

Bridging the gaps in animal movement: hidden behaviors and ecological relationships revealed by integrated data streams

BRETT T. MCCLINTOCK,[†] JOSHUA M. LONDON, MICHAEL F. CAMERON, AND PETER L. BOVENG

Marine Mammal Laboratory, Alaska Fisheries Science Center, NOAA-NMFS, 7600 Sand Point Way NE, Seattle, Washington 98115 USA

Citation: McClintock, B. T., J. M. London, M. F. Cameron, and P. L. Boveng. 2017. Bridging the gaps in animal movement: hidden behaviors and ecological relationships revealed by integrated data streams. *Ecosphere* 8(3):e01751. 10.1002/ecs2.1751

Abstract. Inferences about animal behavior from movement models typically rely solely on location data, but auxiliary biotelemetry and environmental data are powerful and readily available resources for incorporating much more behavioral realism. Integrating multiple data streams can not only reveal hidden behaviors and ecological relationships that would otherwise be difficult or impossible to infer from location data alone, but also facilitate more realistic path reconstruction that respects important ecological features while bridging the information gaps that commonly arise due to measurement error or missing data. Using the bearded seal (*Erignathus barbatus*), a benthic predator associated with Arctic sea ice, we demonstrate how integrating location, dive activity, land cover, bathymetry, and sea ice data in a unified modeling framework allowed us to identify novel behavior states, such as hauling out on seasonal sea ice and those associated with competing foraging strategies (i.e., benthic vs. mid-water prey). By utilizing multiple data streams, ecologists can move beyond conventional two-state models (“foraging” and “transit”) and address more interesting hypotheses about activity budgets, resource selection, and many other areas of movement and behavioral ecology. The generality of our approach provides broad applicability to marine and terrestrial species, as well as many types of biotelemetry and environmental data.

Key words: animal location data; bearded seal; hidden Markov; switching behavior; telemetry.

Received 3 February 2017; **accepted** 8 February 2017. Corresponding Editor: Debra P. C. Peters.

Copyright: © 2017 McClintock et al. This article is a U.S. Government work and is in the public domain in the USA. *Ecosphere* published by Wiley Periodicals, Inc., on behalf of the Ecological Society of America. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

[†] **E-mail:** brett.mcclintock@noaa.gov

INTRODUCTION

Although animal movement is an incredibly complex process influenced by many factors as individuals respond to numerous stimuli, inferences about animal behavior from movement models still overwhelmingly rely on the analysis of animal location data alone (e.g., Blackwell 2003, Morales et al. 2004, Jonsen et al. 2005, Gurarie et al. 2009, McClintock et al. 2012, Fleming et al. 2015). When inference is based solely on location data, however, it can be very difficult to distinguish important movement behaviors. For example, location data from periods of relative inactivity (e.g., resting) can appear identical to

those from periods of intense area-restricted search (e.g., foraging), and inference about such behaviors based solely on horizontal trajectory can be misleading (McClintock et al. 2013, Bestley et al. 2016).

Modern biotelemetry devices provide much more than simple information about animal location. They can now be equipped with various “bio-logging” sensors that record detailed information about animal behavior, physiology, and the surrounding environment, such as temperature, heart rate, dive profiles, altitude, accelerometry, and proximity to conspecifics, competitors, or predators. Publicly available repositories of remotely sensed data (e.g., bathymetry, habitat type) also provide easy access to information about

environments animals inhabit. These auxiliary biotelemetry and environmental data are powerful and readily available resources for incorporating much more ecological realism into animal movement models.

Integrating animal location, biotelemetry, and environmental data can reveal hidden behaviors and ecological relationships that would otherwise be difficult or impossible to infer from location data alone. For example, the bearded seal (*Erignathus barbatus*) is known as a benthic forager whose life history in the Arctic is intricately linked with sea ice, but the nature of that relationship and other ecological connections with habitat is poorly understood. Given the inherent difficulty of direct observation, biologists rely on satellite telemetry as the primary source of behavioral data for this species. However, key components of bearded seal behavioral ecology, such as benthic foraging or association with sea ice, are not recorded by current satellite transmitters. Rather, locations obtained during surface breathing intervals or haul-out periods are typically recorded, often with substantial measurement error. Basic two-state movement models can often be fit to such location data, and these typically include a “foraging” state characterized by area-restricted search-type movement and a “transit” state characterized by high speed and directional persistence (Morales et al. 2004, Jonsen et al. 2005). Whether applied to marine or terrestrial species, the typical “foraging” state in a two-state model would include both foraging and non-foraging “resting” behavior. Marine satellite tags often record additional biotelemetry information such as pressure sensor (i.e., dive activity) and conductivity (i.e., wet/dry) data, and while these data can be integrated to help distinguish resting from foraging (e.g., McClintock et al. 2013, 2015, Joy et al. 2015, Russell et al. 2015), these data sources cannot be used for inferences about different *types* of resting or foraging behavior. The incorporation of readily accessible environmental data (e.g., sea ice cover, land cover, sea floor depth) can help distinguish these otherwise hidden behavioral states and thereby facilitate greater ecological insight than can conventional generic state assignments (e.g., “resting,” “foraging,” “transit”).

When explicitly accounted for in the movement process model, biotelemetry and environmental

data can also facilitate more realistic reconstruction of the latent movement path from location data that are temporally irregular, infrequent, or subject to measurement error. Obvious examples are barriers to movement such as land and large bodies of water for marine and terrestrial animals, respectively, but such important ecological barriers are typically ignored in analyses of animal movement (e.g., Jonsen et al. 2005, Johnson et al. 2008, but see Tremblay et al. 2009, McClintock et al. 2012, Brost et al. 2015). Path reconstruction can also be improved by incorporating biotelemetry data. For example, dive depth (or altitude) or external temperature can be used to constrain movements to only those areas with compatible elevations or temperatures, respectively (e.g., Tremblay et al. 2009).

While some recent animal movement studies have begun to utilize both location and additional sources of biotelemetry or environmental data to identify additional behavioral states (e.g., McClintock et al. 2013, Isojunno and Miller 2015, Joy et al. 2015, Russell et al. 2015) or as covariates in state transition probabilities (e.g., Morales et al. 2004, Bestley et al. 2013), inference remains limited to only a few movement behaviors. This is largely because it is difficult to identify >2 biologically meaningful movement behaviors from relatively few data streams (e.g., Morales et al. 2004, McClintock et al. 2014), particularly in the presence of measurement error (e.g., Breed et al. 2013, Silva et al. 2014). Other approaches have successfully linked animal movement to environmental covariates (e.g., Forester et al. 2007, Hanks et al. 2011, Joy et al. 2015), but these do not explicitly integrate distinct behavioral state relationships with the environment and state-switching mechanisms into the movement process.

Here, we develop a discrete-time, multistate movement modeling framework for inferring many movement behavior states and associated activity budgets from animal location, biotelemetry, and environmental data. By incorporating all data streams into a single movement model, our approach facilitates the identification of novel and ecologically meaningful behaviors that otherwise could not be identified solely from any particular data stream. The combined information from all data streams can also help deal with common challenges, such as location measurement error, missing data, and environmental boundaries, in reconstructing the movement

path. After developing the model in the next section, we provide an example using telemetry data collected from bearded seals tagged in coastal waters of Alaska, United States. Climate change and other factors have prompted studies on the influence of benthic communities and environmental characteristics on bearded seal benthic foraging ecology and sea ice use, but in order to study these relationships, one must first be able to identify when and where these activities occur from the limited data available. We demonstrate how this can be accomplished by integrating location, biotelemetry, and environmental data in a unified inferential framework.

METHODS

As a benthic predator that relies on sea ice, the bearded seal exhibits many behaviors that are not directly observable or classifiable from standard satellite telemetry data. These behaviors include hauling out on ice, resting at sea, hauling out on land, mid-water foraging (e.g., for pelagic prey), benthic foraging, and transit (e.g., migration). Some of these behaviors, such as resting at sea, hauling out on land, or hauling out on ice, could potentially be identified post hoc after fitting a model that incorporates both location and dive activity data (Johnson et al. 2008, Russell et al. 2015), but this is problematic with locations that are temporally irregular, infrequent, or subject to measurement error because the movement model does not incorporate these important environmental features (e.g., land, sea ice) when recreating the movement path. Other behaviors, such as mid-water and benthic foraging, are more difficult to reliably infer post hoc regardless of the regularity, frequency, or accuracy of observed locations. The challenge in identifying latent behaviors such as benthic foraging lies not only in estimating where the animal was, but also whether the animal was diving to the sea floor. This is but one of many examples where the explicit integration of environmental data with location and biotelemetry data can aid not only in identifying hidden movement behavior states, but also in the recreation of the movement path.

Bearded seal data

Between 2009 and 2012, we deployed Argos (Service Argos 2013) satellite-linked tags on $N = 7$

bearded seals near Kotzebue, Alaska, United States (Fig. 1). In addition to location acquisition, the tags were equipped with pressure sensors for recording dive information and conductivity sensors for recording wet/dry information. Location acquisition and data transmission were limited to surface intervals (i.e., during haul-out periods or surface breathing intervals). To cope with limited transmission periods and bandwidth, the dive data were summarized onboard and transmitted as histograms of “dive depth,” “dive duration,” and “time at depth” over 6-h intervals, where each 6-h histogram was divided into 14 pre determined depth (m) or dive duration (min) bins (see London 2016 for further details). Hourly percent dry times were also transmitted. From these dive activity data, we calculated the proportion of time spent diving >4 m below the surface, the proportion of dry time, and the number of dives at 6-h intervals for use as data streams in our movement process model. These 6-h time steps, chosen to roughly correspond to “day,” “night,” and two “crepuscular” periods, are common in longer tag deployments for diving marine mammals due to practical limitations associated with battery life and data transmission.

In addition to location and biotelemetry data, we utilized several environmental data sources that are readily available from public repositories to both help identify movement behavior states and recreate the movement path. These included land cover data for the region defined by the Bering, Chukchi, and Beaufort seas, sea ice concentration data, and bathymetry data (see Appendix S1). Given the nature of the Argos platform, animal behavior leading to limited or irregular exposure to the satellites, and limited bandwidth for transferring data, the location data were subject to measurement error and the dive activity data were subject to missing or incomplete records. A larger proportion of data were missing for seals tagged during 2009 because extra (GPS location) data collected only in that year consumed a portion of the data uplinks to the Argos satellites. We therefore require a model that accounts for these vagaries.

Movement process model

Building on the discrete-time, population-level hierarchical model of McClintock et al. (2013), we developed a six-state movement behavior model for bearded seals, where movement

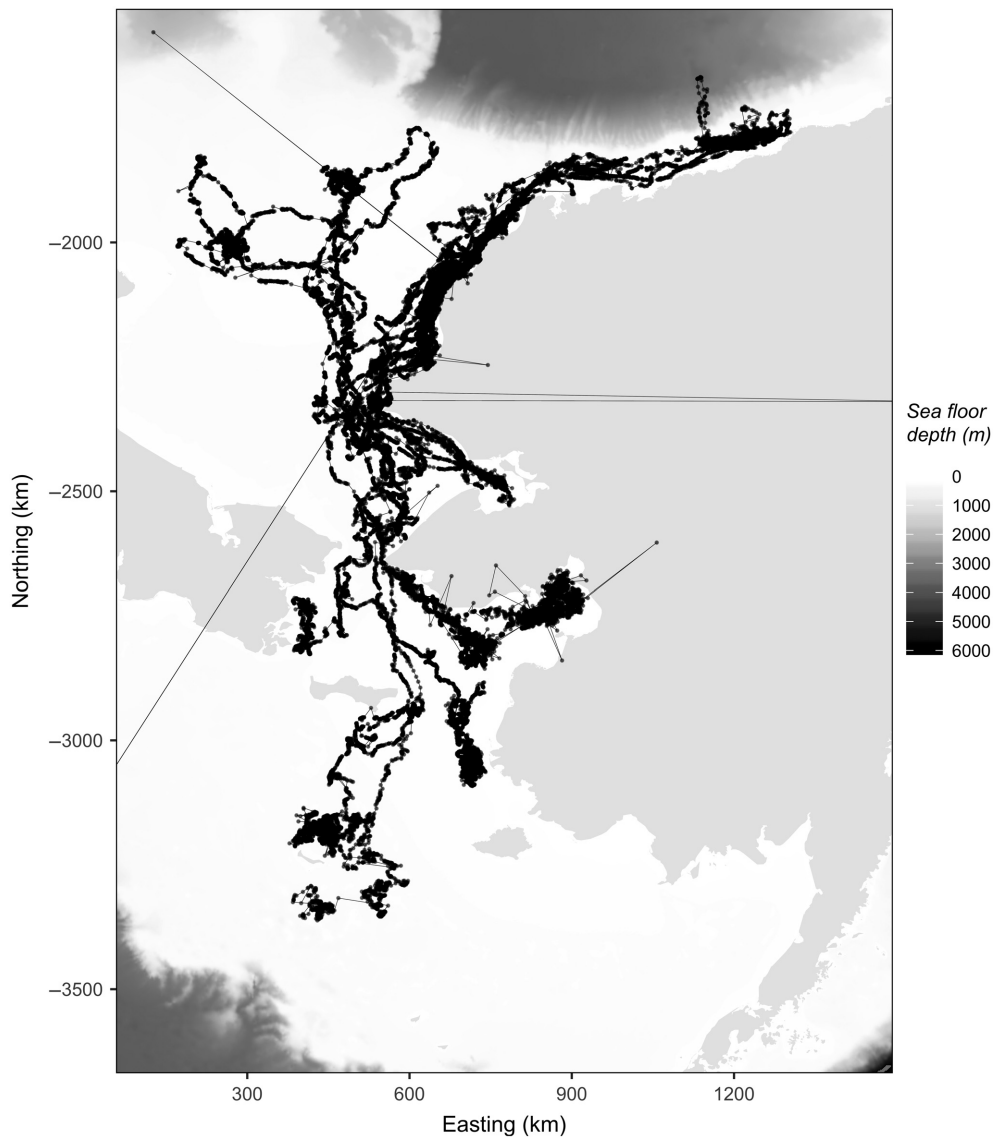


Fig. 1. Observed locations for seven bearded seals obtained from Argos satellite telemetry tags deployed between 2009 and 2012 near Kotzebue, Alaska, United States. The 53,069 locations (black dots) were subject to measurement error and were only attainable during surface intervals. Darker shading offshore indicates deeper sea floor depths (in meters).

behavior states and associated movement parameters were estimated from seven data streams. These data streams included step length ($s_{n,t}$), bearing ($\phi_{n,t}$), the proportion of time spent diving >4 m below the surface ($w_{n,t}$), the proportion of dry time ($d_{n,t}$), the number of dives to the sea floor (i.e., “benthic dives”; $e_{n,t}$), the average proportion of sea ice cover ($c_{n,t}$), and the average proportion of land cover ($l_{n,t}$) for each 6-h time

step $t = 1, \dots, T_n$ and individual $n = 1, \dots, N$. Our goal was to identify and estimate activity budgets to six distinct movement behavior states, $z_{n,t} \in \{I, S, L, M, B, T\}$, where I denotes “hauled out on ice,” S denotes “resting at sea,” L denotes “hauled out on land,” M denotes “mid-water foraging,” B denotes “benthic foraging,” and T denotes “transit,” based on the combined information across all data streams. As a heuristic

example of how the movement process model works, suppose a particular 6-h time step exhibited a short step length, no time spent diving below 4 m, 100% dry time, and no dives to the sea floor; if sea ice cover was >0% and land cover was 0%, one could reasonably expect the animal was hauled out on ice during this time step (state *I*; Table 1).

For horizontal movement, we assumed step length $s_{n,t}|z_{n,t} = i \sim \text{Gamma}(a_{n,i}/b_{n,i}, b_{n,i})$ with state-specific mean step length parameter $a_{n,z} > 0$ and shape parameter $b_{n,z} > 0$ for $z \in \{I, S, L, M, B, T\}$. For bearing, we assumed $\phi_{n,t}|z_{n,t} = i \sim \text{wCauchy}(\phi_{n,t-1}, r_{n,i})$, which is a wrapped Cauchy distribution with state-specific directional persistence parameter $-1 < r_{n,z} < 1$. Based on bearded seal movement behavior, we expect average step length to be smaller for resting (states *I*, *S*, and *L*) and larger for transit. We also expect directional persistence to be largest for transit. As in McClintock et al. (2013), these expected relationships were reflected in prior constraints on the state-dependent parameters (see Table 1; Appendix S1 for full details).

Although movement behavior state assignment could be based solely on horizontal movement characteristics (e.g., Morales et al. 2004, Jonsen et al. 2005, McClintock et al. 2012), we wished to incorporate the additional information about behavior states provided by biotelemetry (i.e., dive activity) and environmental (i.e., bathymetry, land cover, and sea ice concentration) data. Assuming independence between data streams (but still conditional on state), we incorporated $w_{n,t}$, $d_{n,t}$, $e_{n,t}$, $c_{n,t}$, and $l_{n,t}$ into a joint

conditional likelihood whereby each data stream contributes its own state-dependent component. While for simplicity we assume independence of data streams conditional on state, data streams such as proportion of dive ($w_{n,t}$) and dry ($d_{n,t}$) time could potentially be more realistically modeled using multivariate distributions that account for additional (state-dependent) correlations.

For the data streams corresponding to proportions ($w_{n,t}$, $d_{n,t}$, $c_{n,t}$, $l_{n,t}$), we assumed

$$q_{n,t}|\mathbf{v}^q, \delta^q, z_{n,t} = i \sim \text{Beta}(v_i^q, \delta_i^q),$$

where v_i^q and δ_i^q are (state-dependent) shape parameters for $q \in \{w, d, c, l\}$ and $i \in \{I, S, L, M, B, T\}$. For the number of benthic dives ($e_{n,t}$), we assumed

$$e_{n,t}|\lambda, z_{n,t} = i \sim \text{Poisson}(\lambda_i),$$

where λ_i is a (state-dependent) rate or intensity parameter. As with step length and bearing, we imposed prior constraints on these distributions based on the expected relationships of the states and the data streams (see Table 1; Appendix S1).

Although critical for identifying benthic foraging activity, $e_{n,t}$ was not directly observable because the exact locations and depths of the seals during each 6-h time step were unknown. We therefore calculated the number of benthic foraging dives, defined as the number of dives to depth bins with endpoints that included the sea floor, based on the sea floor depths at the estimated start and end locations for each time step. Similarly, $c_{n,t}$ and $l_{n,t}$ were calculated based on the average of the sea ice concentration and land

Table 1. Expected characteristics of six movement behavior states for a bearded seal movement model incorporating seven data streams.

Behavioral state	Horizontal trajectory		Dive	Dry	Benthic	Ice	Land
	Step length	Directional persistence					
"Hauled out on ice" (<i>I</i>)	Shorter		Lower	Higher	Lower	Higher	Lower
"Resting at sea" (<i>S</i>)	Shorter		Lower	Lower	Lower	Lower	Lower
"Hauled out on land" (<i>L</i>)	Shorter		Lower	Higher	Lower	Lower	Higher
"Mid-water foraging" (<i>M</i>)			Higher	Lower	Lower		
"Benthic foraging" (<i>B</i>)			Higher	Lower	Higher		
"Transit" (<i>T</i>)	Longer	Higher	Higher	Lower	Lower		

Notes: These data streams included horizontal trajectory ("step length" and "directional persistence"), the proportion of time spent diving below 4 m ("dive"), the proportion of time spent dry ("dry"), and the number of benthic dives ("benthic") during each 6-h time step. The model incorporated environmental data on the proportion of sea ice and land cover in 25×25 km grid cell(s) containing the start and end locations for each time step ("ice" and "land"), as well as bathymetry data to identify benthic dives. Blank entries indicate no a priori relationships were assumed in the model.

cover values, respectively, for the start and end locations. We estimated start and end locations for each time step by combining our movement process model with an observation process model similar to Jonsen et al. (2005) extended for the Argos error ellipse (McClintock et al. 2015), but, importantly, we also imposed constraints on the predicted locations by prohibiting movements inland and to areas where the sea floor depth was shallower than the maximum observed dive depth for each time step (see *Observation process model*).

To model transitions between behavior states, we assigned a first-order Markov categorical distribution, $z_{n,t}|\psi, z_{n,t-1} = k \sim \text{Categorical}(\psi_{k,I}, \psi_{k,S}, \psi_{k,L}, \psi_{k,M}, \psi_{k,B}, \psi_{k,T})$, for $k = \{I, S, L, M, B, T\}$, where $\psi_{k,j}$ is the transition probability from state k at time $t - 1$ to state j at time t , and $\sum_j \psi_{k,j} = 1$. Thus, the conditional likelihood for our multi-state, discrete-time movement process model integrating location, biotelemetry, and environmental data is

$$f(s, \phi, w, d, e, c, l, z|\theta) = \prod_{n=1}^N \prod_{t=1}^{T_n} f(s_{n,t}|\theta, z_{n,t}) \times f(\phi_{n,t}|\theta, z_{n,t}) \times f(w_{n,t}|\theta, z_{n,t}) \times f(d_{n,t}|\theta, z_{n,t}) \times f(e_{n,t}|\theta, z_{n,t}) \times f(c_{n,t}|\theta, z_{n,t}) \times f(l_{n,t}|\theta, z_{n,t}) \times f(z_{n,t}|\theta, z_{n,t-1})$$

where $f()$ denotes a (conditional) probability density function and θ is the set of all parameters.

Observation process model

For the observation process model, the data consist of the observed locations $(x_{n,t,i}, y_{n,t,i})$ for individual $n = 1, \dots, 7$, time step $t = 1, \dots, T_n$ and observation $i = 1, \dots, k_{n,t}$ (where time steps with $k_{n,t} = 0$ have no observed locations). Similar to Jonsen et al. (2005) and McClintock et al. (2012, 2013), we assumed that individuals travel in a straight line between times $t - 1$ and t . The temporally irregular observed $(x_{n,t,i}, y_{n,t,i})$ and true $(\mu_{x_{n,t,i}}, \mu_{y_{n,t,i}})$ locations were then related to the temporally regular locations $(X_{n,t}, Y_{n,t})$ via:

$$\mu_{x_{n,t,i}} = (1 - j_{n,t,i})X_{n,t-1} + j_{n,t,i}X_{n,t} \\ \mu_{y_{n,t,i}} = (1 - j_{n,t,i})Y_{n,t-1} + j_{n,t,i}Y_{n,t}$$

where $j_{n,t,i} \in (0,1]$ is the proportion of the time interval between locations $(X_{n,t-1}, Y_{n,t-1})$ and $(X_{n,t}, Y_{n,t})$ at which the i th observation was

obtained, and $(X_{n,0}, Y_{n,0})$ is an initial (latent) location for individual n .

Following McClintock et al. (2015), we account for error resulting from location measurement explained by the Argos error ellipse and that attributable to outliers or discretization of the movement path into temporally regular steps using a bivariate t -distribution with γ_1 degrees of freedom:

$$f\left(x_{n,t,i}, y_{n,t,i}, \Sigma_{n,t,i}|\theta, X_{n,0}, Y_{n,0}, \phi_{n,[1:t]}, s_{n,[1:t]}\right) \\ = \frac{\Gamma\left(\frac{\gamma_1+2}{2}\right)}{\gamma_1 \pi \left(\frac{\gamma_1}{2}\right) |\Sigma_{n,t,i}|^{\frac{1}{2}}} \left[1 + \frac{1}{\gamma_1} \left(\frac{x_{n,t,i} - \mu_{x_{n,t,i}}}{y_{n,t,i} - \mu_{y_{n,t,i}}}\right)^T \right. \\ \left. \times \Sigma_{n,t,i}^{-1} \begin{pmatrix} x_{n,t,i} - \mu_{x_{n,t,i}} \\ y_{n,t,i} - \mu_{y_{n,t,i}} \end{pmatrix} \right]^{-\frac{(\gamma_1+2)}{2}}$$

where

$$\Sigma_{n,t,i} = \begin{pmatrix} \sigma_{x_{n,t,i}}^2 & \sigma_{xy_{n,t,i}} \\ \sigma_{xy_{n,t,i}} & \sigma_{y_{n,t,i}}^2 \end{pmatrix},$$

$\phi_{n,[1:t]} = (\phi_{n,1}, \phi_{n,2}, \dots, \phi_{n,t})$, $s_{n,[1:t]} = (s_{n,1}, s_{n,2}, \dots, s_{n,t})$, and θ again denotes the set of all model parameters. Note that $\Sigma_{n,t,i}$ is derived from the components of the error ellipse (provided by Argos) for each location. For the GPS-quality locations obtained from the tags deployed in 2009, we used independent scaled t -distributions:

$$f(x_{n,t,i}|\theta, X_{n,0}, Y_{n,0}, \phi_{n,[1:t]}, s_{n,[1:t]}) \\ = \frac{\Gamma\left(\frac{\gamma_2+1}{2}\right)}{\sqrt{\gamma_2 \pi \sigma_1^2} \Gamma\left(\frac{\gamma_2}{2}\right)} \left(1 + \frac{1}{\gamma_2} \left(\frac{(x_{n,t,i} - \mu_{x_{n,t,i}})^2}{\sigma_1^2}\right)\right)^{-\frac{\gamma_2+1}{2}} \\ f(y_{n,t,i}|\theta, X_{n,0}, Y_{n,0}, \phi_{n,[1:t]}, s_{n,[1:t]}) \\ = \frac{\Gamma\left(\frac{\gamma_2+1}{2}\right)}{\sqrt{\gamma_2 \pi \sigma_2^2} \Gamma\left(\frac{\gamma_2}{2}\right)} \left(1 + \frac{1}{\gamma_2} \left(\frac{(y_{n,t,i} - \mu_{y_{n,t,i}})^2}{\sigma_2^2}\right)\right)^{-\frac{\gamma_2+1}{2}}$$

with degrees of freedom term γ_2 and scale parameters σ_1^2 and σ_2^2 .

In order to maintain meaningful ecological boundaries and consistency among the temporally regular locations $(X_{n,t}, Y_{n,t})$, dive data, and environmental variables (sea floor depth, sea ice concentration, and land cover), predicted locations were constrained to those with depths no

shallower than the deepest dive bin for each time step and with <100% land cover (thus forbidding inland movements). Our observation process model was therefore

$$f(\mathbf{x}, \mathbf{y}, \boldsymbol{\Sigma} | \boldsymbol{\theta}, X_0, Y_0, \boldsymbol{\phi}, \mathbf{s}) = \prod_{n=1}^N \prod_{t=1}^{T_n} I((X_{n,t}, Y_{n,t}) \in D_{n,t}) I((X_{n,t}, Y_{n,t}) \in C) \times \left[\prod_{i=1}^{k_{n,t}} f(x_{n,t,i}, y_{n,t,i}, \boldsymbol{\Sigma}_{n,t,i} | \boldsymbol{\theta}, X_{n,0}, Y_{n,0}, \boldsymbol{\phi}_{n,[1:t]}, \mathbf{s}_{n,[1:t]}) \times (1 - I((x_{n,t,i}, y_{n,t,i}) \in G)) + f(x_{n,t,i}, y_{n,t,i} | \boldsymbol{\theta}, X_{n,0}, Y_{n,0}, \boldsymbol{\phi}_{n,[1:t]}, \mathbf{s}_{n,[1:t]}) \times I((x_{n,t,i}, y_{n,t,i}) \in G) \right]$$

where $I((X_{n,t}, Y_{n,t}) \in D_{n,t})$ and $I((X_{n,t}, Y_{n,t}) \in C)$ are indicators for whether location $(X_{n,t}, Y_{n,t}) = (X_{n,t-1} + s_{n,t} \cos(\phi_{n,t}), Y_{n,t-1} + s_{n,t} \sin(\phi_{n,t}))$ is in the permitted depth ($D_{n,t}$) and land regions (C) for time step t , and $I((x_{n,t,i}, y_{n,t,i}) \in G)$ is an indicator for whether observed location $(x_{n,t,i}, y_{n,t,i})$ is in the set of GPS-quality locations (G).

Model fitting and diagnostics

We fit our Bayesian state-space model, consisting of both movement and observation process components, using Markov chain Monte Carlo (MCMC) methods. We assigned uninformative priors whenever there were no a priori expectations about the relationships of the model

parameters and the movement behavior states (see Appendix S1). Missing or incomplete dive activity data ($w_{n,t}$, $d_{n,t}$, or $e_{n,t}$) were assumed to occur at random and imputed (conditional on state $z_{n,t}$) within each iteration of the MCMC algorithm (see Appendix S2). We initialized six chains from random starting values drawn from the prior distributions. After initial pilot tuning and burn-in, we ran the chains until the univariate Gelman–Rubin potential scale reduction factor (PSRF) for monitored parameters was <1.2 and effective sample sizes exceeded 4000. This required about 6 million iterations per chain, where 1 million iterations for each parallel chain required about 100 h using six cores of a 3.7-GHz processor. Data and code can be found in Data S1.

RESULTS

Based on posterior modes for state assignments, we found these bearded seals spent 70% of their deployment periods in the benthic foraging state (B). They hauled out on ice (state I) and rested at sea (state S) in roughly equal proportions, but there was a clear seasonal component in activity budgets among the hauled out on ice (I), benthic foraging (B), and transit (T) states that coincided with the summer retreat and winter advance of sea ice in the Arctic (Table 2). All seals exhibited intensive benthic foraging along the continental shelves of the Chukchi or Beaufort seas as sea ice receded northward in the summer. We found very little evidence of these

Table 2. Bearded seal activity budgets among six behavioral states (I = “hauled out on ice,” S = “resting at sea,” L = “hauled out on land,” M = “mid-water foraging,” B = “benthic foraging,” and T = “transit”) from a discrete-time, multistate movement model integrating animal location, biotelemetry, and environmental data.

State	Summer			Autumn			Winter			Overall		
	Time	LCI	UCI	Time	LCI	UCI	Time	LCI	UCI	Time	LCI	UCI
I	0.01	0.01	0.01	0.06	0.05	0.06	0.19	0.17	0.20	0.07	0.07	0.08
S	0.06	0.06	0.06	0.06	0.05	0.06	0.07	0.06	0.08	0.06	0.06	0.07
L	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.00	0.00	0.01
M	0.04	0.04	0.04	0.04	0.03	0.04	0.05	0.04	0.06	0.04	0.04	0.04
B	0.80	0.80	0.81	0.70	0.69	0.71	0.57	0.56	0.58	0.70	0.70	0.71
T	0.09	0.08	0.09	0.14	0.14	0.15	0.12	0.11	0.13	0.12	0.11	0.12

Notes: There were seasonal differences in activity budgets between “summer” (from tagging in late June and early July to 30 September), “autumn” (1 October to 31 December), and “winter” (1 January until tag loss between February and April) that coincided with the southern advance of winter sea ice in the Arctic. Activity budget summaries were calculated based on the posterior mode of time spent in each state with lower (LCI) and upper (UCI) credible intervals based on 95% highest posterior density. Individual seal activity budgets are included in Appendix S4.

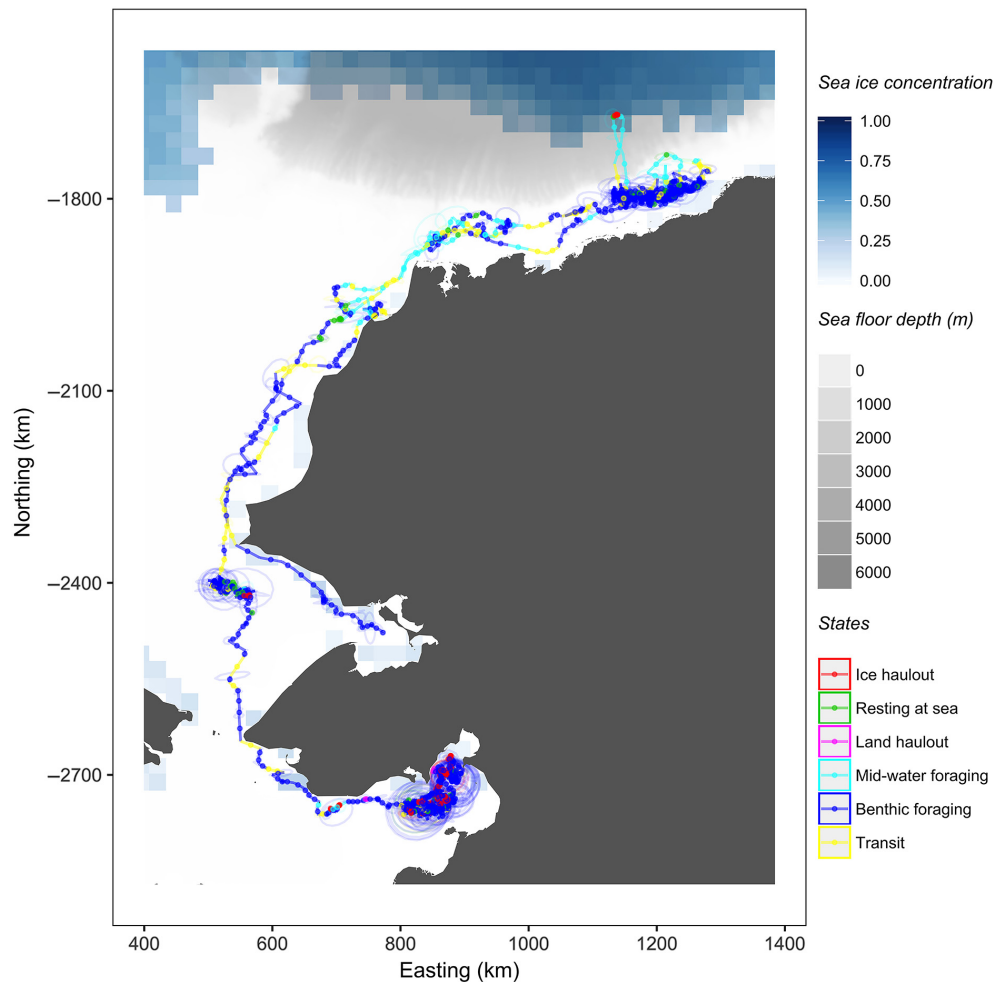


Fig. 2. Predicted locations and state assignments for an adult male bearded seal tag deployment from 25 June 2009 to 9 March 2010. The proportion of sea ice concentration cover in 25×25 km grid cells on 29 July 2009 indicates our model identified a mid-water foraging trip to the deeper Canada Basin waters off the Beaufort Shelf during which the animal hauled out on the northwardly receding sea ice edge. Uncertainty in predicted locations is indicated by 95% normal error ellipses.

ice-associated seals hauling out on land, even when near the coast during the largely ice-free period in the summer. Interestingly, an adult male briefly moved north of the Beaufort coastal shelf in late July 2009 and was predicted to have hauled out on the receding sea ice edge between bouts of mid-water foraging in the deeper waters of the Canada Basin (Fig. 2). We provide a detailed animation of predicted movements and state assignments for all seals relative to bathymetry and sea ice cover in Appendix S3.

As expected, the resting states (*I*, *S*, and *L*) exhibited shorter step lengths, no directional

persistence, and lower proportions of each time step diving >4 m below the surface based on state-specific parameter estimates (Fig. 3; see Appendix S4 for posterior summaries of all monitored parameters). Both the mid-water foraging (*M*) and transit (*T*) states exhibited longer step lengths, lower proportions of dry time, and higher proportions of diving >4 m below the surface. Only the transit state exhibited strong directional persistence, but state *M* exhibited moderate directional persistence for some individuals (Fig. 3). The benthic foraging (*B*) state generally exhibited moderate step lengths, no

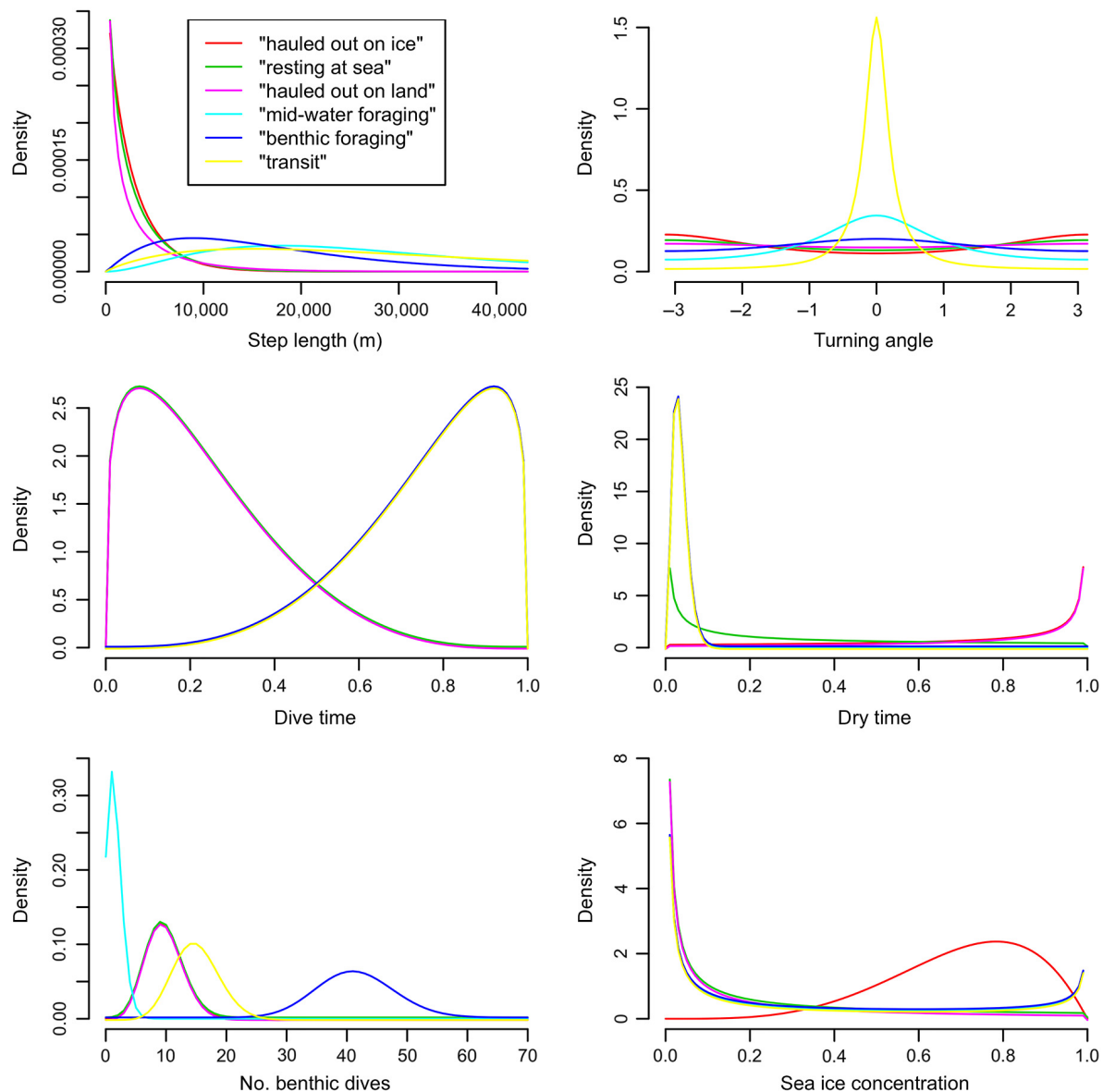


Fig. 3. State-dependent probability densities for the step length, turning angle, dive time, dry time, number of benthic dive, and sea ice concentration data streams. Estimates are based on posterior medians of population-level parameters for $N = 7$ bearded seals.

directional persistence, lower proportions of dry time, higher proportions of diving >4 m below the surface, and considerably higher benthic dive counts (posterior median for $\lambda_B = 41.4$; 95% highest posterior density interval: 41.2–41.7). The resting ($\lambda_I = \lambda_S = \lambda_L = 9.7$; 95% HPDI: 9.4–10.0) and transit ($\lambda_T = 15.0$; 95% HPDI: 14.5–15.5) states exhibited moderate benthic dive counts, while the mid-water foraging state generally had

the lowest benthic dive counts ($\lambda_M = 1.5$; 95% HPDI: 1.3–1.7; Fig. 3). The moderate benthic dive counts for the resting and transit states could be attributable to mid-interval state switches, but there certainly could have been opportunistic benthic foraging during transit. There also could have been high-speed, directionally persistent benthic dives while transiting in relatively shallow shelf waters.

The parameters associated with state L exhibited poor mixing among chains, but this was due to the paucity of time steps assigned to the L state and therefore of little consequence. For a single seal, we also found somewhat poorer mixing for a handful of the step length and bearing distribution parameters $(r_{4,I}, a_{4,M}, a_{4,B}, a_{4,T}, b_{4,I}, b_{4,B})$ that exhibited PSRFs between 1.14 and 1.27. For all other parameters, the point and upper confidence limit estimates of the PSRFs never exceeded 1.08 (median = 1.003) and 1.19 (median = 1.009), respectively (see Appendix S4).

DISCUSSION

We have demonstrated how animal location, biotelemetry, and environmental data can be integrated within a unified framework for inferring more ecologically meaningful behaviors and realistic movement paths. By integrating additional sources of readily available data, ecologists can move beyond the limited inferences provided by conventional two-state movement models (e.g., “foraging” and “transit” sensu Morales et al. 2004, Jonsen et al. 2005). It can also facilitate more realistic path reconstruction by respecting important features that influence movement such as land, water, altitude, and habitat type. The bearded seal served as a model species for our approach because of its associations with sea ice and foraging behaviors that are difficult to observe directly, but our approach has broad applicability to many marine and terrestrial species, as well as to different types of biotelemetry and environmental data. For example, North American elk are primarily grazers but are also known to browse when grass-like plants are unavailable (Christianson and Creel 2007); our approach could be used to identify these different foraging behaviors while linking them to specific habitat types and other environmental factors (e.g., snow cover). Including biotelemetry data from head or mandible accelerometers as an additional data stream could prove particularly useful for identifying different foraging behaviors.

The ultimate objective of our bearded seal analysis was to better understand the influence of benthic communities and environmental characteristics on bearded seal benthic foraging ecology in the face of climate change and other

factors (M. F. Cameron, B. T. McClintock, and A. L. Blanchard, *unpublished manuscript*). Using the output from our model, M. F. Cameron, B. T. McClintock, and A. L. Blanchard, (*unpublished manuscript*) found that bearded seals avoided fine grain sand and selected benthic foraging habitat associated with numerous factors, including higher relative densities of fishes such as small sculpins and snailfishes, as well as certain bivalves and crustaceans. While we are confident in our ability to identify benthic foraging activity for this purpose, we are less confident in our ability to distinguish “mid-water foraging” (state M) from “transit” (state T) using the available data streams. While the animals certainly could have been foraging in mid-water, we are reluctant to conclude time steps assigned to state M actually constituted mid-water foraging and not a slightly slower, less directionally persistent transit state with fewer benthic dives. However, only 4% of all time steps were assigned to state M (see Table 2), and we were less concerned with this as our goal was primarily to distinguish benthic foraging from other behaviors for a subsequent investigation of bearded seal benthic resource selection. In order to better distinguish mid-water foraging from transit, we would likely need to incorporate additional data streams (e.g., dive profile or accelerometer) or prior information. Furthermore, our foraging states are not indicative of success; additional data streams would also be required to identify successful foraging activity (e.g., stomach temperature).

When measurement error and missing data are negligible, our framework can be easily implemented using multivariate hidden Markov model (HMM) fitting techniques (e.g., Bagniewska et al. 2013, Phillips et al. 2015, Zucchini et al. 2016). However, a drawback of our approach is the need to rely on computationally demanding techniques for model fitting (e.g., MCMC) when accounting for measurement error and/or missing data as was necessary in our bearded seal application. Multiple imputation provides a potentially more computationally efficient approach to model fitting while accounting for location measurement error and missing data (Rubin 1987, Hooten et al. 2017). For example, one could repeatedly draw movement path realizations from a single-state movement model that accommodates measurement error and is relatively easy to fit (e.g., Johnson

et al. 2008), and then apply our multistate movement model to each track realization (potentially imputing any missing biotelemetry or environmental data) using the relatively fast and efficient HMM algorithm. This multiple imputation approach remains an avenue of ongoing research.

Our inclusion of six behavior states in the bearded seal example was based on expert knowledge of bearded seal biology, our ecological questions of interest, and the available data. However, selecting the most appropriate number of states in (finite-state) HMMs can be challenging. Standard model selection techniques based on classical information criteria (e.g., Akaike Information Criterion, Bayesian Information Criterion) or Bayesian approaches (e.g., Bayes factors, posterior model probabilities) can tend to favor models with more states than is typically desirable or biologically interpretable. Pohle et al. (2017) suggest a practical approach to addressing this issue that emphasizes study objectives and expert knowledge in addition to standard model selection and model checking procedures.

We used a discrete-time movement model because this made it straightforward to reconcile multiple data streams collected at different time scales with varying degrees of “missingness.” This of course comes at the expense of discretization error, a feature that is absent from continuous-time models (McClintock et al. 2014). Path reconstruction and state assignment could therefore potentially be further improved by extending our approach to continuous time. As demonstrated in our bearded seal example, activity budgets can often be influenced by season and environment. Other promising extensions therefore include modeling state transition probabilities as a hidden semi-Markov process (Zucchini et al. 2016) or as a function of biotelemetry or environmental data.

ACKNOWLEDGMENTS

Bearded seal tagging was funded by the Bureau of Ocean Energy Management Alaska Environmental Studies Program through Interagency Agreement M07RG13317 with the U.S. Department of Commerce, NOAA/NMFS, and conducted under NMFS marine mammal research permits 782-1765-00 and 15126 authorized under IACUC Number A/NW 2010-3. The data were collected in partnership with seal experts

from the Alaska Native villages of Kotzebue, Wainwright, and Utkiagvik (Barrow). The findings and conclusions herein are those of the author(s) and do not necessarily represent the views of NOAA/NMFS.

LITERATURE CITED

- Bagniewska, J. M., T. Hart, L. A. Harrington, and D. W. Macdonald. 2013. Hidden Markov analysis describes dive patterns in semiaquatic animals. *Behavioral Ecology* 24:659–667.
- Bestley, S., I. D. Jonsen, R. G. Harcourt, M. A. Hindell, and N. J. Gales. 2016. Putting the behavior into animal movement modeling: improved activity budgets from use of ancillary tag information. *Ecology and Evolution* 6:8243–8255.
- Bestley, S., I. D. Jonsen, M. A. Hindell, C. Guinet, and J.-B. Charrassin. 2013. Integrative modelling of animal movement: incorporating in situ habitat and behavioural information for a migratory marine predator. *Proceedings of the Royal Society B: Biological Sciences* 280:20122262.
- Blackwell, P. G. 2003. Bayesian inference for Markov processes with diffusion and discrete components. *Biometrika* 90:613–627.
- Breed, G. A., W. D. Bowen, and M. L. Leonard. 2013. Behavioral signature of intraspecific competition and density dependence in colony-breeding marine predators. *Ecology and Evolution* 3:3838–3854.
- Brost, B. M., M. Hooten, E. Hanks, and R. Small. 2015. Animal movement constraints improve resource selection inference in the presence of telemetry error. *Ecology* 96:2590–2597.
- Christianson, D. A., and S. Creel. 2007. Review of environmental factors affecting elk winter diets. *Journal of Wildlife Management* 71:164–176.
- Fleming, C. H., W. F. Fagan, T. Mueller, K. A. Olson, P. Leimgruber, and J. M. Calabrese. 2015. Estimating where and how animals travel: an optimal framework for path reconstruction from autocorrelated tracking data. *Ecology*. <https://doi.org/10.1890/15-1607>
- Forester, J. D., A. Ives, M. Turner, D. Anderson, D. Fortin, H. Beyer, D. Smith, and M. Boyce. 2007. State-space models link elk movement patterns to landscape characteristics in Yellowstone National Park. *Ecological Monographs* 77:285–299.
- Gurarie, E., R. D. Andrews, and K. L. Laidre. 2009. A novel method for identifying behavioural changes in animal movement data. *Ecology Letters* 12: 395–408.
- Hanks, E. M., M. Hooten, D. Johnson, and J. Sterling. 2011. Velocity-based movement modeling for individual and population level inference. *PLoS One* 6:e22795.

- Hooten, M. B., D. S. Johnson, B. T. McClintock, and J. M. Morales. 2017. Animal movement: statistical models for telemetry data. CRC Press, Boca Raton, Florida, USA.
- Isojunno, S., and P. J. Miller. 2015. Sperm whale response to tag boat presence: Biologically informed hidden state models quantify lost feeding opportunities. *Ecosphere* 6:6.
- Johnson, D. S., J. M. London, M.-A. Lea, and J. W. Durban. 2008. Continuous-time correlated random walk model for animal telemetry data. *Ecology* 89: 1208–1215.
- Jonsen, I. D., J. M. Flemming, and R. A. Myers. 2005. Robust state-space modeling of animal movement data. *Ecology* 86:2874–2880.
- Joy, R., M. Dowd, B. Battaile, P. Lestenkof, J. Sterling, A. Trites, and R. Routledge. 2015. Linking northern fur seal dive behavior to environmental variables in the eastern Bering Sea. *Ecosphere* 6:75.
- London, J. M. 2016. kotzeb0912: v1.0 [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.57290>
- McClintock, B. T., D. S. Johnson, M. B. Hooten, J. M. Ver Hoef, and J. M. Morales. 2014. When to be discrete: the importance of time formulation in understanding animal movement. *Movement Ecology* 2:21.
- McClintock, B. T., R. King, L. Thomas, J. Matthiopoulos, B. J. McConnell, and J. M. Morales. 2012. A general discrete-time modeling framework for animal movement using multi-state random walks. *Ecological Monographs* 82:335–349.
- McClintock, B. T., J. M. London, M. F. Cameron, and P. L. Boveng. 2015. Modelling animal movement using the Argos satellite telemetry location error ellipse. *Methods in Ecology and Evolution* 6:266–277.
- McClintock, B. T., D. J. Russell, J. Matthiopoulos, and R. King. 2013. Combining individual animal movement and ancillary biotelemetry data to investigate population-level activity budgets. *Ecology* 94: 838–849.
- Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell. 2004. Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* 85:2436–2445.
- Phillips, J. S., T. A. Patterson, B. Leroy, G. M. Pilling, and S. J. Nicol. 2015. Objective classification of latent behavioral states in bio-logging data using multivariate-normal hidden Markov models. *Ecological Applications* 25:1244–1258.
- Pohle, J., R. Langrock, van Beest F. M., and N. M. Schmidt. 2017. Selecting the number of states in hidden Markov models: pitfalls, practical challenges, and pragmatic solutions. arXiv preprint: arXiv:1701.08673.
- Rubin, D. 1987. Multiple imputation for nonresponse in surveys. Wiley, New York, New York, USA.
- Russell, D. J., B. McClintock, J. Matthiopoulos, P. Thompson, D. Thompson, P. Hammond, E. Jones, M. MacKenzie, S. Moss, and B. McConnell. 2015. Intrinsic and extrinsic drivers of activity budgets in sympatric grey and harbour seals. *Oikos* 124: 1462–1472.
- Service Argos. 2013. Argos User's Manual. CLS. <http://www.argos-system.org/manual>
- Silva, M. A., I. Jonsen, D. Russell, R. Prieto, D. Thompson, and M. Baumgartner. 2014. Assessing performance of Bayesian state-space models fit to Argos satellite telemetry locations processed with Kalman filtering. *PLoS One* 9:e92277.
- Tremblay, Y., P. W. Robinson, and D. P. Costa. 2009. A parsimonious approach to modeling animal movement data. *PLoS One* 4:e4711.
- Zucchini, W., I. MacDonald, and R. Langrock. 2016. Hidden Markov models for time series: an introduction using R. Second edition. CRC Press, Boca Raton, Florida, USA.

SUPPORTING INFORMATION

Additional Supporting Information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1751/full>