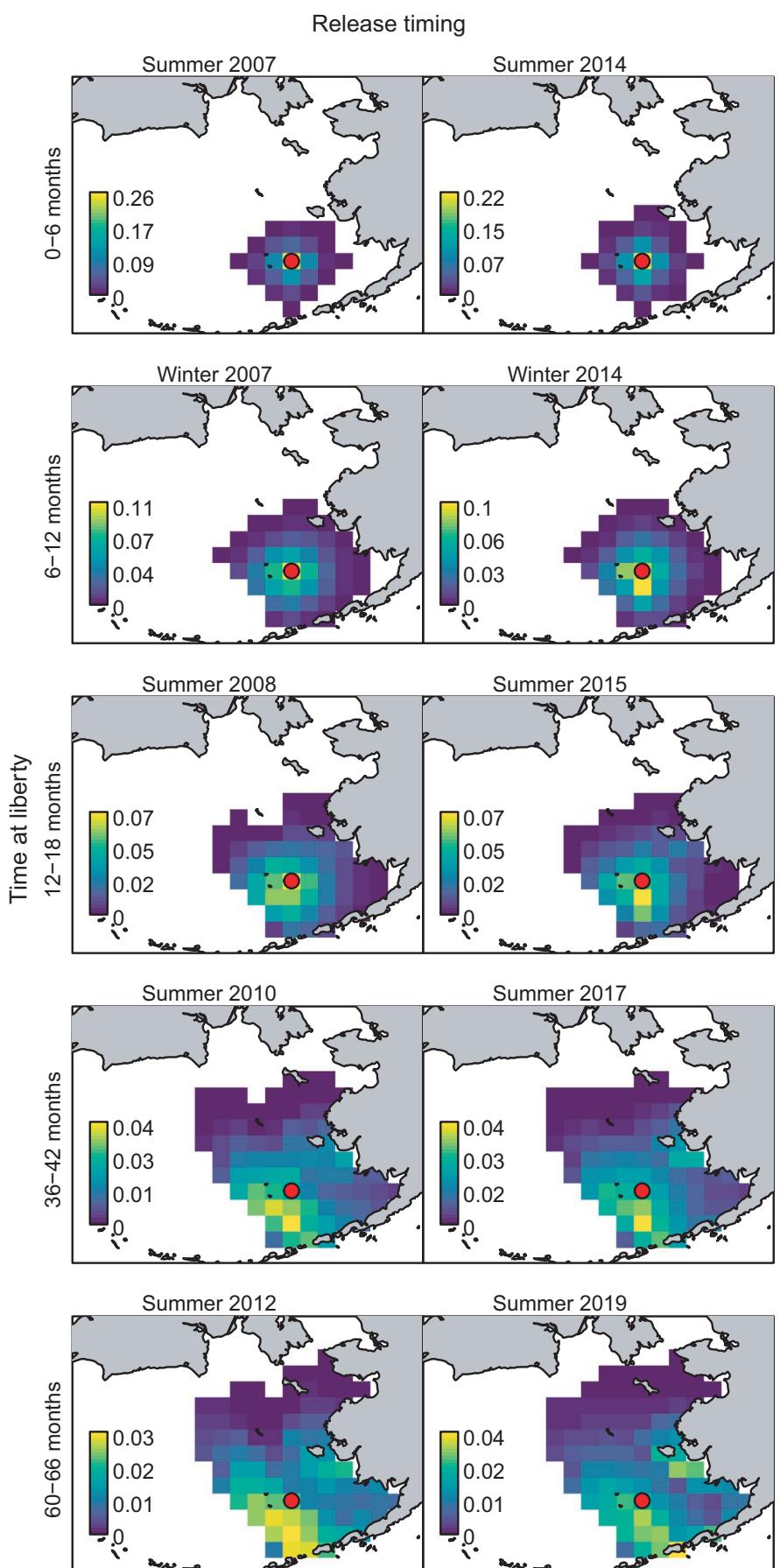
sufficient to explain over 2/3 of variation (0.69, see Appendix S4: Figure S6) in centred habitat preferences under ideal circumstances (i.e. given that the model is correctly specified).

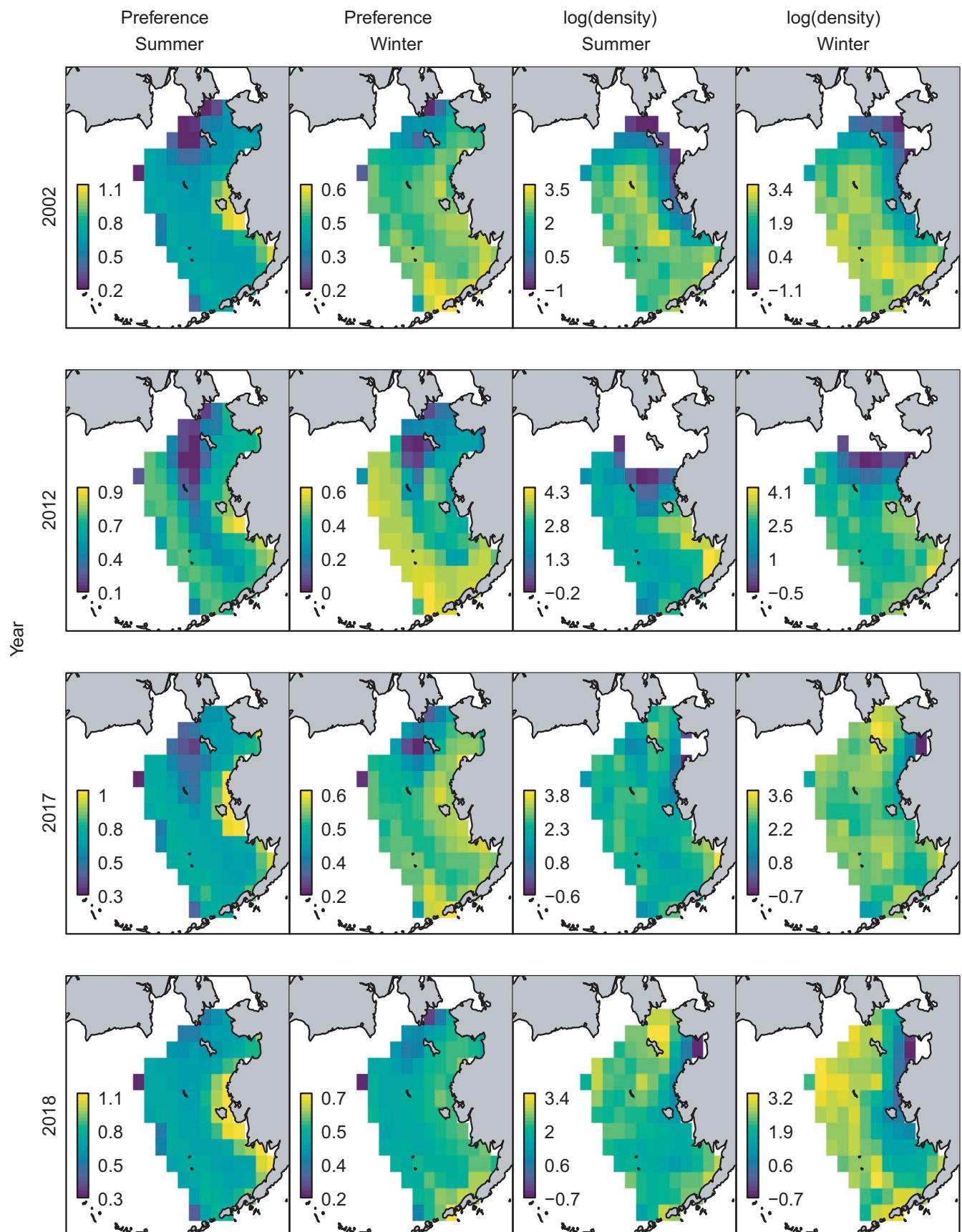
The estimated habitat preference function (and resulting taxis) shows only small differences in movement for individuals during a cold stanza (2007–2012) and warm stanza (2014–2019) (Figure 4, comparing columns), including between summer and the subsequent winter distribution (Figure 4, comparing 1st and 2nd rows). Despite this broad similarity in movement between warm and cold stanzas, the model does predict slight increases in northward movement during warm stanzas, for example elevated movement probabilities from the southern middle domain release location to south of St. Lawrence Island for a release in 2014 relative to that in 2007 (comparing bottom row of Figure 4). Differences in habitat preference are more clearly illustrated by comparing the estimated winter preference function in warm years (2002/2019) and a cold year (2012) (Figure 5, 2nd column). These differences are above the approximately 0.1 standard error in estimated preferences (see Appendix

S4: Figure S7). The predicted densities show greater differences among years, where the northern Bering Sea has elevated densities in 2017/2019 relative to 2002/2012 in both summer and winter seasons (Figure 5, 3rd and 4th columns).

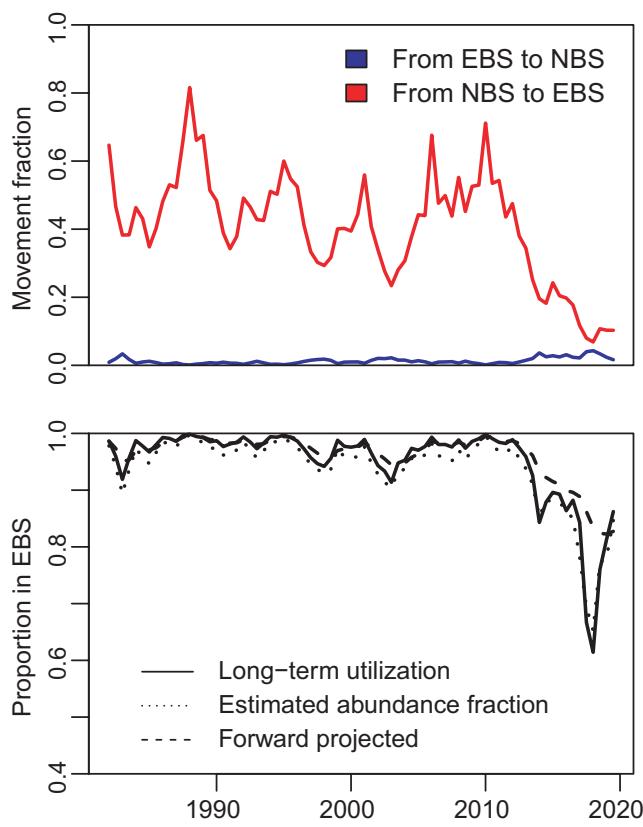
Next, we visualize movement between the eastern and northern Bering Sea to understand whether we can predict the recent increase in proportion in the northern stratum of the most recent stock assessment. The model predicts that 30%–70% of individuals in the northern Bering Sea would move back to the eastern Bering Sea during a given year from 1982 to 2012 (Figure 6). This proportion started dropping in 2013 and was 5%–10% in 2017–2019, with an associated but smaller increase in the proportion predicted to move from the eastern to northern Bering Sea strata. Given those movement fractions, the *equilibrium* metric for proportion closely matched the *abundance* metric for the proportion in the northern stratum, reaching 20%–40% from 2017 to 2019 (see Methods for detailed calculation). As expected, the *forward projection* metric lagged behind the *equilibrium* metric and reached 20% in the

**FIGURE 4** Maps of movement probabilities (see colour bar in each panel and with probabilities <1% of maximum plotted as white) given a southern release location in summer 2007 during a cold stanza (left column) and summer 2014 during a warm stanza (right column)





**FIGURE 5** Maps of habitat preference in summer (1st column) and winter (2nd column), as well as predicted log-abundance in summer (3rd column) and winter (4th column) for selected years (rows) highlighting a warm year (2002), cold year (2012) and then recent warm years (2017–2018)



**FIGURE 6** Estimated summer-to-summer movement fraction (y-axis) from NBS to EBS (red line) and EBS to NBS (blue line) in each year (x-axis; top panel), as well as the equilibrium fraction (solid line), estimated abundance fraction (dotted line) and forward projected proportion (dashed line) in the eastern Bering Sea relative to both eastern and northern Bering Sea (bottom panel)

northern stratum by 2019. We therefore conclude that including abundance-expanded movement among strata within a spatially stratified assessment model would capture some (but not all) of the northward shift that is apparent from survey and fishery data.

As a sensitivity analysis, we examine these same results for a model that sequentially drops each major data type (fishery catch-and-effort, survey or tagging data). This shows two major points. First, removing tagging data results in a much smaller estimate of diffusion rate (Figure 7, bottom-left panel). This then results in lower movement among strata, in turn causing a slower response of the *forward projection* metric to changing environmental conditions and a degraded match between that metric and the *abundance* metric (Figure 7, bottom-right panel). Second, removing survey data results in an implausible increase in the estimated density in the northern Bering Sea (Figure 7, 2nd row right column). By contrast, removing fishery data results in relatively little change in the qualitative results (Figure 7, top row), suggesting that fishery catch-and-effort data are less important than survey and tagging data for reconstructing movement for this population.

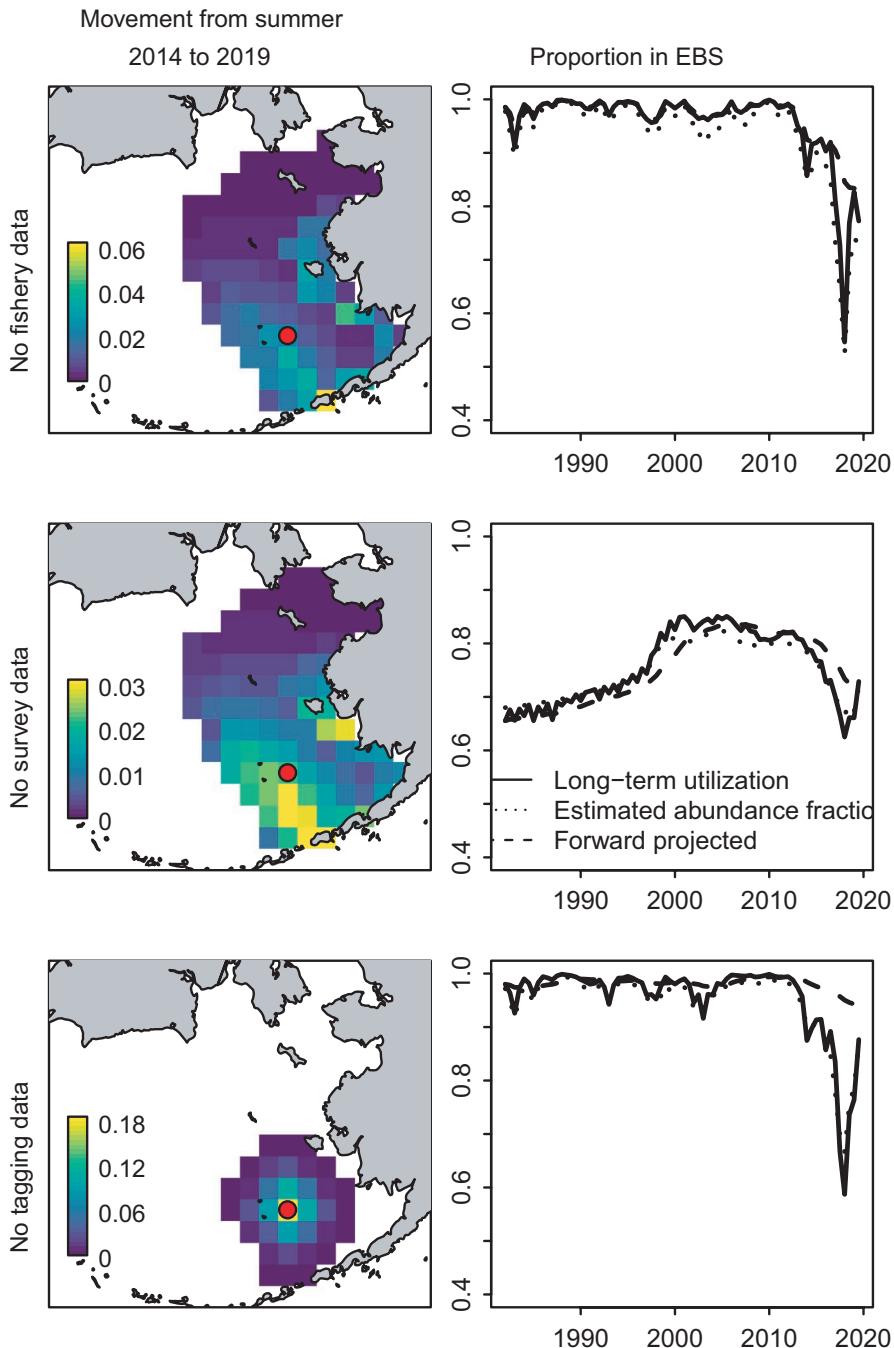
Finally, we demonstrate the impact of including estimates of annual movement fractions in the most recent spatially stratified stock

assessment model for Pacific cod (Figure 8). The existing “two-box” model (Figure 8 grey line) estimates a high proportion (10%–50%) of abundance in the northern Bering Sea throughout the entire period, despite a high-quality systematic survey in 2010 detecting almost no Pacific cod in that area and earlier surveys in the 1980s similarly recording low densities in the NBS (Figure 8 black bullets). This existing assessment model attributes the disagreement between predicted abundance and survey data to low catchability for cod in the northern Bering Sea prior to 2017. By contrast, the same model using movement fractions estimated here predicts that almost all abundance is in the eastern Bering Sea throughout 1990–2010 (Figure 8 blue line). It then predicts an increasing proportion after 2013 with patterns that match the “forward-projected” proportions shown previously (Figure 6). We therefore conclude (i) that movement fractions are easily included in the widely used stock assessment platform used in this assessment, and (ii) that this information can greatly impact predictions of population movement for this commercially important stock. However, the age-structured model using movement rates estimated here still substantially underestimates the proportion observed in the northern Bering Sea in 2017–2019 based on survey data (Figure 8 black bullets).

#### 4 | DISCUSSION

We have demonstrated a “hybrid” species distribution modelling framework (*sensu* Dormann et al., 2012) that has characteristics in common with both correlative SDMs (i.e. fitting to data statistically using covariates) and mechanistic methods (i.e. including individual movement processes). We used this hybrid SDM to integrate climate and habitat variables with data from conventional tags, fishery and survey operations to estimate fine-scale seasonal movement patterns for a marine species and also presented future extensions to integrate movement gates, archival tags, ecogegeochemical tracers and other emerging data types. We also demonstrated how to coarsen the spatial and seasonal resolution to calculate annualized movement fractions among large spatial strata that can be easily integrated within existing spatially stratified stock assessment models. Finally, we demonstrated the benefits of this approach for Bering Sea Pacific cod, a commercially important fish stock experiencing climate-driven northward range shifts. In the following, we next (i) discuss avenues to improve our ecological understanding of Pacific cod in particular, (ii) summarize improvements of this instantaneous diffusion–advection–taxis approach relative to previous methods integrating tag and survey data, (iii) review how the DAT model can be connected to IPM and stock assessment models and (iv) envision how this approach could be used to optimize ongoing deployment of tagging efforts.

We have used Pacific cod in the Bering Sea to demonstrate our integrated approach, but note that we have restricted environmental preferences to bathymetry and seasonal bottom temperature. We recommend that future research explores including additional covariates influencing seasonal and interannual variability in Pacific

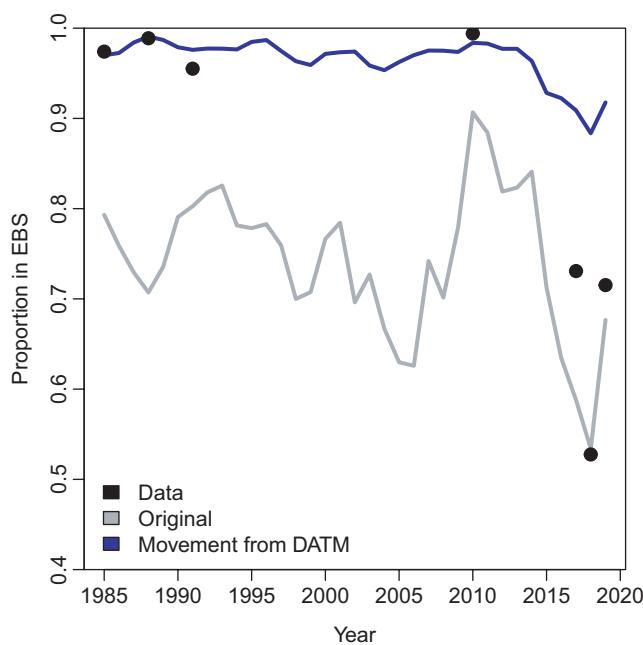


**FIGURE 7** Sensitivity analysis exploring the impact of excluding each major data type including fishery catch-and-effort (top row), survey data (middle row) or tagging data (bottom row), showing movement (see Figure 4 caption for legend details) from summer 2014 to summer 2019 (left column; see Figure 4 bottom-right panel for plot for base model) and proportion of Bering Sea abundance in the eastern Bering Sea in each year (right column; see Figure 6 bottom panel for plot for base model)

cod distribution. The processes that drive seasonal and interannual variation in distribution likely differ between feeding and spawning seasons, and the timing of spawning migration likely varies between years. Variation in the winter spawning distribution is influenced by the timing and extent of ice coverage and bottom temperatures (Neidetcher et al., 2014), while variation in spatial distribution outside of the spawning season is likely influenced by the variability in distribution of forage. The distribution of forage fishes, in turn, is associated with bottom depth and temperature, but also salinity, water column stability, concentration of chlorophyll-a, zooplankton biomass and composition, as well as annual indices of regional oceanography (Parker-Stetter et al., 2016). Future developments

of the DAT model could incorporate these regional oceanographic indices using a spatially varying response when estimating habitat preferences (Thorson, 2019b). The timing of Pacific cod spawning is also variable (Neidetcher et al., 2014), perhaps in response to the narrow window of temperatures resulting in optimal egg hatch success (Laurel & Rogers, 2020), and this could be resolved using finer seasonal increments (e.g. Thorson et al., 2020).

The DAT model builds upon existing advection-diffusion-reaction (ADR) analyses used in fisheries. For example, SEAPODYM estimates habitat preference separately for spawning and feeding stages based on tagging data, which are then used to drive advection within an advective-diffusive movement model (Lehodey et al.,



**FIGURE 8** Illustrating survey and stock assessment model estimates of the proportion of Bering Sea Pacific cod summary abundance (defined as abundance for ages  $\geq 1$ ) occurring in the eastern Bering Sea relative to the total summary abundance for northern and eastern Bering Sea strata (y-axis) in each year 1990–2019 (x-axis) with legend in bottom-left corner (black bullets: survey abundance converted to a proportion; grey line: original two-box age-structured stock assessment model estimate; blue line: assessment model estimate after including the movement fractions estimated by the DAT model). Note that the blue line deviates from the “forward-projected” estimates from the movement model (Figure 6 bottom panel) because the stock assessment model includes additional demographic structure (age-structure) and data types (age and length composition samples)

2008; Senina et al., 2019). Similarly, SPM integrates tagging and survey data to fit an age-structured population model, while projecting movement based on parameters representing movement probabilities and an estimated seasonal movement kernel (Dunn et al., 2015). However, we are not aware of other fisheries models that estimated advection-diffusion movement while also incorporating spatially correlated process errors and fitting to point-count data from fisheries and surveys. Identifying environmental associations using a hybrid SDM can be used directly within fisheries management, to identify long-term “essential fish habitat” and short-term bycatch hotspots (Maxwell et al., 2015; Rooper et al., 2020). Parameterizing hybrid SDMs is increasingly feasible given recent advances in biologging, as well as detailed habitat and oceanographic mapping efforts (Lowerre-Barbieri et al., 2019).

As demonstrated here, fine-scale predictions of movement can be coarsened to estimate seasonal or annualized movement fractions among larger spatial strata, and these movement fractions can be inputted as data into spatial IPMs or used to specify priors on movement parameters. These movement fractions are conditioned upon both (i) the estimated advective-diffusion movement process and (ii)

the fine-scale distribution of numerical abundance. Estimating fine-scale distribution is facilitated by fitting survey and fishery data, and spatially correlated process errors are useful to accurately incorporate interannual variation in distribution (Thorson, 2019a). In the future, we envision that the DAT model itself could be merged into an integrated stock assessment model to provide improved estimates of movement among coarsened strata, while enabling the incorporation of a wider array of data sources. Along these lines, we recommend future research to extend the DAT model to better account for stage based movement dynamics, such as migratory behaviour (i.e. spawning migrations); these could be modelled by estimating separate preference functions for different stages. We also recommend ongoing research to incorporate spatial variation in demographics (e.g. habitat-specific fishing mortality rates) within the DAT model.

Once a hybrid SDM has been developed and validated, analysts can then use the parameterized model as a simulation tool to optimize the location and timing of tag releases to maximize information gain with respect to an explicit management goal. Conventional tags are often deployed systematically with large sample sizes as a component of a standardized fisheries survey, randomly tagging fish proportionally to catch through space and time (e.g. Echave et al., 2013). This method can be useful for understanding general movement patterns when developing an initial hybrid SDM, but recapture rates are typically low (Shimada & Kimura, 1994) resulting in little statistical power for estimating processes affecting movement within small subareas or habitats of particular management interest. By contrast, newer archival and satellite tags remain expensive and typically are deployed in targeted campaigns without a population-scale probabilistic design. We recommend using a hybrid SDM as an operating model within a formal optimization exercise, using results (along with other operational and management constraints) to inform the deployment of targeted tagging efforts using additional electronic tags. We hypothesize that times and locations with historically less data or high variability in a fitted hybrid SDM could be a useful guidepost for future tagging efforts, but formal optimization may also yield surprises contrary to this pattern. For example, sampling optimization using occupancy models has recommended increasing sampling in locations with intermediate densities (rather than proportional to estimation variance) to better determine range limits (e.g. Reich et al., 2018).

In conclusion, we foresee a tremendous decrease in cost and increase in deployment for technologies tracking movement during the “Biologging decade.” We also see a growing need to integrate these novel movement-sampling technologies and programmes with conventional data types. This synthesis is needed, for example, to provide “best available science” advice for fisheries management and marine spatial planning, including the overlap between the changing footprint of human activities and changing seasonal and spatial habitat utilization. We therefore encourage researchers to develop a “community of practice” using diffusion-advection-taxis models to link individual movement to habitat and climate variables within hybrid SDMs. Defining annual movement from instantaneous habitat preference and diffusion parameters is necessary to

integrate data with large differences in temporal frequencies, for example daily records (archival tags, movement gates and selection experiments) and seasonal records (surveys and conventional tags). A common model then allows for improved sharing of information among sampling programmes as well as among species (based on similarities in traits and phylogeny) and promises to allow improved learning between marine regions and research communities worldwide (Thorson et al., 2021).

## ACKNOWLEDGMENTS

We thank many scientists for helpful discussions regarding this research, including L. Britt, S. McDermott, P. Conn and D. Johnson. JT is grateful to K. Kristensen and H. Skaug for their prior guidance regarding diffusion-taxis models and the developers of Template Model Builder for making these ambitious models possible. We thank many others for contributing to the data used here, including S. McDermott, D. Bryan and D. Nichol and others for the archival tagging efforts; M. Zimmermann, J. Pirtle, C. Rooper and others for the bathymetric layer; countless fishery observers and portside samplers for the longline fishery CPUE data; A. Hermann, W. Cheng and D. Pilcher for the ROMS bottom temperature layer; and B. Lauth, S. Kotwicki and others for the bottom trawl survey data. We thank C. Braun, P. Conn, E. Hazen, B. Gardner, M. Mangel and one anonymous reviewer for helpful comments on an earlier draft. This publication is partially funded by the Joint Institute for the Study of the Atmosphere and Ocean (JISAO) under NOAA Cooperative Agreement NA15OAR4320063, Contribution No. 2020-1135.

## DATA AVAILABILITY STATEMENT

Survey catches, tagging data and environmental layers (depth and seasonal bottom temperatures) are publicly available from the Alaska Fisheries Science Center and are reposed as an example in R package ATM (<https://github.com/James-Thorson-NOAA/ATM>). Fishery catch-per-unit-effort (CPUE) data are private information and are available from the Alaska Fisheries Science Center upon request and suitable data-sharing agreement. The aggregated fishery CPUE data used here are reposed in R package ATM, after excluding grid cells and intervals with three or fewer samples. The fishery effort data are private information and again are available from the Alaska Fisheries Science Center upon request and suitable data-sharing agreement. See Appendix S3 for detailed description of the data used.

## ORCID

- James T. Thorson  <https://orcid.org/0000-0001-7415-1010>
- Steven J. Barbeaux  <https://orcid.org/0000-0001-9498-0369>
- Daniel R. Goethel  <https://orcid.org/0000-0003-0066-431X>
- Kelly A. Kearney  <https://orcid.org/0000-0002-6152-5236>
- Edward A. Laman  <https://orcid.org/0000-0003-2663-9243>
- Julie K. Nielsen  <https://orcid.org/0000-0002-8384-4158>
- Matthew R. Siskey  <https://orcid.org/0000-0002-7834-4749>
- Kevin Siwickie  <https://orcid.org/0000-0003-4555-5272>
- Grant G. Thompson  <https://orcid.org/0000-0002-3240-6705>

## REFERENCES

- Berger, A. M., Goethel, D. R., Lynch, P. D., Quinn, T., Mormede, S., McKenzie, J., & Dunn, A. (2017). Space oddity: The mission for spatial integration. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(11), 1698–1716. <https://doi.org/10.1139/cjfas-2017-0150>
- Brillinger, D. R. (2012). Learning a potential function from a trajectory. In P. Guttorm & D. Brillinger (Eds.), *Selected works of david brillinger* (pp. 361–364). Springer. [https://doi.org/10.1007/978-1-4614-1344-8\\_20](https://doi.org/10.1007/978-1-4614-1344-8_20)
- Bruneel, S., Gobeyn, S., Verhelst, P., Reubens, J., Moens, T., & Goethals, P. (2018). Implications of movement for species distribution models - Rethinking environmental data tools. *Science of The Total Environment*, 628–629, 893–905. <https://doi.org/10.1016/j.scitenv.2018.02.026>
- Cao, J., Thorson, J. T., Punt, A. E., & Szuwalski, C. (2020). A novel spatio-temporal stock assessment framework to better address fine-scale species distributions: Development and simulation testing. *Fish and Fisheries*, 21(2), 350–367. <https://doi.org/10.1111/faf.12433>
- Dormann, C. F., Schymanski, S. J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin, X., Römermann, C., Schröder, B., & Singer, A. (2012). Correlation and process in species distribution models: Bridging a dichotomy. *Journal of Biogeography*, 39(12), 2119–2131. <https://doi.org/10.1111/j.1365-2699.2011.02659.x>
- Dunn, A., Rasumussen, S., & Mormede, S. (2015). *Spatial Population Model User Manual* (NIWA Technical Report 138; p. 208). National Institute of Water & Atmospheric Research Ltd.
- Echave, K., Hanselman, D. H., & Maloney, N. E. (2013). *Report to industry on the Alaska sablefish tag program, 1972-2012* (NOAA Tech. Memo. NMFS-AFSC-254; p. 47).
- Fournier, D. A., Hampton, J., & Sibert, J. R. (1998). MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. *Canadian Journal of Fisheries and Aquatic Sciences*, 55(9), 2105–2116. <https://doi.org/10.1139/f98-100>
- Fox, J., & Weisberg, S. (2018). *An R companion to applied regression*, (2nd ed.). Sage Publications.
- Goethel, D. R., Bosley, K. M., Langseth, B. J., Deroba, J. J., Berger, A. M., Hanselman, D. H., & Schueler, A. M. (2021). Where do you think you're going? Accounting for ontogenetic and climate-induced movement in spatially stratified integrated population assessment models. *Fish and Fisheries*, 22(1), 141–160. <https://doi.org/10.1111/faf.12510>
- Hampton, J., & Fournier, D. A. (2001). A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Marine and Freshwater Research*, 52(7), 937–963. <https://doi.org/10.1071/mf01049>
- Hanks, E. M., Hooten, M. B., & Alldredge, M. W. (2015). Continuous-time discrete-space models for animal movement. *The Annals of Applied Statistics*, 9(1), 145–165. <https://doi.org/10.1214/14-AOAS803>
- Hanselman, D. H., Heifetz, J., Echave, K. B., & Dressel, S. C. (2015). Move it or lose it: Movement and mortality of sablefish tagged in Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(2), 238–251. <https://doi.org/10.1139/cjfas-2014-0251>
- Hauser, D. D. W., Laidre, K. L., & Stern, H. L. (2018). Vulnerability of Arctic marine mammals to vessel traffic in the increasingly ice-free Northwest Passage and Northern Sea Route. *Proceedings of the National Academy of Sciences*, 115(29), 7617–7622. <https://doi.org/10.1073/pnas.1803543115>
- Hays, G. C., Bailey, H., Bograd, S. J., Bowen, W. D., Campagna, C., Carmichael, R. H., Casale, P., Chiaradia, A., Costa, D. P., Cuevas, E., Nico de Bruyn, P. J., Dias, M. P., Duarte, C. M., Dunn, D. C., Dutton, P. H., Esteban, N., Friedlaender, A., Goetz, K. T., Godley, B. J., ... Sequeira, A. M. M. (2019). Translating Marine Animal Tracking Data into Conservation Policy and Management. *Trends in Ecology & Evolution*, 34(5), 459–473. <https://doi.org/10.1016/j.tree.2019.01.009>

- Heifetz, J., & Fujioka, J. T. (1991). Movement dynamics of tagged sablefish in the northeastern Pacific. *Fisheries Research*, 11(3), 355–374. [https://doi.org/10.1016/0165-7836\(91\)90009-5](https://doi.org/10.1016/0165-7836(91)90009-5)
- Hermann, A. J., Gibson, G. A., Bond, N. A., Curchitser, E. N., Hedstrom, K., Cheng, W., Wang, M., Cokelet, E. D., Stabeno, P. J., & Aydin, K. (2016). Projected future biophysical states of the Bering Sea. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 134, 30–47. <https://doi.org/10.1016/j.dsr2.2015.11.001>
- Johnson, D. S., London, J. M., Lea, M.-A., & Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89(5), 1208–1215. <https://doi.org/10.1890/07-1032.1>
- Kass, R. E., & Steffey, D. (1989). Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical bayes models). *Journal of the American Statistical Association*, 84(407), 717–726. <https://doi.org/10.2307/2289653>
- Kearney, K., Hermann, A., Cheng, W., Ortiz, I., & Aydin, K. (2020). A coupled pelagic-benthic-sympagic biogeochemical model for the Bering Sea: Documentation and validation of the BESTNPZ model (v2019.08.23) within a high-resolution regional ocean model. *Geoscientific Model Development*, 13(2), 597–650. <https://doi.org/10.5194/gmd-13-597-2020>
- Kéry, M., & Schaub, M. (2021). *Integrated population models: Theory and ecological applications with R and JAGS* (1st ed.). Academic Press.
- Kristensen, K., Nielsen, A., Berg, C. W., Skaug, H., & Bell, B. M. (2016). TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software*, 70(5), 1–21. <https://doi.org/10.18637/jss.v070.i05>
- Kristensen, K., Thygesen, U. H., Andersen, K. H., & Beyer, J. E. (2014). Estimating spatio-temporal dynamics of size-structured populations. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(2), 326–336. <https://doi.org/10.1139/cjfas-2013-0151>
- Lam, C. H., Nielsen, A., & Sibert, J. R. (2008). Improving light and temperature based geolocation by unscented Kalman filtering. *Fisheries Research*, 91(1), 15–25. <https://doi.org/10.1016/j.fishres.2007.11.002>
- Laurel, B. J., & Rogers, L. A. (2020). Loss of spawning habitat and pre-recruits of Pacific cod during a Gulf of Alaska heatwave. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(4), 644–650. <https://doi.org/10.1139/cjfas-2019-0238>
- Laurel, B. J., Stoner, A. W., & Hurst, T. P. (2007). Density-dependent habitat selection in marine flatfish: The dynamic role of ontogeny and temperature. *Marine Ecology Progress Series*, 338, 183–192.
- Lee, M. S. Y., & Palci, A. (2015). Morphological phylogenetics in the genomic age. *Current Biology*, 25(19), R922–R929. <https://doi.org/10.1016/j.cub.2015.07.009>
- Lehodey, P., Senina, I., & Murtugudde, R. (2008). A spatial ecosystem and populations dynamics model (SEAPODYM) – Modeling of tuna and tuna-like populations. *Progress in Oceanography*, 78(4), 304–318. <https://doi.org/10.1016/j.pocean.2008.06.004>
- Lindgren. (2012). Continuous domain spatial models in R-INLA. *The ISBA Bulletin*, 19(4), 14–20.
- Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4), 423–498. <https://doi.org/10.1111/j.1467-9868.2011.00777.x>
- Lowerre-Barbieri, S. K., Kays, R., Thorson, J. T., & Wikelski, M. (2019). The ocean's movescape: fisheries management in the bio-logging decade (2018–2028). *ICES Journal of Marine Science*, 76(2), 477–488. <https://doi.org/10.1093/icesjms/fsy211>
- Maunder, M. N., & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, 142, 61–74. <https://doi.org/10.1016/j.fishres.2012.07.025>
- Maureaud, A. A., Frelat, R., Pécuchet, L., Shackell, N., Mérigot, B., Pinsky, M. L., Amador, K., Anderson, S. C., Arkhipkin, A., Auber, A., Barri, I., Bell, R. J., Belmaker, J., Beukhof, E., Camara, M. L., Guevara-Carrasco, R., Choi, J., Christensen, H. T., Conner, J., ... Thorson, J. T. (2021). Are we ready to track climate-driven shifts in marine species across international boundaries? - A global survey of scientific bottom trawl data. *Global Change Biology*, 27(2), 220–236. <https://doi.org/10.1111/gcb.15404>
- Maxwell, S. M., Hazen, E. L., Lewison, R. L., Dunn, D. C., Bailey, H., Bograd, S. J., Briscoe, D. K., Fossette, S., Hobday, A. J., Bennett, M., Benson, S., Caldwell, M. R., Costa, D. P., Dewar, H., Eguchi, T., Hazen, L., Kohin, S., Sippel, T., & Crowder, L. B. (2015). Dynamic ocean management: Defining and conceptualizing real-time management of the ocean. *Marine Policy*, 58, 42–50. <https://doi.org/10.1016/j.marpol.2015.03.014>
- Method, R. D., & Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, 142, 86–99.
- Michelot, T., Gloaguen, P., Blackwell, P. G., & Étienne, M.-P. (2019). The Langevin diffusion as a continuous-time model of animal movement and habitat selection. *Methods in Ecology and Evolution*, 10(11), 1894–1907. <https://doi.org/10.1111/2041-210X.13275>
- Mormede, S., Dunn, A., Hanchet, S. M., & Parker, S. (2014). Spatially explicit population dynamics models for Antarctic toothfish in the Ross Sea region. *CCAMLR Science*, 21, 19–37.
- Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., & Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences of the United States of America*, 105(49), 19052–19059. <https://doi.org/10.1073/pnas.0800375105>
- Neidetcher, S. K., Hurst, T. P., Ciannelli, L., & Logerwell, E. A. (2014). Spawning phenology and geography of Aleutian Islands and eastern Bering Sea Pacific cod (*Gadus macrocephalus*). *Deep Sea Research Part II: Topical Studies in Oceanography*, 109, 204–214. <https://doi.org/10.1016/j.dsr2.2013.12.006>
- Nichol, D. G., Honkalehto, T., & Thompson, G. G. (2007). Proximity of Pacific cod to the sea floor: Using archival tags to estimate fish availability to research bottom trawls. *Fisheries Research*, 86(2), 129–135. <https://doi.org/10.1016/j.fishres.2007.05.009>
- O'Leary, C. A., Thorson, J. T., Ianelli, J. N., & Kotwicki, S. (2020). Adapting to climate-driven distribution shifts using model-based indices and age composition from multiple surveys in the walleye pollock (*Gadus chalcogrammus*) stock assessment. *Fisheries Oceanography*, 29(6), 541–557. <https://doi.org/10.1111/fog.12494>
- O'Leary, C. A., Thorson, J. T., Miller, T. J., & Nye, J. A. (2020). Comparison of multiple approaches to calculate time-varying biological reference points in climate-linked population-dynamics models. *ICES Journal of Marine Science*, 77(3), 930–941. <https://doi.org/10.1093/icesjms/fsz215>
- Parker-Stetter, S., Urmy, S., Horne, J., Eisner, L., & Farley, E. (2016). Factors affecting summer distributions of Bering Sea forage fish species: Assessing competing hypotheses. *Deep Sea Research Part II: Topical Studies in Oceanography*, 134, 255–269. <https://doi.org/10.1016/j.dsr2.2016.06.013>
- Pebesma, E. (2019). stars: Spatiotemporal Arrays, Raster and Vector Data Cubes. *R Package Version 0.4-0*.
- Pinsky, M. L., Worm, B., Fogarty, M. J., Sarmiento, J. L., & Levin, S. A. (2013). Marine taxa track local climate velocities. *Science*, 341(6151), 1239–1242.
- Preisler, H. K., Ager, A. A., & Wisdom, M. J. (2013). Analyzing animal movement patterns using potential functions. *Ecosphere*, 4(3), art32. <https://doi.org/10.1890/ES12-00286.1>
- R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Reich, B. J., Pacifici, K., & Stallings, J. W. (2018). Integrating auxiliary data in optimal spatial design for species distribution

- modelling. *Methods in Ecology and Evolution*, 9(6), 1626–1637. <https://doi.org/10.1111/2041-210X.13002>
- Rooper, C. N., Ortiz, I., Hermann, A. J., Laman, N., Cheng, W., Kearney, K., & Aydin, K. (2020). Predicted shifts of groundfish distribution in the Eastern Bering Sea under climate change, with implications for fish populations and fisheries management. *ICES Journal of Marine Science*, 78(1), 220–234. <https://doi.org/10.1093/icesjms/fsaa215>
- Santora, J. A., Mantua, N. J., Schroeder, I. D., Field, J. C., Hazen, E. L., Bograd, S. J., Sydeman, W. J., Wells, B. K., Calambokidis, J., Saez, L., Lawson, D., & Forney, K. A. (2020). Habitat compression and ecosystem shifts as potential links between marine heatwave and record whale entanglements. *Nature Communications*, 11(1), 536. <https://doi.org/10.1038/s41467-019-14215-w>
- Senina, I., Lehodey, P., Sibert, J., & Hampton, J. (2019). Integrating tagging and fisheries data into a spatial population dynamics model to improve its predictive skills. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(3), 576–593. <https://doi.org/10.1139/cjfas-2018-0470>
- Shertzer, K. W., & Bacheler, N. M. (2020). Estimating population abundance at a site in the open ocean: combining information from conventional and telemetry tags with application to gray triggerfish (*Balistes capriscus*) - Canadian Journal of Fisheries and Aquatic Sciences. *Canadian Journal of Fisheries and Aquatic Sciences*, 77(1), 34–43.
- Shimada, A. M., & Kimura, D. K. (1994). Seasonal movements of Pacific cod, *Gadus macrocephalus*, in the eastern Bering Sea and adjacent waters based on tag-recapture data. *Fish Bull*, 92, 800–816.
- Sibert, J. R., Hampton, J., Fournier, D. A., & 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*). *Canadian Journal of Fisheries and Aquatic Sciences*, 56(6), 925–938. <https://doi.org/10.1139/f99-017>
- Smith, J. Q. (1985). Diagnostic checks of non-standard time series models. *Journal of Forecasting*, 4(3), 283–291. <https://doi.org/10.1002/for.3980040305>
- Spies, I., Gruenthal, K. M., Drinan, D. P., Hollowed, A. B., Stevenson, D. E., Tarpey, C. M., & Hauser, L. (2020). Genetic evidence of a northward range expansion in the eastern Bering Sea stock of Pacific cod. *Evolutionary Applications*, 13(2), 362–375. <https://doi.org/10.1111/eva.12874>
- Thompson, G., Conner, J., Shotwell, K., Fissel, B., Hurst, T., Laurel, B., Rogers, L., & Siddon, E. (2020). Assessment of the Pacific cod stock in the Eastern Bering Sea. In *Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea and Aleutian Islands [NPFMC Bering Sea and Aleutian Islands SAFE]*. North Pacific Fishery Management Council. Retrieved from <https://archive.afsc.noaa.gov/refm/docs/2019/EBSpcod.pdf>
- Thorson, J. T. (2019a). Forecast skill for predicting distribution shifts: A retrospective experiment for marine fishes in the Eastern Bering Sea. *Fish and Fisheries*, 20(1), 159–173. <https://doi.org/10.1111/faf.12330>
- Thorson, J. T. (2019b). Measuring the impact of oceanographic indices on species distribution shifts: The spatially varying effect of cold-pool extent in the eastern Bering Sea. *Limnology and Oceanography*, 64(6), 2632–2645. <https://doi.org/10.1002/lo.11238>
- Thorson, J. T. (2020). Predicting recruitment density dependence and intrinsic growth rate for all fishes worldwide using a data-integrated life-history model. *Fish and Fisheries*, 21(2), 237–251. <https://doi.org/10.1111/faf.12427>
- Thorson, J. T., Adams, C. F., Brooks, E. N., Eisner, L. B., Kimmel, D. G., Legault, C. M., Rogers, L. A., & Yasumiishi, E. M. (2020). Seasonal and interannual variation in spatio-temporal models for index standardization and phenology studies. *ICES Journal of Marine Science*, 77(5), 1879–1892. <https://doi.org/10.1093/icesjms/fsaa074>
- Thorson, J. T., & Barnett, L. A. K. (2017). Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science*, 74(5), 1311–1321. <https://doi.org/10.1093/icesjms/fsw193>
- Thorson, J. T., Hermann, A. J., Siwicke, K., & Zimmermann, M. (2021). Grand challenge for habitat science: stage-structured responses, nonlocal drivers, and mechanistic associations among habitat variables affecting fishery productivity. *ICES Journal of Marine Science*, fsaa236. <https://doi.org/10.1093/icesjms/fsaa236>
- Thorson, J. T., Shelton, A. O., Ward, E. J., & Skaug, H. J. (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science: Journal Du Conseil*, 72(5), 1297–1310. <https://doi.org/10.1093/icesjms/fsu243>
- Webster, R. A., Clark, W. G., Leaman, B. M., & Forsberg, J. E. (2013). Pacific halibut on the move: A renewed understanding of adult migration from a coastwide tagging study. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(4), 642–653. <https://doi.org/10.1139/cjfas-2012-0371>
- Wikle, C. K. (2003). Hierarchical bayesian models for predicting the spread of ecological processes. *Ecology*, 84(6), 1382–1394. [https://doi.org/10.1890/0012-9658\(2003\)084%5B1382:HBMFP%5D2.0.CO;2](https://doi.org/10.1890/0012-9658(2003)084%5B1382:HBMFP%5D2.0.CO;2)
- Wilson, K., Hanks, E., & Johnson, D. (2018). Estimating animal utilization densities using continuous-time Markov chain models. *Methods in Ecology and Evolution*, 9(5), 1232–1240. <https://doi.org/10.1111/2041-210X.12967>
- Zimmermann, M., & Benson, J. L. (2013). *Smooth sheets: How to work with them in a GIS to derive bathymetry, features and substrates* (NOAA Technical Memorandum NMFS-AFSC-249; p. 52). US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Alaska Fisheries Science Center.
- Zimmermann, M., & Prescott, M. M. (2018). Bathymetry and canyons of the eastern Bering Sea slope. *Geosciences*, 8(5), 184. <https://doi.org/10.3390/geosciences8050184>
- Zipkin, E. F., Rossman, S., Yackulic, C. B., Wiens, J. D., Thorson, J. T., Davis, R. J., & Grant, E. H. C. (2017). Integrating count and detection-nondetection data to model population dynamics. *Ecology*, 98(6), 1640–1650. <https://doi.org/10.1002/ecy.1831>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Thorson, J. T., Barbeaux, S. J., Goethel, D. R., Kearney, K. A., Laman, E. A., Nielsen, J. K., Sisley, M. R., Siwicke, K., & Thompson, G. G. (2021). Estimating fine-scale movement rates and habitat preferences using multiple data sources. *Fish and Fisheries*, 00, 1–18. <https://doi.org/10.1111/faf.12592>