A Continuous-Time Semi-Markov Model for Animal Movement in a Dynamic Environment

Devin S. Johnson¹, Noel A. Pelland, and Jeremy T. Sterling

Alaska Fisheries Science Center, NOAA Fisheries,
Seattle, Washington, U.S.A.

June 21, 2018

¹Email: devin.johnson@noaa.gov

Abstract

1 We consider an extension to discrete-space continuous-time models animal
2 movement that have previously be presented in the literature. The exten-
3 sion from a continuous-time Markov formulation to a continuous-time *semi*-
4 Markov formulation allows for the inclusion of temporally dynamic habitat
5 conditions as well as temporally changing movement responses by animals
6 to that environment. We show that with only a little additional considera-
7 tion, the Poisson likelihood approximation for the Markov version can still be
8 used within the multiple imputation framework commonly employed for anal-
9 ysis of telemetry data. In addition, we consider a Bayesian model selection
10 methodology with the imputation framework. The model selection method
11 uses a Laplace approximation to the posterior model probability to provide
12 a computationally feasible approach. The full methodology is then used to
13 analyze movements of 15 northern fur seal (*Callorhinus ursinus*) pups with
14 respect to surface winds, geostrophic currents, and sea surface temperature.
15 The highest posterior model probabilities belonged to those models contain-
16 ing only winds and current, SST did not seem to be a significant factor for
17 modeling their movement.

18 **Key words:** Animal movement, semi-Markov model, dynamic habitat, model
19 selection, multiple imputation, posterior model probability, resource selection

20 1 Introduction

21 Studying the movement of animals has become ubiquitous in the ecological
22 literature in the past decade. One traditionally thorny issue for developing
23 statistical models for telemetry data is relating spatial habitat conditions
24 to movement processes. This difficulty arises because of the mismatch in
25 support for the telemetry data (continuous spatial domain) and the observed
26 habitat variables (usually a spatial grid or raster) and the often random times
27 when location is observed. As the habitat variables are the lowest spatial
28 resolution of the two types, thus, it makes sense to bring the telemetry data
29 to the level of the habitat data, so-to-speak. We propose a generalization of

the spatially-discrete, continuous-time movement of Hanks et al. (2015) that can relate movement between cells to a dynamically changing environment evaluated at the cell level.

In recent years, there have been several methodologies based on translating telemetry data to a discrete version and modeling movement between cells based on temporally static habitat variables (see Hooten et al. 2010, Hanks et al. 2015, Hanks and Hughes 2016, and Avgar et al. 2016). In addition, Johnson et al. (2013) illustrated that for static environments spatial point process models could also be used by collapsing over the temporal index of the telemetry data.

The continuous-time Markov chain (CTMC) movement model of Hanks et al. (2015) is especially appealing due to the fact that the continuous-time formulation means that the model interpretation is not subject to the choice of time discretization. In addition, the CTMC model can be fit with GLM or GAM software for efficient parameter estimation. The CTMC model has been used to analyze mountain lion (*Puma concolor*) and northern fur seal movements (*Callorhinus ursinus*) (e.g., Buderman et al. 2018; Hanks and Hughes 2016; Hanks et al. 2015). However, there are some limitations in the CTMC model that prohibit use of temporally indexed covariates. We propose a generalization of the CTMC to a continuous-time *semi*-Markov chain (CTSMC) movement model for which habitat can vary in time.

The inspiration for this research is the migratory movements of northern fur seal pups in relation to space and time-dependent physical environmental conditions, namely, ocean surface currents, marine surface winds, and sea surface temperature (SST). Each of these habitat variables are derived from satellite remote sensing or atmospheric reanalysis products. These variables were chosen based on previous evidence of their influence on migratory movements or physiological processes in this or other similar pinniped species. Ream et al. (2005) found evidence in satellite-tagged adult female northern fur seals of migratory movement alignment with mesoscale oceanographic surface currents. Anecdotal evidence and traditional knowledge of the influence of marine winds on the migratory trajectories of northern fur seal pups extends back millennia, as documented in testimonials of Aleut hunters to US

63 government officials during study of the Bering Sea northern fur seal popula-
 64 tion in the late 1800s (Hooper, 1895). These hunters were in agreement that
 65 seals always traveled with a fair wind and sea, that the prevailing patterns of
 66 these variables influenced the passes through which migratory animals exited
 67 the Bering Sea, and that the timing of fall storms could influence the onset
 68 of migration. Widespread satellite tagging has provided the opportunity to
 69 confirm and quantify these effects; Lea et al. (2009) found differing departure
 70 and dispersal patterns of Pribilof Islands northern fur seal pups in years of
 71 contrasting atmospheric forcing, and evidence of more rapid movement in the
 72 presence of a tailwind. Sterling et al. (2014) and Pelland et al. (2014) found
 73 statistical evidence that surface winds influence adult northern fur seals dur-
 74 ing migration, with high winds associated with linear, directed movement,
 75 though directionality in relation to the wind was not explored.

76 While no direct evidence of the influence of SST on migratory movement
 77 exists for this species, results from the study of metabolic rates of captive
 78 northern fur seal pups imply that given the climatological SST patterns in the
 79 North Pacific, their thermoregulatory capabilities are insufficient for them to
 80 remain in the Bering Sea and northwestern Gulf of Alaska in winter (Rosen
 81 and Trites, 2014). Water temperature has also been found to influence the
 82 metabolic rate and likely the fasting duration of newly-weaned Antarctic fur
 83 seal pups (Rutishauser et al., 2004), which have a similar morphology to
 84 northern fur seals and also undergo a long post-weaning migration through
 85 subarctic waters that initiates in the fall. The metabolic constraints imposed
 86 by cold SSTs, in combination with the observed general southward movement
 87 of northern fur seal pups toward warmer SST after their departure from
 88 the Pribilof Islands, suggest that SST is possibly an additional migratory
 89 environmental cue for this species.

90 In this paper our objectives are to develop the CTSMC model from first
 91 principles such that we do not need to rely on the distributional assumptions
 92 of Hanks et al. (2015) with respect to cell residence time (Section 2). Al-
 93 though, see Hooten et al. (2017, Sec. 7.4) for an alternative development of
 94 the CTMC model. Because the habitat can change before the animal moves
 95 cells we can not rely on a constant rate of movement while the animal re-

96 mains in a cell. Following CTSMC development, in Section 3, we show that
 97 with only slight modification, the imputation and GLM approach of Hanks
 98 et al. (2015) can still be used for parameter inference. In addition to model
 99 fitting we illustrate methodology for Bayesian model selection using this 2
 100 stage method. In Section 4, we illustrate the full methodology for fitting and
 101 model selection using migration data from 15 northern fur seal pups tagged
 102 in Alaska, USA.

103 2 Continuous-Time Semi-Markov Model for 104 Movement

105 The description of the data and notation generally follows that of Hanks
 106 et al. (2015), however, we have made a few alterations to accommodate the
 107 temporal dynamics of the semi-Markov generalization of their model. In a
 108 continuous-time, discrete-space movement model the geographical study do-
 109 main is partitioned into a cells indexed by $\mathcal{G} = \{1, \dots, n\}$. These cells can
 110 be thought of as a mathematical graph with nodes represented by the cell
 111 centroids. Each node, i , has a set of neighboring nodes $\mathcal{N}_i = \{j \in \mathcal{G} : i \sim j\}$,
 112 where $j \sim i$ means that cells i and j are directly connected and the animal
 113 can move from cell i to cell j in one move. For example, neighboring cells
 114 might share a border in a raster partition, but, cells might also represent spa-
 115 tially separate patches that the animal might visit. As with all semi-Markov
 116 models, the complete path of the animal, $\mathcal{P}$, can decomposed into the times
 117 at which cell transitions are made (jump times), $\boldsymbol{\tau} = \{\tau_0, \tau_1, \dots, \tau_K\}$, and
 118 the sequence of cells visited (embedded Markov chain) $\mathbf{G} = \{G_0, \dots, G_K\}$.
 119 We will assume that $\tau_0 = 0$ and G_0 is the known starting cell. We use the
 120 notation $G(t)$ to represent the location of the animal for any time in $[0, T]$,
 121 e.g., $G(t) = G_k$ for $t \in [\tau_k, \tau_{k+1})$.

2.1 General likelihood formulation

The heart of the CTSMC model is the emigration rate function

$$\lambda_{ij}(t|\mathcal{P}_t) = \lim_{h \rightarrow 0} [G(t+h) = j|\mathcal{P}_t]/h, \quad (1)$$

where $\mathcal{P}_t$ is the path of the animal up to time t , i.e., $\mathcal{P}_t = \{\boldsymbol{\tau}_t, \mathbf{G}_t\}$, where $\boldsymbol{\tau}_t = \{\tau_k : \tau_k < t\}$ and $\mathbf{G}_t = \{G_k : \tau_k < t\}$. Herein, we use the notation ‘ $[A|B]$ ’ to represent the probability density (distribution) function of A given B . Therefore, $[G(t+h) = j|\mathcal{P}_t]$ is the probability that $G(t+h) = j$ given the path up to time t . The total rate of emigration from cell i at time t is the accumulation of the emigration rate $\lambda_{ij}(t|\mathcal{P}_t)$ over $\mathcal{N}_i$,

$$\Lambda_i(t|\mathcal{P}_t) = \sum_{j \in \mathcal{N}_i} \lambda_{ij}(t|\mathcal{P}_t). \quad (2)$$

Using the total rate of emigration from a cell i at time t , we can obtain the distribution of the time of the next move, τ_k given the animal moved to cell i at time τ_{k-1} . Using standard results from temporal point process methodology (see Hooten et al., 2017, Section 3.1.6), where the emigration rate function is mathematically equivalent to a point process intensity function, the density function of next movement time is,

$$[\tau_k|\mathcal{P}_{\tau_k}] = \Lambda_i(\tau_k|\mathcal{P}_{\tau_k}) \exp \left\{ - \int_{\tau_{k-1}}^{\tau_k} \Lambda_i(u|\mathcal{P}_u) du \right\}. \quad (3)$$

There is one complication that needs to be addressed at the end of the telemetry device deployment, the censoring of the last observation. If the individual is followed from time $\tau_0 = 0$ to time $T \geq \tau_K$, the last cell movement, then the time between τ_K and T also needs to be accommodated. So, let τ_{K+1} be the unobserved ‘next’ transition occurring after the animal is no longer being observed at time T . The quantity necessary to accommodate the censoring is,

$$\begin{aligned} [\tau_{K+1} > T|\mathcal{P}_T] &= 1 - \int_{\tau_K}^T [\tau|\mathcal{P}_\tau] d\tau \\ &= \exp \left\{ - \int_{\tau_K}^T \Lambda_i(u|\mathcal{P}_u) du \right\} \end{aligned} \quad (4)$$

(Hooten et al., 2017). For the remainder of the model development we will notationally set $\tau_{K+1} = T$, the end of the observation window for the telemetry deployment.

Now, if one conditions on the fact that the next move will be at time τ_k , then the cell to which the move is made, is a categorical variable over $\mathcal{N}_i$,

$$[G_k = j | \tau_k, \mathcal{P}_{\tau_k}] = \frac{\lambda_{ij}(\tau_k | \mathcal{P}_{\tau_k})}{\Lambda_i(\tau_k | \mathcal{P}_{\tau_k})} \quad (5)$$

(Norris, 1998). If we now assume that the emigration hazard function, $\lambda_{ij}(t | \mathcal{P}_t)$, is also a function of a parameter vector, say $\boldsymbol{\theta}$ then, given an observed path $\mathcal{P}$, the likelihood is,

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta} | \mathcal{P}) &= \left\{ \prod_{k=1}^K [G_k | \tau_k, \mathcal{P}_{\tau_k}] \times [\tau_k | \mathcal{P}_{\tau_k}] \right\} \times [\tau_{K+1} > T | \mathcal{P}_T] \\ &= \prod_{k=1}^K \frac{\lambda_{G_{k-1}, G_k}(\tau_k | \mathcal{P}_{\tau_k})}{\Lambda_{G_{k-1}}(\tau_k | \mathcal{P}_{\tau_k})} \times \Lambda_{G_{k-1}}(\tau_k | \mathcal{P}_{\tau_k}) \exp \left\{ - \int_{\tau_{k-1}}^{\tau_k} \Lambda_{G_{k-1}}(u | \mathcal{P}_u) du \right\} \\ &\quad \times \exp \left\{ - \int_{\tau_K}^T \Lambda_{G_K}(u | \mathcal{P}_u) du \right\} \\ &= \left\{ \prod_{k=1}^K \lambda_{G_{k-1}, G_k}(\tau_k | \mathcal{P}_{\tau_k}) \right\} \times \exp \left\{ - \int_0^T \sum_{j \in \mathcal{N}_{G(u)}} \lambda_{G(u), j}(u | \mathcal{P}_u) du \right\} \end{aligned} \quad (6)$$

There are a couple of notes here,

2.2 A proportional rate model

Now that we have a general form for the likelihood of the movement process, we can consider specifying a model that includes the external variables which might be thought to influence movement. Here we follow the same model form given by Hanks et al. (2015) and Hanks and Hughes (2016), however,

we generalize it to accommodate a time-varying structure and place it in a survival analysis context with a rich history of model development. In the survival modeling context, one of the most popular class is the proportional hazards models (Cox and Oakes, 1984). In the CTSMC movement context, for $G(t) = i$ and $j \in \mathcal{N}_i$, the proportional rates (hazards) model with a time-varying covariate is given by

$$\lambda_{ij}(t|\mathcal{P}_t) = \lambda_0(t|\mathcal{P}_t) \exp \{ \mathbf{m}'_i(t)\boldsymbol{\delta} + \mathbf{x}'_{ij}(t)\boldsymbol{\gamma} \}, \quad (7)$$

where $\lambda_0(t|\mathcal{P}_t)$ a a baseline hazard function, $\mathbf{m}_i(t)$ is a covariate associated only with cell i at time t , $\mathbf{x}_{ij}(t)$ is a covariate associated with both cell i and j (e.g., a gradient of an environmental variable) at time t , and $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are regression coefficients associated with the respective covariates. The $\mathbf{m}$ variables are motility drivers in that they only affect the rate at which the animal will leave a cell, they have no influence on whether an animal is attracted to another cell; they disappear in equation (5). Whereas, the ‘gradient’ variables, $\mathbf{x}$, influence rate of movement by drawing animals to other neighboring cells.

At this point, the model (7) is similar to that given by Hanks and Hughes (2016), however, there are some notable differences. First, and most obvious, the covariates are potentially time-varying. Second, the baseline transition rate need not be constant over time. The transition hazard can change with time since the last transition. This leads to clustered (or regular) transition times with respect to those that are exponentially distributed in the CTMC model. Another difference between the CTMC and CTSMC models are how the model treats temporally varying coefficients. In (7), the coefficients, $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are constant though time. But, as with the Markov version, one might consider $\boldsymbol{\gamma}(t)$ or $\boldsymbol{\delta}(t)$. However, the difference in the CTSMC extension is that the coefficients can change smoothly through time even as the animal remains within the cell. The coefficients of the CTMC can only change when the animal transitions cells. This seems more mathematically appropriate, however, the realized difference in effect may be small in practice.

In addition to the obvious link between the CTSMC model and its CTMC special case, there is also a similarity with the continuous-space and time

point process analysis of movement data presented by Johnson et al. (2013). The spatio-temporal point process likelihood is very similar to the CTSMC likelihood in (6) with the exception that the summation in the right hand side is replaced with an integration over the continuous spatial domain of neighboring sites. It is obvious that if we allow the cells to shrink in area while simultaneously increasing the number of neighbors to maintain equal neighborhood area, the summation will become an integral in the limit. Therefore, we can think of the CTSMC model as a spatially discrete approximation of the spatio-temporal point process of Johnson et al. (2013). The benefit of the CTSMC model is that by defining transitions between a small number of neighboring cells, the computation of the likelihood becomes much less demanding.

3 Statistical Inference

3.1 Likelihood calculation

The log-likelihood function in equation (6) is often challenging to evaluate in general. Hanks et al. (2015) present an approximation that allows users to obtain maximum likelihood estimates using standard GLM software for the CTMC model. Further, this approximation can be made as accurate as desired. In the context of survival analysis, this same approach was used by Holford (1980) and Laird and Olivier (1981). Following the same development used in the survival analysis, we can form an approximation for calculating the CTSMC movement model likelihood.

To begin the approximation, we first select a set of quadrature points, q_l , to numerically evaluate the integral in (6). These quadrature points include times when cell transitions occurred, $\tau = \{\tau_0, \dots, \tau_K\}$, and times where dynamic covariates changed values. In addition, the quadrature set also includes any additional points that the practitioner feels necessary to approximate the integral. For example, in Section 4, a 15 minute grid of times was also included in the quadrature set to model seal movement. Then, the approximation assumes $\lambda(t|\mathcal{P}_t) = \lambda(q_l|\mathcal{P}_{q_l})$ for $t \in (q_{l-1}, q_l]$. To ease the

226 notational burden, we also set,

$$\begin{aligned} G_l &= G(q_l) \text{ for } l = 0, \dots, L, \\ \text{and} \\ \lambda_{lj} &= \lambda_{G_{l-1},j}(q_l | \mathcal{P}_{q_l}) \text{ for } l = 1, \dots, L \text{ and } j \text{ in } \mathcal{N}_{G_{l-1}}. \end{aligned} \quad (8)$$

228 Now the integral is approximated with the summation,

$$229 \quad \int_0^T \sum_{j \in \mathcal{N}_{G(u)}} \lambda_{G(u),j}(u | \mathcal{P}_u) \approx \sum_{l=1}^L \sum_{j \in \mathcal{N}_{G_{l-1}}} \delta_l \lambda_{lj} \quad (9)$$

230 as arbitrarily close as desired, where $\delta_l = q_l - q_{l-1}$. After placing the
231 approximation (9) into (5), one obtains the approximate likelihood function,

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta} | \mathcal{P}) &\approx \prod_{l=1}^L \prod_{j \in \mathcal{N}_{G_{l-1}}} \lambda_{lj}^{Z_{lj}} \exp\{-\delta_l \lambda_{lj}\} \\ &\propto \prod_{l=1}^L \prod_{j \in \mathcal{N}_{G_{l-1}}} \tilde{\lambda}_{lj}^{Z_{lj}} \exp\{-\tilde{\lambda}_{lj}\} \end{aligned} \quad (10)$$

233 where $\tilde{\lambda}_{lj} = \delta_l \lambda_{lj}$ and $Z_{lj} = 1$ if $q_l \in \boldsymbol{\tau}$ and $G_l = j$ for some $k = 1, \dots, K$,
234 that is, q_l is a transition time and j is the cell transitioned to at time q_l ,
235 else $Z_{lj} = 0$. Notice, as with the model specification of Hanks and Hughes
236 (2016) and Hanks et al. (2015), that the likelihood function formed by (10)
237 is proportional to a Poisson likelihood function where Z_{lj} are the ‘data’ and
238 $\tilde{\lambda}_{lj}$ are the rate parameters. The only difference between this formulation
239 and that of Hanks et al. (2015) is that there are times, q_l where $Z_{lj} = 0$ for
240 all j , that is, there was no movement to a cell, but the λ_{ij} have changed.

241 3.2 GLM/GAM and the proportional rate model

242 Using the the Poisson approximation (10) one can gain powerful computa-
243 tional assistance if the proportional hazards model (7) is used. Because of
244 the linear structure on the log scale, standard GLM software can be used to

245 fit the CTSMC movement model. On the log scale, the Poisson rate used for
246 the Z_{lj} data is

$$\begin{aligned} \log \tilde{\lambda}_{lj} = & \log \delta_l + \log \lambda_0(q_{l-1} - \tau_l^*) \\ & + \mathbf{m}'_{G_{l-1}}(q_{l-1})\boldsymbol{\delta}(q_{l-1}) + \mathbf{x}'_{G_{l-1},j}(q_{l-1})\boldsymbol{\gamma}(q_{l-1}), \end{aligned} \quad (11)$$

248 where $\log \delta_l$ takes the form of a known offset and $\lambda_0(q_{l-1} - \tau_l^*)$ is the baseline
249 cell transition rate that depends on τ_l^* , the time of the last transition prior
250 to q_l . Note, that we have now allowed the coefficients to vary with time now.
251 Assuming λ_0 is constant (we will discuss this momentarily), the parameters
252 are all linear on the log scale, so any GLM fitting software can be used.
253 One may also estimate time-varying coefficients using a ‘varying coefficients’
254 model as illustrated in the present analysis of northern fur seal migration in
255 Section 4.

256 We now examine the baseline rate function, λ_0 . To begin, if one models
257 $\log \lambda_0(q_{l-1} - \tau_l^*) = \beta_0 + \beta_1 \log(q_{l-1} - \tau_l^*)$ then the proportional rate model is
258 still in a log-linear form, so standard GLM software can be used. This log-
259 linear model for λ_0 implies the waiting times between cell transitions follows a
260 Weibull distribution if covariates are not considered (Cox and Oakes, 1984).
261 If $\beta_1 < 0$ then the rate function decreases with time since last transition
262 and the times of transition will tend to be clustered in time (Hooten et al.,
263 2017). The reverse is true for $\beta_1 > 0$; transitions will occur at more regular
264 intervals. For $\beta_1 = 0$ the residence times will be exponentially distributed
265 as in the CTMC model (Hanks and Hughes, 2016). If $q_{l-1} - \tau_l^*$ is used
266 instead of the log version, the residence times will be distributed according
267 to the Gompertz-Makeham distribution (Cox and Oakes, 1984). However,
268 if neither of these simple models are useful, it is straightforward to model
269 λ_0 nonparametrically using a GAM model to directly estimate λ_0 treating
270 $q_{l-1} - \tau_l^*$ as a covariate to smooth.

3.3 Path uncertainty and the process imputation approach

Until this point, we have constructed the model and likelihood as if, in practice, we would observe $\mathcal{P}$ as data. This may be the case in the future as satellite telemetry devices improve, e.g., see Liu et al. (2014) for path reconstruction at subsecond intervals using accelerometer tags. However, for the most part, locations are observed sparsely and irregularly throughout the course of deployment. Therefore, we recommend the use of an *imputation* approach (Scharf et al., 2017) for analysis of traditional telemetry data in this framework. The imputation approach in movement analysis was initially proposed by Hooten et al. (2010) and is also used by Hanks et al. (2015) for the Markov version of the discrete space movement model. Although not necessarily proposed for fully Bayesian inference, we take that approach here because it provides a coherent method of analysis from estimation to model selection.

The process imputation approach proceeds by considering the true quasi-continuous (in space and time) path of the animal, $\boldsymbol{\mu}(t)$. If $\boldsymbol{\mu}(t)$ were observed for all t ($\boldsymbol{\mu}$ from here on) then we could summarize it into the data we desire, namely $\mathcal{P} = \mathcal{P}(\boldsymbol{\mu})$. However, locations are usually only observed continuously in space, but sporadically in time with error, say, $\mathbf{y} = \{\mathbf{y}_{\tau_1^*}, \dots, \mathbf{y}_{\tau_n^*}\}$, where τ^* is an observation time in $[0, T]$. To obtain the correct Bayesian posterior distribution for the movement parameters, $\boldsymbol{\theta}$, the posterior distribution must be marginalized over the missing $\boldsymbol{\mu}$ process, that is,

$$\begin{aligned} [\boldsymbol{\theta}|\mathbf{y}] &= \int [\boldsymbol{\theta}|\boldsymbol{\mu}] [\boldsymbol{\mu}|\mathbf{y}] d\boldsymbol{\mu} \\ &\propto \int \frac{\mathcal{L}(\boldsymbol{\theta}|\mathcal{P}(\boldsymbol{\mu})) [\boldsymbol{\theta}]}{[\mathcal{P}(\boldsymbol{\mu})]} [\boldsymbol{\mu}|\mathbf{y}] d\boldsymbol{\mu}. \end{aligned} \tag{12}$$

In practice, a two step procedure is used to describe $[\boldsymbol{\theta}|\mathbf{y}]$ by first drawing a sample from the imputation distribution $\boldsymbol{\mu}_r \sim [\boldsymbol{\mu}|\mathbf{y}]$, $r = 1, \dots, R$, computing the discrete space path from the imputed data, $\mathcal{P}_r = \mathcal{P}(\boldsymbol{\mu}_r)$, then summarizing $[\boldsymbol{\theta}|\mathcal{P}_r]$ in some way. For example, Hanks et al. (2015) describes obtaining the posterior mean and variance of $\boldsymbol{\theta}|\mathcal{P}_r$ using the asymptotic dis-

tribution of the MLE, $\hat{\theta}_r$ as an approximation of the posterior from which one can obtain the unconditional mean and variance of $E(\theta|\mathbf{y}) \approx \tilde{\theta} = \text{mean}(\hat{\theta}_r)$ and $\text{Var}(\theta|\mathbf{y}) \approx \tilde{\Sigma} = \text{mean}(\hat{\Sigma}_r) + \text{var}(\hat{\theta}_r)$, where $\hat{\Sigma}_r$ is the estimated asymptotic covariance matrix for the MLE $\hat{\theta}$. Herein we use ‘mean’ and ‘var’ to represent sample mean and variance over the imputation index, r .

As with imputation approaches in general, calculating interval estimates is not always straightforward (Rubin and Schenker, 1986), as there are many ways one may combine the data to obtain appropriate intervals. For example, Hanks et al. (2015) uses $\tilde{\theta} \pm 2 * \text{diag}(\tilde{\Sigma})$ to obtain marginal intervals. However, this approach assumes that $[\theta|\mathbf{y}]$ is generally normal looking to obtain 95% intervals. We propose an alternative approach based on approximating the posterior with a stochastic sample from the marginal posterior. Because the the imputed paths were drawn from $[\mu|\mathbf{y}]$, a sample from $[\theta|\mathbf{y}]$ can be obtained by drawing samples from each $[\theta|\mu_r]$ and collecting them into one large sample. Then one can simply calculate a credible interval from the joint sample. This of course also applies to any function of the parameters as in any posterior sample. However, note, if different samples sizes are obtained for each $[\theta|\mu_r]$, then for each sample less than the largest, one can resample with replacement to obtain comparable weighting for all imputations. If GLM or GAM software is used to fit the model, one could easily approximate $[\theta|\mu_r] \approx N(\hat{\theta}_r, \hat{\Sigma}_r)$, the large sample MLE or posterior distribution (see Section 4.2).

Another aspect of movement model inference using process imputation is that has received relatively little attention is that of model selection. In the first CTMC paper Hanks et al. (2015) describes use of a regularization prior (penalty) to shrink insignificant coefficients to zero, such that selection can be accomplished for nested models. However, there has been little discussion in the imputation literature with respect to selection from a discrete set of models. There have been previous proposals to perform model selection using *ad hoc* AIC weighting (Nakagawa and Freckleton, 2011). However, we propose maintaining the Bayesian paradigm inspiring the imputation approach and use posterior model probabilities to perform selection and model averaging to form a coherent process in analysis.

333 To begin formulating Bayesian model selection we now suppose that there
 334 is a set of P models, $\mathcal{M} = M_1, \dots, M_P$ for the movement rate, $\lambda_{ij}(t|\mathcal{P}_t, M_p)$
 335 from which the likelihood, $\mathcal{L}(\boldsymbol{\theta}_p|\mathcal{P}, M_p)$, is formed. Now, using the same
 336 decomposition as (12), we can formulate

$$337 \quad [M_p|\mathbf{y}] = \int [M_p|\mathcal{P}_r] [\boldsymbol{\mu}_r|\mathbf{y}] d\boldsymbol{\mu}, \quad (13)$$

338 where,

$$339 \quad [M_p|\mathcal{P}_r] \propto [\mathcal{P}_r|M_p] [M_p] = \left\{ \int \mathcal{L}(\boldsymbol{\theta}_p|\mathcal{P}_r, M_p) [\boldsymbol{\theta}_p|M_p] d\boldsymbol{\theta}_p \right\} [M_p], \quad (14)$$

340 $[\boldsymbol{\theta}_p|M_p]$ is the parameter prior distribution under model M_p , and $[M_p]$ is
 341 the prior model probability. Bayesian model selection has a large volume of
 342 literature devoted to its methodology. However for application within the
 343 imputation framework, we seek a method to evaluate $[M_p|\mathcal{P}_r]$ that is com-
 344 putationally efficient even if it is approximate, as the whole model selection
 345 procedure needs to be replicated for each $\boldsymbol{\mu}_r$. Therefore, we propose using
 346 Laplace approximations to equation (14; Kass and Raftery 1995),

$$347 \quad \log[\mathcal{P}_r|M_p] \approx B(M_p|\mathcal{P}_r) = \log \mathcal{L}(\hat{\boldsymbol{\theta}}_{rp}|\mathcal{P}_r, M_p) + \log[\hat{\boldsymbol{\theta}}_{rp}|M_p] \quad (15)$$

$$- \log |\hat{\mathbf{H}}_{rp}|/2 + d_p \log(2\pi)/2,$$

348 where $\hat{\boldsymbol{\theta}}_{rp}$ is the posterior mode of $[\boldsymbol{\theta}_p|\mathcal{P}_r, M_p]$, $\hat{\mathbf{H}}_{rp}$ is the Hessian matrix
 349 of the negative log posterior density, and d_p is the dimension of $\boldsymbol{\theta}_p$. Kass
 350 and Raftery (1995) also note that the approximation is still of the same
 351 order if the MLE of $\boldsymbol{\theta}$ is used provided the prior exerts little influence on
 352 the posterior relative to the likelihood. The benefit of that is one can use
 353 standard GLM software to fit the model without worrying about the prior
 354 distribution. However, proper priors are still necessary (highly advised) for
 355 terms that are being selected over (Bayarri et al., 2012).

356 Placing all the pieces back together we obtain an approximation for the
 357 posterior model probability that we can use for model selection

$$358 \quad [M_p|\mathbf{y}] \approx \text{mean} \left(\frac{\exp\{B(M_p|\mathcal{P}_r)\} [M_p]}{\sum_{p'} \exp\{B(M_{p'}|\mathcal{P}_r)\} [M_{p'}]} \right). \quad (16)$$

359 The Laplace approximation allows us to have a coherent approach to aver-
360 aging an information criterion within the imputation framework.

361 4 Northern Fur Seal Pup Migration

362 To demonstrate analysis of movement data, we analyzed telemetry locations
363 from 15 female northern fur seal pups tagged in the Pribilof Is., Alaska prior
364 to their annual migration. Female ‘355’ from this data set was previously
365 analyzed by Johnson et al. (2008) using a continuous time random walk
366 (CTCRW). However, the previous analysis did not include any habitat re-
367 lated covariates to associate with movement. So, here we extend their anal-
368 ysis with the use of spatio-temporal covariates. Code for reproducing this
369 analysis is available in an online appendix ². These data were analyzed using
370 the R Statistical Environment (R Core Team, 2018) due to the availability
371 of all necessary spatial, data manipulation, movement modeling, and GLM
372 fitting functions, however, nothing precludes development in other environ-
373 ments that contain spatial data manipulation and GLM capabilities.

374 4.1 A dynamic environment

375 Environmental covariates were defined on a hexagonal spatial grid covering
376 the domain of the pup tracks that departed during fall 2005. For the process
377 imputation, 20 tracks were generated using draws from the posterior distribu-
378 tion of locations every 15 minutes from a continuous-time correlated random
379 walk model with a correlated random drift process (Johnson et al., 2011)
380 using the R package `crawl` (Johnson and London, 2018) (Figure 1). Other
381 methods of approximating the imputation distribution, $[\mu|y]$, could be used
382 as well, e.g., basis function movement models (Hooten and Johnson, 2017;
383 Scharf et al., 2017; Buderman et al., 2015). A regular hexagonal lattice, with
384 40 km center point lateral separation, and associated hexagonal polygons
385 (“cells”) were generated. Only those cells that contain part of an imputed

²Upon acceptance a DOI will be issued for the repository, for now, code can be accessed
at: https://github.com/dsjohnson/ctsmc_movement

track or neighbors of those cells are necessary for fitting, so, only those cells are retained hereafter (see Figure 2 for an example with data from female 355). Model covariates were defined on the hexagonal grid using environmental variables obtained from remote sensing observations or atmospheric reanalysis. Along with all of the data, an animation of the spatio-temporal environmental variables is available in the online repository³.

Estimates of geostrophic ocean surface currents – slowly-varying oceanic motions (cm/s) forced by mesoscale (~ 10 -200 km) gradients of sea surface height (Vallis, 2006) – were obtained from the AVISO All-Satellite Absolute Dynamic Topography (ADT) product (<http://aviso.altimetry.fr/index.php?id=1269>). The ADT product provides a smoothed, gridded estimate of the sea surface height at $1/4^\circ$ spatial and daily temporal resolution based on measurements of the sea level anomaly from orbiting radar altimeters, combined with estimates of the mean sea surface obtained from *in situ* hydrographic measurements (Rio et al., 2013). At each grid point, s , the surface geostrophic current $\mathbf{x}_{curr}(s, t) = (u_{curr}(s, t), v_{curr}(s, t))'$ has a component u_{curr} in the east direction along unit vector and v_{curr} in the north direction.

Atmospheric surface (10 m height) wind estimates $\mathbf{x}_{wind} = (u_{wind}(s, t), v_{wind}(s, t))'$ (in m/s) were obtained from the National Centers for Environmental Prediction/National Center for Atmospheric Research Reanalysis 1 (R1) product (<http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html>). This product gives estimates of atmospheric variables at 6-hourly resolution on a roughly 2° global spatial grid and is derived from assimilating observational data into a global forecast model (Kalnay et al., 1996). Comparison of R1 to independent observations in the North Pacific Ocean has shown that, away from topographic features, R1 winds are well-correlated with observed winds with a weak positive bias in magnitude (Ladd and Bond, 2002).

Sea surface temperature ($^\circ\text{C}$) estimates, $x_{sst}(s, t)$, were obtained from the NOAA Optimal Interpolation V2 High Resolution dataset (<http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html>).

³<https://figshare.com/account/home#/collections/3995112>

These estimates are derived from *in situ* data and the Advanced Very High Resolution Radiometer satellite instrument (Reynolds et al., 2007) and provided on a $1/4^\circ$ global grid at daily temporal resolution.

Each of the environmental variables described above was spatially interpolated to the hexagonal grid cells at 6-hourly intervals. If a cell overlapped a grid boarder, the weighted average value was used based the areas contained in the hex. The 6-hour interval corresponds to the shortest resolved interval among the available variables (R1 winds); for variables obtained at daily resolution, values are repeated in multiple time intervals in each cell. For the gradient based variables, the covariates for cell transitions were defined by the projection of the gradient vector to neighboring cell centroids, that is,

$$\mathbf{wind}_{ijt} = \mathbf{x}_{wind}(s_i, t)' \mathbf{w}_{ij} \quad \text{and} \quad \mathbf{curr}_{ijt} = \mathbf{x}_{curr}(s_i, t)' \mathbf{w}_{ij}, \quad (17)$$

where s_i is the location of the i th hex centroid and $\mathbf{w}_{ij}$ is a unit length vector pointing from the hex i centroid to the hex j centroid (Hanks et al., 2015). SST was treated as a motility covariate to test whether pups might adjust speed if they encounter favorable conditions with respect to SST, $\mathbf{SST}_{it} = x_{sst}(s_i, t)$

In addition, to wind and current, we defined three additional “kinematic” covariates to account for the possible effects of any background drift/movement tendency independent of environmental variables:

- $\mathbf{residence}_t = t - \tau_t^*$
- $\mathbf{previous}_{ijt} = \mathbf{w}_t' \mathbf{w}_{ij},$
- $\mathbf{north}_{ij} = (0, 1)' \mathbf{w}_{ij},$
- $\mathbf{east}_{ij} = (1, 0)' \mathbf{w}_{ij},$

where τ_t^* is the time of the previous transition (to the current cell), $\mathbf{w}_t$ is the unit vector of the last transition prior to t . $\mathbf{north}$ and $\mathbf{east}$ were included to account for pup migration tendency. The $\mathbf{residence}$ variable is the residence time within the cell (up to t). $\mathbf{previous}$ was added to reflect correlated movement on a finer scale than general migration patterns and

446 is the direction of the previous cell transition. It is typical of animal move-
447 ment that individuals tend to move in the same direction for some time in a
448 correlated fashion.

449 In addition to the cell transition times and covariate changing times,
450 locations on each imputed path, μ_r , were generated at 15 minute intervals
451 to approximate the likelihood integral. Thus, for each quadrature time on
452 each imputed path, q_l , a six row data set ($j = 1 \dots 6$) was created with
453 variables: $\mathbf{z} = Z_{lj}$, $\text{delta} = q_l - q_{l-1}$, $\text{time} = q_l$, $\text{residence} = q_l - \tau_{q_l}^*$,
454 $\text{previous}_{ij,q_l}$, north_{ij} , east_{ij} , sst_{i,q_l} , wind_{ij,q_l} , and curr_{ij,q_l} . The data set
455 was then concatenated over all q_l within an imputation, this forms the model
456 data, say ‘glm_data’ for the r th imputation. From this data all models can
457 be fitted using, say glm or gam with family=‘poisson’.

458 4.2 Model formulation, estimation, and selection

459 In pseudo-R notation, we defined the kinematic ‘Base’ movement model to be
460 $\sim \log(\text{residence}) * \text{previous} + \text{B:north} + \text{B:east}$. The B notation refers
461 to a basis matrix used to describe the smoothly varying (over time) coeffi-
462 cients with which they interact. Here, the λ_0 term in equation (7) is modeled
463 with a $\log(\text{residence}) * \text{previous}$ interaction effect to account for the syn-
464 ergistic effect of kinetic movement on the base residence time within a cell.
465 That is, if there is a strong relationship between the previous movement and
466 the current one, then the residence time is likely to be longer because the
467 animal is probably transiting the entire cell. However, if the animal is go-
468 ing back and forth across a border, residence time will likely be short and
469 relationship to the previous movement will likely be negative. Because of
470 the length of deployment, it is natural to inquire whether an individual’s
471 movement in response to the environment changes over the course of time.
472 The **previous** and **residence** terms were modeled as constant due to the
473 fact that initial investigation showed no improvement by allowing their effect
474 to vary over time. All other model terms were modeled as time varying. To
475 accomplish this we used a radial basis function model for $\delta(t)$ and $\gamma(t)$, that

476 is,

$$477 \quad \boldsymbol{\delta}(t) = \sum_k r(t, \kappa_k, \phi) \boldsymbol{\delta}_k, \quad (18)$$

478 where $r(t, \kappa_k, \phi) = \exp\{(t - \kappa_k)^2 / \phi^2\}$, κ_k is a knot in $[0, T]$, $\boldsymbol{\delta}_k$ are the basis
479 weights, and ϕ is a basis scaling parameter. Herein we chose 3 equally spaced
480 interior knots plus 2 additional knots at equal spacing before and after $[0, T]$
481 for a 5 basis model. The scalar ϕ set to the knot spacing interval.

482 For this analysis we used the R package `mgcv` and the `gam` function to
483 estimate the posterior mode $\hat{\boldsymbol{\theta}}$, and obtain the Hessian matrix $\hat{\mathbf{H}}$, for the
484 Laplace approximation in the posterior model probability (PMP) calculation
485 as well as the approximating the conditional posterior

$$486 \quad [\boldsymbol{\theta}_{rp} | \mathcal{P}_r, M_p] = N(\hat{\boldsymbol{\theta}}_{rp}, -\hat{\mathbf{H}}_{rp}^{-1}) \quad (19)$$

487 used for effect inference. To increase computational efficiency, for each ani-
488 mal fits were completed in parallel over the 20 imputed paths. This was ac-
489 complished with use of the R packages: `future` (Bengtsson, 2018), `doFuture`
490 (Bengtsson, 2017), and `foreach` (Microsoft and Weston, 2017).

491 The posterior model probability (PMP) approximation to select between
492 the 8 models composed of all subsets of the environmental variables (Table
493 1). For each model and imputation, we used the prior distribution

$$494 \quad [\boldsymbol{\theta}_{rp} | M_p] = N(\mathbf{0}, (\xi_{rp} \mathbf{X}'_{rp} \mathbf{X}_{rp})^{-1}), \quad (20)$$

495 where $\mathbf{X}_{rp}$ is the design matrix for model (r, p) and ξ_{rp} is a scalar. A flat
496 prior distribution was used for ξ_{rp} in each model.

497 4.3 Results

498 There was strong evidence for an influence of surface winds on movement with
499 the posterior probability of the wind only model close or equal to 1 for 10 of
500 15 animals (Table 2). Of those animals in which the wind only model was not
501 the maximum *a posteriori* model, the model with wind and geostrophic current

502 effects was the maximum PMP model for 3 of the remaining 5. Finally, the
503 remaining 2 animals were basically split between models 2 and 5.

504 We will now take a look at the temporally varying effects of animal ‘355’.
505 Here we will examine the effects for model 8 even though model 2 was the
506 highest PMP model for this animal. However, it is instructive to see the
507 general pattern of all covariates (Figure 3). For animal 355, the wind effect
508 is overall positive throughout the course of the deployment implying that
509 the animal generally travels with the prevailing winds although, the strength
510 of this effect seems to vary with time. There also seems to be a generally
511 positive effect of geostrophic current on the rate of movement, although for
512 only a small fraction of the deployment is it significantly positive according
513 to the 95% credible interval. This small interval is not enough to make the
514 covariate overall significant as the wind only model has $PMP = 0.99$. SST
515 has a negative but insignificant effect meaning that the animal tends to slow
516 in warmer seas, but it is highly uncertain. The kinetic nuisance effects, north
517 and east are also illustrated and describe the general trend in movement as
518 expected, first southerly movement followed by generally easterly movement.

519 5 Discussion

520 Here we’ve presented an extension to the CTMC model of Hanks et al. (2015)
521 for animal movement modeling. In addition to the ability to handle dynamic
522 habitat covariates the semi-Markov version (CTSMC) relaxes the assumption
523 of exponentially distributed residence times for each cell. However, even after
524 the extensions, the Poisson likelihood approximation still holds with only
525 slight considerations, that is at quadrature points without cell transition,
526 none of the Z_{ij} values are set to 1. It can be thought of as the absence of
527 movement at that time. Thus, there is no additional difficulty in practice to
528 fit a temporally dynamic model (covariates and coefficients) as it is to fit the
529 time-constant habitat CTMC model.

530 To fully account for path uncertainty in the analysis, we used a very sim-
531 ilar 2-stage imputation approach used by Hanks et al. (2015). However, an

alternative is to use a ‘stacked’ likelihood (e.g., White et al. 2011; Hanks and Hughes 2016). Thus, one can make inference using this weighted likelihood and avoid having to recombine separate model inference. However, the spatio-temporal data sets built using the methodology of Section 4.1 might often preclude this as the complete data set might become too large to deal with efficiently for, say, a complex GAM model. It would be more efficient to analyze each imputed data set separately and recombine the results in a buffered fashion. In this way the separate fits could be parallelized for increased computational speed as we have done in the example herein.

In addition to extension of the CTMC model to a semi-Markov version, we have also explored model selection in the context of a 2-stage imputation approach. To our knowledge, this general topic has not been explored outside of Nakagawa and Freckleton (2011), and only with respect to AIC model averaging therein. Here we propose a fully Bayesian model selection approach in an imputation setting using an approximated version of the posterior model probability (PMP). PMPs can then be averaged over imputations in accordance with the Bayesian interpretation of multiple imputation to account for path uncertainty in selection.

The Laplace approximation used in the PMP calculations allows for computational efficiency over all of the imputed paths. One might question whether the approximation provides values close enough to the true marginal distribution $[M_p|\mathcal{P}_r]$. While there is no definitive way to know whether it is the case for any particular data set, Kass and Raftery (1995) note that sample sizes $< 5d_p$ are worrisome, while those $> 20d_p$ should be sufficiently accurate. The question for these models is, “what is the sample size?” Volinsky et al. (1997) notes that the appropriate quantity for proportional hazard models is the number of observed events in the data. In the CTSMC models this would be the number of observed cell transitions $n_r = \sum_l Z_{lj}$ for each $\mathcal{P}_r$. In the fur seal example the largest model (model 8) contains 30 parameters, thus sample sizes > 600 should cause little worry for any of the models fitted here. For the fur seal data, only 1 animal fell below this threshold and its minimum was $n_r = 560$ over all 20 imputed paths. So, we believe that n_r was sufficient in all animals for this analysis. However, this is something to

keep in mind for other data. It should be noted that for each imputation full PMPs must be calculated so that imputation dependent normalization constants in (14) cancel. If one uses, say, B (eqn. 15) or BIC, alone the value will not be comparable over imputations because $[\mathcal{P}_r]$ has not been included. This is essentially the same reason that comparing AIC or BIC values between datasets is not valid.

The CTSMC model presents a computationally feasible approach for making inference with respect to the influence of dynamically changing habitat on animal movement. The ratio λ_{ij}/Λ_i essentially represents a discrete choice resource selection model (McCracken et al., 1998) with the additional inference on rate of movement modeled as well. The imputation methodology allows one to disentangle the selection from the rate of movement by regularizing the modeled locations of the animal. Thus, we recommend the CTSMC approach as a general purpose resource selection and movement analysis tool.

Acknowledgments

The findings and conclusions in the paper are those of the authors and do not necessarily represent the views of the National Marine Fisheries Service, NOAA. Reference to trade names does not imply endorsement by the National Marine Fisheries Service, NOAA.

References

- Avgar, T., Potts, J. R., Lewis, M. A., and Boyce, M. S. (2016). Integrated step selection analysis: bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution*, In press.
- Bayarri, M. J., Berger, J. O., Forte, A., García-Donato, G., et al. (2012). Criteria for bayesian model choice with application to variable selection. *The Annals of statistics*, 40(3):1550–1577.

- 591 Bengtsson, H. (2017). *doFuture: A Universal Foreach Parallel Adaptor using*
592 *the Future API of the 'future' Package*. R package version 0.6.0.
- 593 Bengtsson, H. (2018). *future: Unified Parallel and Distributed Processing in*
594 *R for Everyone*. R package version 1.8.1.
- 595 Buderman, F. E., Hooten, M. B., Alldredge, M., Hanks, E. M., and Ivan,
596 J. S. (2018). Predatory behavior is primary predictor of movement of
597 wildland-urban cougars. *bioRxiv*, page 257295.
- 598 Buderman, F. E., Hooten, M. B., Ivan, J. S., and Shenk, T. M. (2015).
599 A functional model for characterizing long-distance movement behaviour.
600 *Methods in Ecology and Evolution*, 7:264–273.
- 601 Cox, D. R. and Oakes, D. (1984). *Analysis of survival data*, volume 21. CRC
602 Press.
- 603 Hanks, E. M., Hooten, M. B., and Alldredge, M. W. (2015). Continuous-
604 time discrete-space models for animal movement. *The Annals of Applied*
605 *Statistics*, 9:145–165.
- 606 Hanks, E. M. and Hughes, D. A. (2016). Flexible discrete space models of
607 animal movement. *arXiv*, 1606.07986.
- 608 Holford, T. R. (1980). The analysis of rates and of survivorship using log-
609 linear models. *Biometrics*, 36:299–305.
- 610 Hooper, C. L. (1895). Report of Captain Hooper, dated November 21, 1892.
611 In *US Senate, 53rd Congress, 2nd session. Fur-seal arbitration, proceedings*
612 *of the Tribunal of Arbitration, volume VII, Senate executive document 177*
613 *(serial set 3166)*, pages 228–233. Government Printing Office, Washington,
614 D. C.
- 615 Hooten, M. B. and Johnson, D. S. (2017). Basis function models for nonsta-
616 tionary continuous-time trajectories. *Journal of the American Statistical*
617 *Association*, 512:578–589.

- 618 Hooten, M. B., Johnson, D. S., Hanks, E. M., and Lowry, J. H. (2010). Agent-
619 based inference for animal movement and selection. *Journal of Agricultural*
620 *Biological and Environmental Statistics*, 15(4):523–538.
- 621 Hooten, M. B., Johnson, D. S., Mcclintock, B. T., and Morales, J. M. (2017).
622 *Animal Movement: Statistical Models for Telemetry Data*. Chapman and
623 Hall/CRC, Boca Raton, Florida.
- 624 Johnson, D. S., Hooten, M. B., and Kuhn, C. E. (2013). Estimating animal
625 resource selection from telemetry data using point process models. *Journal*
626 *of Animal Ecology*, 82:1155–1164.
- 627 Johnson, D. S. and London, J. M. (2018). *crawl*: an r package for fitting
628 continuous-time correlated random walk models to animal movement data.
- 629 Johnson, D. S., London, J. M., and Kuhn, C. E. (2011). Bayesian inference
630 for animal space use and other movement metrics. *Journal of agricultural,*
631 *biological, and environmental statistics*, 16(3):357–370.
- 632 Johnson, D. S., London, J. M., Lea, M.-A., and Durban, J. W. (2008).
633 Continuous-time correlated random walk model for animal telemetry data.
634 *Ecology*, 89:1208–1215.
- 635 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin,
636 L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M.,
637 Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K., Ropelewski, C., Wang,
638 J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D. (1996). The
639 NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Me-*
640 *teorological Society*, 77(3):437–471.
- 641 Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American*
642 *Statistical Association*, 90:773–795.
- 643 Ladd, C. and Bond, N. (2002). Evaluation of the NCEP/NCAR reanalysis
644 in the NE Pacific and the Bering Sea. *Journal of Geophysical Research*,
645 107(C10):doi:10.1029/2001JC001157.

- 646 Laird, N. and Olivier, D. (1981). Covariance analysis of censored survival
647 data using log-linear analysis techniques. *Journal of the American Statis-*
648 *tical Association*, 76(374):231–240.
- 649 Lea, M.-A., Johnson, D., Ream, R., Sterling, J., Melin, S., and Gelatt, T.
650 (2009). Extreme weather events influence dispersal of naive northern fur
651 seals. *Biology Letters*, 5:252–257.
- 652 Liu, Y., Battaile, B. C., Zidek, J. V., and Trites, A. (2014). Bayesian melding
653 of the dead-reckoned path and gps measurements for an accurate and high-
654 resolution path of marine mammals. *arXiv preprint arXiv:1411.6683*.
- 655 McCracken, M. L., Manly, B., and Vander Heyden, M. (1998). The use of
656 discrete-choice models for evaluating resource selection. *Journal of Agri-*
657 *cultural, Biological and Environmental Statistics*, 3:268 – 279.
- 658 Microsoft and Weston, S. (2017). *foreach: Provides Foreach Looping Con-*
659 *struct for R*. R package version 1.4.4.
- 660 Nakagawa, S. and Freckleton, R. P. (2011). Model averaging, missing data
661 and multiple imputation: a case study for behavioural ecology. *Behavioral*
662 *Ecology and Sociobiology*, 65(1):103–116.
- 663 Norris, J. R. (1998). *Markov Chains*. Cambridge University Press, 2nd
664 edition edition.
- 665 Pelland, N., Sterling, J., Lea, M.-A., Bond, N., Ream, R., Lee, C., and
666 Eriksen, C. (2014). Fortuitous Encounters between Seaglidors and Adult
667 Female Northern Fur Seals (*Callorhinus ursinus*) off the Washington (USA)
668 Coast: Upper Ocean Variability and Links to Top Predator Behavior. *PLoS*
669 *ONE*, 9(8):e101268.
- 670 R Core Team (2018). *R: A Language and Environment for Statistical Com-*
671 *puting*. R Foundation for Statistical Computing, Vienna, Austria.

- 672 Ream, R. R., Sterling, J. T., and Loughlin, T. R. (2005). Oceanographic fea-
673 tures related to northern fur seal migratory movements. *Deep-Sea Research*
674 *II*, 52:823–843.
- 675 Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and
676 Schlax, M. G. (2007). Daily High-Resolution Blended Analyses for Sea
677 Surface Temperature. *Journal of Climate*, 20:5473–5496.
- 678 Rio, M.-H., Mulet, S., Greiner, E., Picot, N., and Pascual, A. (2013). New
679 global Mean Dynamic Topography from a GOCE geoid model, altimeter
680 measurements and oceanographic *in situ* data. OSTST 2013 Meeting,
681 Boulder CO. Available: [http://www.aviso.altimetry.fr/fileadmin/
682 documents/OSTST/2013/oral/mulet_MDT_CNES_CLS13.pdf](http://www.aviso.altimetry.fr/fileadmin/documents/OSTST/2013/oral/mulet_MDT_CNES_CLS13.pdf).
- 683 Rosen, D. A. S. and Trites, A. W. (2014). Thermal limits in young northern
684 fur seals, *Callorhinus ursinus*. *Marine Mammal Science*, 30(3):1014–1028.
- 685 Rubin, D. B. and Schenker, N. (1986). Multiple imputation for interval esti-
686 mation from simple random samples with ignorable nonresponse. *Journal*
687 *of the American Statistical Association*, 81(394):366–374.
- 688 Rutishauser, M. R., Costa, D. P., Goebel, M. E., and Williams, T. M. (2004).
689 Ecological Implications of Body Composition and thermal Capabilities in
690 Young Antarctic Fur Seals (*Arctocephalus gazella*). *Physiological and Bio-*
691 *chemical Zoology*, 77(4):669–681.
- 692 Scharf, H., Hooten, M. B., and Johnson, D. S. (2017). Imputation approaches
693 for animal movement modeling. *Journal of Agricultural Biological and*
694 *Environmental Statistics*, In press.
- 695 Sterling, J. T., Spring, A. M., Iverson, S. J., Johnson, S. P., Pelland, N. A.,
696 Johnson, D. S., Lea, M.-A., and Bond, N. A. (2014). The Sun, Moon,
697 Wind, and Biological Imperative—Shaping Contrasting Wintertime Migra-
698 tion and Foraging Strategies of Adult Male and Female Northern Fur Seals
699 (*Callorhinus ursinus*). *PLoS ONE*, 9(4):e93068.

- 700 Vallis, G. K. (2006). *Atmospheric and Oceanic Fluid Dynamics*. Cambridge
701 University Press.
- 702 Volinsky, C. T., Madigan, D., Raftery, A. E., and Kronmal, R. A. (1997).
703 Bayesian model averaging in proportional hazard models: assessing the
704 risk of a stroke. *Journal of the Royal Statistical Society: Series C (Applied*
705 *Statistics)*, 46(4):433–448.
- 706 White, I. R., Royston, P., and Wood, A. M. (2011). Multiple imputation
707 using chained equations: issues and guidance for practice. *Statistics in*
708 *medicine*, 30(4):377–399.

Table 1: Model set and population average posterior model probabilities. The model column list the models in the model set considered for this analysis. The base model is, in R notation, $\sim \log(\text{residence}) * \text{previous} + \text{B:north} + \text{B:east}$. The notation **B** indicates a radial basis matrix for allowing temporally varying coefficients. PMP column is the average PMP over all 15 animals. The range column is the range of PMP values over animals.

Model No.	Covariates	Avg. PMP	Range
1	base	0.003	[0.000–0.040]
2	base + B:wind	0.714	[0.028–1.000]
3	base + B:current	0.004	[0.000–0.053]
4	base + B:sst		
5	base + B:wind + B:current	0.275	[0.000–0.961]
6	base + B:wind + B:sst	0.004	[0.000–0.049]
7	base + B:current + B:sst		
8	base + B:wind + B:current + B:sst	0.001	[0.000–0.011]

Table 2: Posterior model probabilities (PMP) for the CTSMC movement analysis. All models included the ‘Base’ model for movement, where **base** = **log residence * previous + B:north + B:east**. Model numbers correspond to rows and associated models in Table 1.

ID	Model							
	1	2	3	4	5	6	7	8
330		0.14			0.86			
335	0.04	0.34	0.05		0.57			
349		1.00						
351		0.92			0.08			
355		0.99			0.01			
362		0.99			0.01			
367		0.77			0.23			
377		1.00						
380		0.50			0.45	0.05		
381		1.00						
385		1.00						
388		1.00						
394		1.00						
405		0.03			0.96			0.01
408		0.05			0.95

Figure 1: Imputed paths for analysis by the CTSMC model. Each color represents a different northern fur seal pup. Each imputation contains 20 paths for each individual.

Figure 2: Area of deployment for the northern fur seal 355. Blue points indicate the observed locations recorded from the telemetry device. Red lines indicate the paths used for the process imputation methodology. The hex grid are the discrete locations used for CTSMC movement analysis. These hexes are all the ones used for animal 355. Not all of these are used for each imputation run.

Figure 3: Time varying coefficient estimates for the CTSMC model fit to individual 355. Estimates and credible intervals were calculated using the process imputation methodology presented in Section 3.3 using a normal approximation to the posterior for each imputation.