

# A CONTINUOUS-TIME SEMI-MARKOV MODEL FOR ANIMAL MOVEMENT IN A DYNAMIC ENVIRONMENT

BY DEVIN JOHNSON<sup>\*</sup>, NOEL PELLAND<sup>†</sup> AND JEREMY STERLING<sup>‡</sup>

Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, <sup>\*</sup>[devin.johnson@noaa.gov](mailto:devin.johnson@noaa.gov);  
<sup>†</sup>[noel.pelland@noaa.gov](mailto:noel.pelland@noaa.gov); <sup>‡</sup>[jeremy.sterling@noaa.gov](mailto:jeremy.sterling@noaa.gov)

We consider an extension to discrete-space, continuous-time models for animal movement that have previously been presented in the literature. The extension from a continuous-time Markov formulation to a continuous-time *semi*-Markov formulation allows for the inclusion of temporally dynamic habitat conditions as well as temporally changing movement responses by animals to that environment. We show that, with only a little additional consideration, the Poisson likelihood calculation for the Markov version can still be used within the multiple imputation framework commonly employed for analysis of telemetry data. In addition, we consider a Bayesian model selection methodology within the imputation framework. The model selection method uses a Laplace approximation to the posterior model probability to provide a computationally feasible approach. The full methodology is then used to analyze movements of 15 weaned northern fur seal (*Callorhinus ursinus*) pups with respect to surface winds, geostrophic currents and sea surface temperature. The highest posterior model probabilities belonged to those models containing only winds and current; SST was not a significant factor for modeling their movement.

**1. Introduction.** Studying the movement of animals has become ubiquitous in the ecological literature in the past decade. One traditionally thorny issue for developing statistical models for telemetry data is relating spatial habitat conditions to movement processes. This difficulty arises because of the mismatch in support for the telemetry data (continuous spatial domain) and the observed habitat variables (usually a spatial grid or raster) and the often random times when location is observed. Because the habitat variables are the coarsest spatial resolution of the two types, it makes sense to bring the telemetry data to the level of the habitat data, so to speak. We propose a generalization of the spatially-discrete, continuous-time movement models of Hanks, Hooten and Alldredge (2015), so researchers can now relate animal movement to a dynamically changing environment.

In recent years there have been several methodologies based on translations of continuous-space telemetry data to a discrete space and modeling movement between discrete locations based on temporally static habitat variables (see Hooten et al. (2010); Hanks, Hooten and Alldredge (2015); Hanks and Hughes (2016), and Avgar et al. (2016)). In addition, Johnson, Hooten and Kuhn (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, Hooten and Alldredge (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 fitted 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,

---

Received May 2019; revised October 2020.

*Key words and phrases.* Animal telemetry, semi-Markov model, movement model, multiple imputation, northern fur seal.

Hooten and Alldredge (2015), Hanks and Hughes (2016)). 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). 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, Sterling and Loughlin (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 U.S. government officials during study of the Bering Sea northern fur seal population in the late 1800s (Hooper (1895)). These hunters were in agreement that seals always traveled with the wind and sea currents at their back pushing them forward. Widespread satellite tagging has provided the opportunity to confirm and quantify these effects; Lea et al. (2009) found differing departure and dispersal patterns of Pribilof Islands northern fur seal pups in years of contrasting atmospheric forcing and evidence of more rapid movement in the presence of a tailwind. Sterling et al. (2014) and Pelland et al. (2014) found statistical evidence that surface winds influence adult northern fur seals during migration, with high winds associated with linear, directed movement, though directionality in relation to the wind was not explored.

In this paper our objectives are to develop the CTSMC model such that we do not need to rely on the distributional assumptions of Hanks, Hooten and Alldredge (2015) with respect to cell residence time (Section 2). Because the habitat can change before the animal moves cells, we can not rely on a constant rate of movement while the animal remains in a cell. Following CTSMC development, in Section 3 we show that, with only slight modification, the two-stage imputation and GLM approach of Hanks, Hooten and Alldredge (2015) can still be used for parameter inference. In addition to model fitting, we develop methodology for Bayesian model selection using this two-stage method. In Section 4 we illustrate the full methodology for fitting and model selection using migration data from 15 northern fur seal pups tagged in Alaska, U.S.A.

**2. Continuous-time semi-Markov model for movement.** In a continuous-time, discrete-space movement model the geographical study domain is partitioned into a cells indexed by  $\mathcal{C} = \{1, \dots, n\}$ . Each cell,  $i$ , has a set of neighboring cells  $\mathcal{C}_i = \{j \in \mathcal{C} : i \sim j\}$ , where  $j \sim i$  means that cells  $i$  and  $j$  are directly connected and the animal can go from cell  $i$  to cell  $j$  in one move. For example, neighboring cells might share a border in a raster partition, but cells might also represent spatially separate patches. As with all semi-Markov models, the complete path of the animal,  $P$ , can be decomposed into the times at which cell transitions are made (jump times),  $\tau = \{\tau_0, \tau_1, \dots, \tau_M, \tau_{M+1}\}$ , and the sequence of cells visited (embedded Markov chain),  $\mathcal{G} = \{g_0, \dots, g_M\}$ , such that  $P = \{(\tau_m, g_m) : m = 0, \dots, M\}$ . We will assume that  $\tau_0 = 0$ ,  $\tau_M = T$  is the known end of the telemetry deployment, and  $g_0$  is the known starting cell. We will use  $P_t$  to denote the history of the path up to time  $t$ , that is,  $P_t = \{t, (\tau_0, g_0), \dots, (\tau_m, g_m) : \tau_m < t\}$ . In a similar style we also use the location and neighborhood notation  $g_t$  and  $\mathcal{C}_t$  (i.e., indexed by continuous time,  $t$ ) to represent the location and neighborhood of the animal for any time in  $[0, T]$ , that is,  $g_t = g_m$  and  $\mathcal{C}_t = \mathcal{C}_{g_m}$  for  $t \in [\tau_m, \tau_{m+1})$ . It is apparent which context is being used in a particular situation.

2.1. *General likelihood formulation.* The heart of the CTSMC model is the emigration rate function

$$(1) \quad \lambda_{ij}(t|P_t) = \lim_{h \rightarrow 0} [g_{t+h} = j|P_t]/h,$$

where we use the notation “[ $A|B$ ]” to represent the probability density (distribution) function of  $A$  given  $B$ . Therefore, [ $g_{t+h} = j|P_t$ ] 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

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

Using the total rate of emigration from a cell  $i$  at time  $t$ , we can obtain the distribution of  $\tau_m$ , given the animal moved to cell  $i$  at time  $\tau_{m-1}$ . 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

$$(3) \quad [\tau_m|P_{\tau_m}] = \Lambda_i(\tau_m|P_{\tau_m}) \exp\left\{-\int_{\tau_{m-1}}^{\tau_m} \Lambda_i(u|P_u) du\right\}.$$

If one conditions on the fact that the next move is at time  $\tau_m$ , then the cell, to which the move is made, is a categorical variable with distribution

$$(4) \quad [g_m = j|\tau_m, P_{\tau_m}] = \frac{\lambda_{ij}(\tau_m|P_{\tau_m})}{\Lambda_i(\tau_m|P_{\tau_m})}$$

(Norris (1998)). If we now assume that the emigration hazard function,  $\lambda_{ij}(t|P_t)$ , is also a function of a parameter vector, say  $\theta$ , then, given an observed path  $P$ , the likelihood is calculated as a product of all the conditional move and transition time probability density functions,

$$(5) \quad [P|\theta] = \left\{ \prod_{m=1}^K \lambda_{g_{m-1}, g_m}(\tau_m|P_{\tau_m}) \right\} \times \exp\left\{-\int_0^T \sum_{j \in \mathcal{C}_u} \lambda_{g_u, j}(u|P_u) du\right\}.$$

See Section A.1 of the Supplementary Material (Johnson, Pelland and Sterling (2021)) for more details on derivation of the likelihood function.

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 generalize the model form given by Hanks, Hooten and Alldredge (2015) and Hanks and Hughes (2016) to accommodate a time-varying structure. In the CTSMC movement context, for  $g_t = i$  and  $j \in \mathcal{C}_i$ , a proportional rates (hazards) model (Cox and Oakes (1984)) with a time-varying covariate is given by

$$(6) \quad \lambda_{ij}(t|P_t) = \exp\{b_0(t|P_t) + \mathbf{x}'_{ijt} \boldsymbol{\beta}_t\},$$

where  $b_0(t|P_t)$  a baseline log-hazard function,  $\mathbf{x}_{ijt}$  is a covariate vector associated with transition from cell  $i$  to  $j$  at time  $t$  and  $\boldsymbol{\beta}_t$  are time-indexed regression coefficients associated with the respective covariates. If a covariate is only indexed by  $i$  and  $t$ , then they are known as motility drivers. 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 (4). Whereas, the variables, which are indexed by  $j$  as well, influence rate of movement by drawing (repelling) animals to (from) other neighboring cells.

There are some notable differences between the model in (6) and that considered by [Hanks and Hughes \(2016\)](#). 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, hereafter,  $r_t$  (residence time in cell  $g_t$ ). This leads to clustered (or regular) transition times with respect to those that are exponentially distributed in the CTMC model. A more important difference is how temporal dynamics affect movement. In the CTMC model the rate function can change through time, but only when an animal transitions between cells (see [Hoover et al. \(2019\)](#)). In the CTSMC model the rate function can change *while* the animal remains in the cell, potentially triggering movement.

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, Hooten and Kuhn \(2013\)](#). The spatiotemporal point process likelihood is very similar to the CTSMC likelihood in (5), with the exception that the summation in the right-hand side is replaced with an integration over the continuous spatial domain of a neighboring region. 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 spatiotemporal point process of [Johnson, Hooten and Kuhn \(2013\)](#). 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 (5) is often challenging to evaluate in general. Under the CTMC assumption of a constant rate function while the animal remains in a cell, [Hanks, Hooten and Alldredge \(2015\)](#) show that practitioners can formulate the exact likelihood as a Poisson regression to obtain maximum likelihood estimates using standard GLM software. In the case of the CTSMC model, this is not possible to calculate the exact likelihood using this trick. However, an approximate likelihood can be calculated with this same method. Moreover, 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 \(1981\)](#) for temporally varying individual covariates. Following the same development, we can form an approximation for calculating the CTSMC movement model likelihood using a Poisson regression model.

To begin the approximation, we first select a set of quadrature points,  $q_n$ ,  $n = 0, \dots, N$ , to numerically evaluate the integral in (5). These quadrature points include times when cell transitions occurred,  $\tau = \{\tau_0, \dots, \tau_K\}$ , times where dynamic covariates changed values and any additional points that the practitioner feels necessary to approximate the integral in (4). For example, in Section 4 a 15-minute grid of times was also included in the quadrature set to model seal movement. For the augmented quadrature times we maintain the same end points,  $q_0 = \tau_0$  and  $q_N = T$ . Then, the approximation assumes  $\lambda(t|P_t) = \lambda(q_{n-1}|P_{q_{n-1}})$  for  $t \in (q_{n-1}, q_n]$ . To ease some of the notational burden and eliminate excess subscripts, we further set

$$(7) \quad \lambda_{nj} = \lambda_{g_{q_{n-1}},j}(q_{n-1}|P_{q_{n-1}}) \quad \text{for } n = 1, \dots, N \text{ and } j \text{ in } \mathcal{C}_{q_{n-1}}.$$

Now, the integral is approximated with the summation

$$(8) \quad \int_0^T \sum_{j \in \mathcal{C}_u} \lambda_{g_u,j}(u|P_u) \approx \sum_{n=1}^N \sum_{j \in \mathcal{C}_{q_{n-1}}} \delta_n \lambda_{nj},$$

where  $\delta_n = q_n - q_{n-1}$ . After placing the approximation (8) into (4), one obtains the approximate likelihood function (up to a proportional constant),

$$(9) \quad [P|\theta] \approx \prod_{n=1}^N \prod_{j \in \mathcal{C}_{q_{n-1}}} \tilde{\lambda}_{nj}^{z_{nj}} \exp\{-\tilde{\lambda}_{nj}\},$$

where  $\tilde{\lambda}_{nj} = \delta_n \lambda_{nj}$  and  $z_{nj} = 1$  if  $q_n$  is a transition time and  $j$  is the cell transitioned to at time  $q_n$ , else  $z_{nj} = 0$ . The likelihood function formed by (9) is proportional to a Poisson likelihood function where  $z_{nj}$  are the “data” and  $\tilde{\lambda}_{nj}$  are the rate parameters. The only difference between this formulation and that of Hanks, Hooten and Alldredge (2015) is that there are times,  $q_n$  where  $z_{nj} = 0$  for all  $j$ , that is, there was no movement to a cell, but the  $\lambda_{ij}$  have changed. This likelihood approximation can be made as accurate as desired by simply increasing the temporal resolution of the quadrature points. Or, it will be exact if the rate function actually is constant between quadrature points, for example, covariates only change values on a temporal grid of times.

**3.2. GLM/GAM and the proportional rate model.** Using the Poisson approximation (9), one can gain powerful computational assistance if the proportional hazards model (6) is used. Because of the linear structure on the log scale, standard GLM software can be used to fit the CTSMC movement model. On the log scale, the Poisson rate used for the  $z_{nj}$  data is

$$(10) \quad \log \tilde{\lambda}_{nj} = \log \delta_n + b_0(r_{q_n} | P_{q_n}) + \mathbf{x}'_{g_{q_n}, j, q_n} \boldsymbol{\beta}_{q_n},$$

where  $\log \delta_n$  takes the form of a known offset. Note that we have now assumed that the baseline rate function is parameterized by the residence time. Assuming  $b_0$  is constant (we will discuss this momentarily), the parameters are all linear on the log scale, so any GLM fitting software can be used. One may also estimate time-varying coefficients using a “varying coefficients” model, as illustrated in the present analysis of northern fur seal migration in Section 4.

We now examine the baseline rate function,  $b_0$ . To begin, if one models  $b_0(r_{q_n}) = \beta_0 + \beta_1 \log r_{q_n}$ , then the proportional rate model is still in a log-linear form, so standard GLM software can be used. This log-linear model for  $b_0$  implies the waiting times between cell transitions follow a Weibull distribution when covariates are not considered (Cox and Oakes (1984)). If  $\beta_1 < 0$ , then the rate function decreases with time since last transition, and the times of transition will tend to be clustered in time (Hooten et al. (2017)). The reverse is true for  $\beta_1 > 0$ ; transitions will occur at more regular intervals. For  $\beta_1 = 0$  the residence times will be exponentially distributed, as in the CTMC model (Hanks and Hughes (2016)). If  $r_{q_n}$  is used instead of  $\log r_{q_n}$ , the residence times will be distributed according to the Gompertz-Makeham distribution (Cox and Oakes (1984)). However, if neither of these simple models are useful, it is straightforward to model  $b_0$  nonparametrically using a GAM to directly estimate  $b_0$  treating  $r_{q_n}$  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  $P$  as data. This may be the case in the future as satellite telemetry devices improve, for example; see Liu et al. (2016) 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, Hooten and Johnson (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 the used by Hanks, Hooten and Alldredge (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,  $\mu(t)$ . If  $\mu(t)$  were observed for all  $t$  ( $\mu$  from here on), then we could summarize it into the data we desire, namely,  $P = P(\mu)$ . However, locations are usually only observed sporadically in time with error, say,  $\mathbf{y} = \{\mathbf{y}_{\tau_1^*}, \dots, \mathbf{y}_{\tau_n^*}\}$ , where  $\tau^*$  is an observation time in  $[0, T]$ . To obtain the correct Bayesian posterior distribution for the movement parameters,  $\theta$ , the posterior distribution must be marginalized over the missing  $\mu$  process, that is,

$$(11) \quad [\theta|\mathbf{y}] = \int [\theta|\mu][\mu|\mathbf{y}] d\mu \propto \int \frac{[P(\mu)|\theta][\theta]}{[P(\mu)]} [\mu|\mathbf{y}] d\mu.$$

In the movement modeling context, however, the true distribution,  $[\mu|\mathbf{y}]$  and thus  $[\theta|\mathbf{y}]$ , can virtually never be evaluated or even sampled from (Scharf, Hooten and Johnson (2017)). Therefore, in practice, a two-step procedure is used to describe  $[\theta|\mathbf{y}]$  by first selecting an approximation to  $\mu$ , say  $\mu^*$ , that is sufficiently similar such that samples from  $[\mu^*|\mathbf{y}]$  are similar to samples from  $[\mu|\mathbf{y}]$ . The approximation  $[\mu^*|\mathbf{y}]$  is known as the *imputation distribution*. Then, a sample is drawn from the imputation distribution  $\mu_k \sim [\mu^*|\mathbf{y}]$ ,  $k = 1, \dots, K$ , computing the discrete space path from the imputed data,  $P^{(k)} = P(\mu_k)$ , then summarizing  $[\theta|P^{(k)}]$  in some way, depending on the inference desired. From hereafter, we will just use  $[\mu|\mathbf{y}]$  to denote the imputation distribution while recognizing this is not the exact distribution. In addition, we also note that, because  $P(\mu)$  is a deterministic function of  $\mu$ ,  $[\theta|P(\mu)] = [\theta|\mu]$ .

With imputation approaches in general, calculating interval estimates is not always straightforward (Rubin and Schenker (1986)), as there are many ways one may combine first-stage analyses to obtain unconditional intervals. We propose an 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  $[\theta|\mu_k]$  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 with any posterior sample. If GLM or GAM software is used to fit the model, one could easily approximate  $[\theta|\mu_k] \approx N(\hat{\theta}_k, \hat{\Sigma}_k)$ , the large sample MLE or posterior distribution (see Section 4.2).

Another aspect of movement model inference using process imputation that has received little attention is that of model selection. 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 begin formulating Bayesian model selection, we suppose that there is a set of  $L$  models,  $\mathcal{M} = \{M_1, \dots, M_L\}$  for the movement rate,  $\lambda_{ij}(t|P_t, M_l)$  from which the likelihoods,  $[P|\theta_l, M_l]$ , are formed. The posterior model probabilities are given by

$$(12) \quad [M_l|\mathbf{y}] = \int [M_l|P][\mu|\mathbf{y}] d\mu,$$

where

$$(13) \quad [M_l|P] \propto [P|M_l][M_l] = \left\{ \int [P|\theta_l, M_l][\theta_l|M_l] d\theta_p \right\} [M_l],$$

$[\theta_l|M_l]$  is the parameter prior distribution under model  $M_l$  and  $[M_l]$  is the prior model probability. Bayesian model selection has a large volume of literature devoted to its methodology. However, for application within the imputation framework, we seek a method to evaluate

$[P|M_l]$  that is computationally efficient, even if it is approximate, as the whole model selection procedure needs to be replicated for each  $\mu_k$ . Therefore, we propose using the Laplace approximation (Kass and Raftery (1995)) to equation (13),

$$(14) \quad \begin{aligned} \log[P^{(k)}|M_l] &\approx B(M_l|P^{(k)}) \\ &= \log[\hat{\theta}_{kl}|P^{(k)}, M_l] + \log[\hat{\theta}_{kl}|M_l] \\ &\quad - \log|\hat{\mathbf{H}}_{kl}|/2 + d_l \log(2\pi)/2, \end{aligned}$$

where  $\hat{\theta}_{kl}$  is the posterior mode of  $[\theta_l|P^{(k)}, M_l]$ ,  $\hat{\mathbf{H}}_{kl}$  is the Hessian matrix of the negative log posterior density and  $d_l$  is the dimension of  $\theta_l$ . Kass and Raftery (1995) also note that the approximation is still of the same order if the MLE of  $\theta$  is used, provided the prior exerts little influence on the posterior relative to the likelihood. The benefit of that is one can use standard GLM software to fit the model without worrying about the prior distribution. However, proper priors are still necessary (highly advised) for terms that are being selected over (Bayarri et al. (2012)).

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

$$(15) \quad [M_l|y] \approx \text{mean}\left(\frac{\exp\{B(M_l|P^{(k)})\}[M_l]}{\sum_{l'} \exp\{B(M_{l'}|P^{(k)})\}[M_{l'}]}\right),$$

where the mean is over the  $K$  imputations. The Laplace approximation allows us to have a coherent approach to averaging an information criterion within the imputation framework.

**4. Northern fur seal pup migration.** To demonstrate analysis of movement data, we analyzed telemetry locations from 15 female northern fur seal pups tagged in the Pribilof Is., Alaska, prior to their annual migration. Female “355” from this data set was previously analyzed by Johnson et al. (2008) using a continuous time random walk. We will highlight its results here; results for the remaining animals can be viewed in Section A.2 of the Supplementary Material. Source code and data for reproducing this analysis are available in the Supplementary Material (Sections B and C; Johnson, Pelland and Sterling (2021)). These data were analyzed using the R Statistical Environment (R Core Team (2019)) due to the availability of all necessary spatial, data manipulation, movement modeling; and GLM fitting functions; however, nothing precludes development in other software.

**4.1. A dynamic environment.** Environmental covariates were defined on a hexagonal spatial grid covering the domain of the pup tracks that departed during fall 2005. For the process imputation, 20 tracks were generated using draws from the posterior distribution of locations every 15 minutes from a continuous-time correlated random walk model with a correlated random drift process (Johnson, London and Kuhn (2011)) using the R package `crawl` (Johnson and London (2018)). A regular hexagonal lattice, with 40 km center point lateral separation, and associated hexagonal polygons (“cells”) were generated. Figure 1 illustrates the imputed tracks and hex grid for animal 355. Notice that only hex cells in the neighborhood around the tracks are necessary for fitting the CTSMC model. So, we only need the habitat covariates for those cells. Model covariates were projected onto the hexagonal grid using environmental variables obtained from remote sensing observations or atmospheric re-analysis.

Estimates of geostrophic ocean surface currents—slowly varying oceanic motions (cm/s) forced by mesoscale ( $\sim 10$ – $200$  km) gradients of sea surface height (Vallis (2006))—were obtained from the AVISO All-Satellite Absolute Dynamic Topography (ADT) product

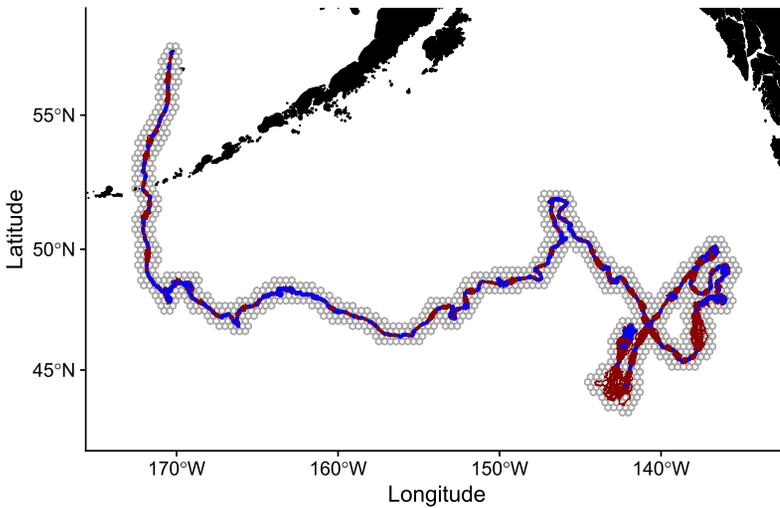


FIG. 1. Area of deployment for the northern fur seal 355. Blue points indicate the observed locations recorded from the telemetry device. Dark red lines indicate the paths used for the process imputation methodology. The hex grid are the discrete locations used for CTSMC movement analysis. These hexes are all the ones used for animal 355. Not all of these are used for each imputation run. For this animal, and all the others, the tracks begin in the Pribilof Islands, Alaska in the northwest portion of the map.

(<https://www.aviso.altimetry.fr/en/home.html>). At each grid point,  $s$ , the surface geostrophic current  $\mathbf{x}_{\text{curr}}(s, t) = (u_{\text{curr}}(s, t), v_{\text{curr}}(s, t))'$  has a component  $u_{\text{curr}}$  in the east direction along unit vector and  $v_{\text{curr}}$  in the north direction. Current measurements varied daily at midnight UTC.

Atmospheric surface (10 m height) wind vector estimates  $\mathbf{x}_{\text{wind}}(s, t)$  (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 six-hour resolution.

Sea surface temperature ( $^{\circ}\text{C}$ ) estimates,  $x_{\text{sst}}(s, t)$ , 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>). This product varies daily at midnight UTC as well.

Each of the environmental variables described above was spatially interpolated to the hexagonal grid cells at six-hourly intervals. If a cell overlapped a raster grid border, the weighted average value was used based the areas contained in the hex. The six-hour interval corresponds to the shortest resolved interval among the available variables (surface wind); 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,

$$(16) \quad \text{wind}_{ijt} = \mathbf{x}_{\text{wind}}(s_i, t)' \mathbf{w}_{ij} \quad \text{and} \quad \text{curr}_{ijt} = \mathbf{x}_{\text{curr}}(s_i, t)' \mathbf{w}_{ij},$$

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, Hooten and Alldredge (2015)). SST was treated as a motility covariate to test whether pups might adjust speed if they encounter favorable conditions with respect to SST,  $\text{sst}_{it} = x_{\text{sst}}(s_i, t)$ .

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:

- $\log r_t$  (time since last transition),
- $\text{previous}_{ijt} = \mathbf{w}_t' \mathbf{w}_{ij}$ ,

- $\text{north}_{ij} = (0, 1)' \mathbf{w}_{ij}$ ,
- $\text{east}_{ij} = (1, 0)' \mathbf{w}_{ij}$ ,

where  $\mathbf{w}_t$  is the unit vector of the last transition prior to  $t$ . north and east were included to account for pup migration tendency. The covariate  $\text{previous}_{ijt}$  was added to reflect correlated movement on a finer scale than general migration patterns and is the direction of the previous cell transition. It is typical of animal movement that individuals tend to move in the same direction for some time in a correlated fashion.

In addition to the cell transition times and covariate changing times, locations on each imputed path,  $\boldsymbol{\mu}_k$ , were generated at 15-minute intervals to approximate the likelihood integral. Thus, for each quadrature time,  $q_n, n = 0, \dots, N$ , on each imputed path, a six row data set ( $j = 1 \dots, 6$ ) was created where the  $j$ th row contains:  $z_{nj}, \delta_n, q_n, r_{q_n}, \text{previous}_{g_{q_n}, j, q_n}, \text{north}_{g_{q_n}, j}, \text{east}_{g_{q_n}, j}, \text{sst}_{g_{q_n}, q_n}, \text{wind}_{g_{q_n}, j, q_n}$  and  $\text{curr}_{g_{q_n}, j, q_n}$ . The data set was then concatenated over all  $q_n$  within an imputation; this forms the model data for the  $k$ th imputation. From this data all models can be fitted using a GLM (or GAM) using a standard Poisson family with log-link function.

4.2. *Model formulation, estimation, and selection.* For the analysis of the fur seal migration data, we will use the log-linear form represented in equations (6) and (10).

Further, we break the model into two pieces for model section purposes. The ‘‘Base’’ portion of the model contains all the kinetic variables described in the previous section, and the ‘‘Habitat’’ portion contains all the environmental variables, that is,

$$\log \lambda_{ij}(t|P_t) = \mathbf{x}'_{\text{Base},ijt} \boldsymbol{\beta}_{\text{Base},t} + \mathbf{x}'_{\text{Habitat},ijt} \boldsymbol{\beta}_{\text{Habitat},t},$$

where  $\mathbf{x}_{\text{Base},ijt}$  includes:  $\log r_t, \text{previous}_{ijt}, \text{north}_{ij}, \text{east}_{ij}$ , plus an intercept and interaction term  $\log r_t \times \text{previous}_{ijt}$  and  $\mathbf{x}_{\text{Habitat},ijt}$  includes:  $\text{wind}_{ijt}, \text{curr}_{ijt}$  and  $\text{sst}_{it}$ . In addition to  $\log r_t$ , the base rate function,  $b_0(t|P_t)$ , in equation (6) is modeled with a  $\log r_t \times \text{previous}$  interaction effect to account for the synergistic effect of kinetic movement on the base residence time within a cell. That is, if there is a strong relationship between the previous movement and the current one, then the residence time is likely to be longer because the animal is probably transiting the entire cell. However, if the animal is going back and forth across a border, residence time will likely be short and relationship to the previous movement will likely be negative. Because of the length of deployment, it is natural to inquire whether an individual’s movement in response to the environment changes over the course of time. The  $\text{previous}$  and  $\log r_t$  terms were modeled as constant due to the fact that initial investigation showed no improvement by allowing their effect to vary over time. All other model terms were modeled as time varying. To accomplish this, we used a radial basis function model for the coefficients. That is, for a general covariate  $v$ , the coefficient is modeled as

$$(17) \quad \beta_{vt} = \mathbf{b}'_t \boldsymbol{\theta}_v,$$

where  $\mathbf{b}_t = (\exp\{(t - \kappa_1)^2/\phi^2\}, \dots, \exp\{(t - \kappa_M)^2/\phi^2\})'$  is a vector of radial basis functions,  $\kappa_m$  is a knot in  $[0, T]$ ,  $\boldsymbol{\theta}_v$  is a vector of basis weights and  $\phi$  is a basis scaling parameter. We chose three equally spaced interior knots plus two additional end knots at equal spacing before and after  $[0, T]$  for a five basis model. The scalar  $\phi$  set to the knot spacing interval. If the covariate is not modeled with a time-varying coefficient, simply set  $\mathbf{b}_t \equiv 1$ . In practice, one combines the basis vectors and covariates by  $\mathbf{x}'_{ijt} \otimes \mathbf{b}'_t$  to form the rows of the full design matrix  $\mathbf{X}$  such that in vector form notation  $\log \boldsymbol{\lambda} = \mathbf{X}\boldsymbol{\theta}$ , where  $\boldsymbol{\theta}$  is all the  $\boldsymbol{\theta}_v$  concatenated into a single vector.

For each model,  $l$  and imputation,  $k$ , we used the  $g$ -prior distribution

$$(18) \quad [\boldsymbol{\theta}_{kl}|M_l] = N(\mathbf{0}, (\xi_{kl} \mathbf{X}'_{kl} \mathbf{X}_{kl})^{-1}),$$

where  $\mathbf{X}_{kl}$  is the design matrix for imputation  $k$  and model  $l$  and  $\xi_{kl}$  is a scalar. A flat prior distribution was used for  $\xi_{kl}$  in each model.

The R package `mgcv` (Wood (2011)) and the `gam` function was used to estimate the posterior mode  $\hat{\theta}$  and to obtain the Hessian matrix  $\hat{\mathbf{H}}$  for the Laplace approximation in the posterior model probability (PMP) calculation and approximating the posterior

$$(19) \quad [\theta_{kl} | P^{(k)}, M_l] \approx N(\hat{\theta}_{kl}, -\hat{\mathbf{H}}_{kl}^{-1})$$

for effect inference. GAMs have a Bayesian interpretation in which the traditional penalty can be thought of as a Gaussian prior distribution over the coefficients (Miller (2019)). The `mgcv::gam` function allows the user to provide their own penalty, which we use the  $g$ -prior precision matrix,  $\mathbf{X}'_{kl}\mathbf{X}_{kl}$  and  $\xi_{kl}^{-1}$  is estimated as the GAM penalty parameter. To increase computational efficiency, for each animal fits were completed in parallel over the 20 imputed paths. This was accomplished with use of the R packages: `future` (Bengtsson (2018)), `doFuture` (Bengtsson (2017)), and `foreach` (Microsoft and Weston (2017)).

**4.3. Results.** There was strong evidence for an influence of surface winds on movement with an average posterior probability of 0.714 (over animals; Table 1) for the model with only wind and 0.275 for the model with wind and current effects. When broken down to individual animals the PMP of the wind only model was close or equal to one for 10 of 15 animals. Of those animals in which the wind-only model was not the maximum a posteriori model, the model with wind and current effects was the maximum PMP model for three of the remaining five. Finally, the remaining two animals were basically split between models with wind only and wind plus current effects. Table A.1 of the Supplementary Material provides full PMP results for all animals. Thus, at a population level, surface winds are a primary driver of movement for fur seal pups. There was only one animal that did not have *all* PMP mass on models with wind effects included and that animal had only 0.09 PMP mass on nonwind effect models. For some pups in addition to wind effects, geostrophic current plays a secondary driver of movement, but this was not the case for all animals. The pups may use it opportunistically, that is, if the current is going where they want to go, they may use it to their advantage. Finally, sea surface temperature does not significantly affect the residency time for pups at a given location. None of the pups had more than 0.05 PMP mass on models containing SST effects, and only two pups had any mass on SST containing models. Therefore, pups do not select warmer SST in the range of temperatures they experienced during migration.

We will now take a look at the temporally varying effects of animal “355.” Here, we will examine the effects for model 8 (Table 1; full model) even though model 2 was the highest

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 contains effects for:  $\log r_t$ ,  $\text{previous}_{ijt}$ ,  $\text{north}_{ijt}$ ,  $\text{east}_{ijt}$ , plus an intercept and interaction term  $\log r_t \times \text{previous}_{ijt}$ . 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 + $\text{wind}_{ijt}$	0.714	[0.028–1.000]
3	base + $\text{curr}_{ijt}$	0.004	[0.000–0.053]
4	base + $\text{sst}_{it}$		
5	base + $\text{wind}_{ijt}$ + $\text{curr}_{ijt}$	0.275	[0.000–0.961]
6	base + $\text{wind}_{ijt}$ + $\text{sst}_{it}$	0.004	[0.000–0.049]
7	base + $\text{curr}_{ijt}$ + $\text{sst}_{it}$		
8	base + $\text{wind}_{ijt}$ + $\text{curr}_{ijt}$ + $\text{sst}_{it}$	0.001	[0.000–0.011]

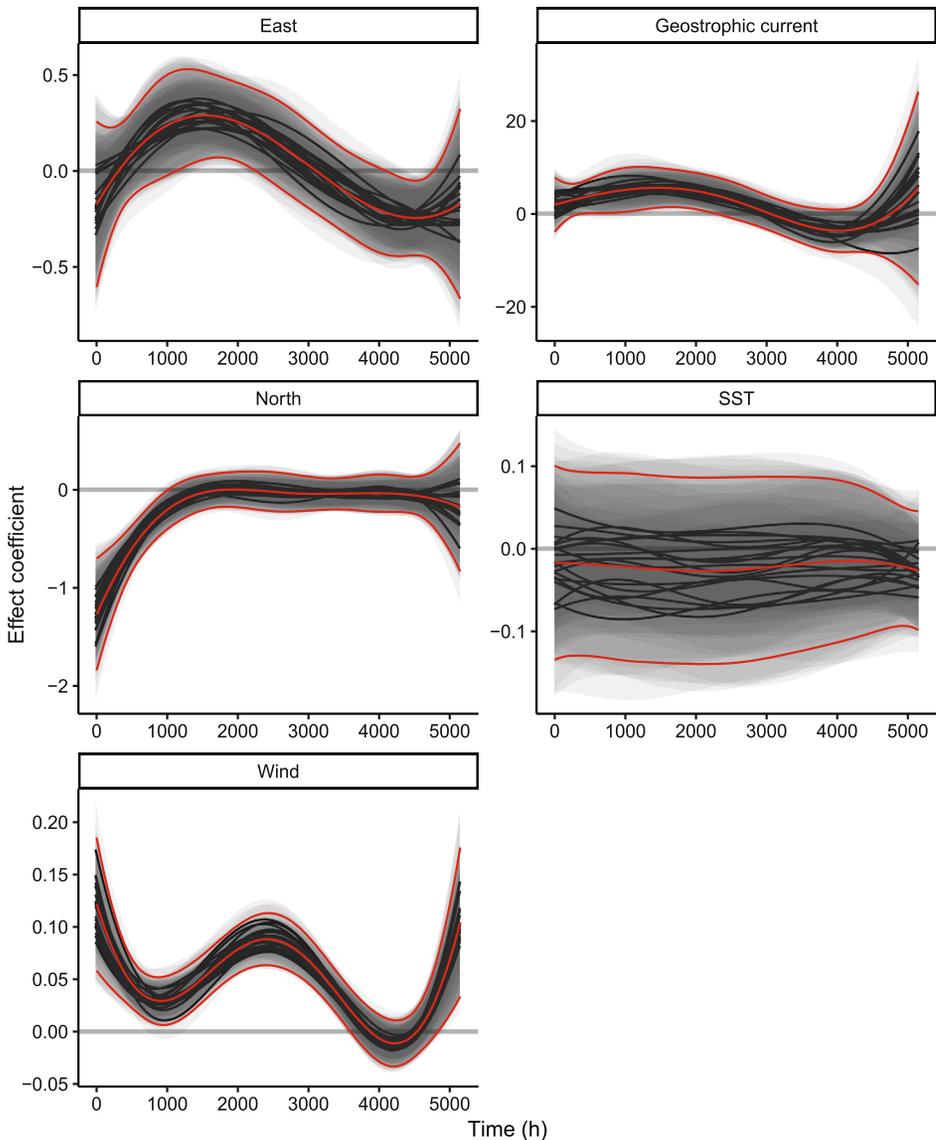


FIG. 2. Time varying coefficient estimates for the CTSMC model fit to individual 355. Estimates and credible intervals were calculated using the process imputation methodology presented in Section 3.3 using a normal approximation to the posterior for each imputation. The black lines represent the effect estimates for each of the 20 imputations. The transparent gray envelopes are the 95% credible intervals for each imputation. The red lines represent the estimate and credible interval when marginalizing over the imputations.

PMP model for this animal (Model 2: 0.99 and model 5: 0.01). However, it is instructive to see the general pattern of all covariates (Figure 2). For animal 355 the wind effect is overall positive throughout the course of the deployment implying that the animal generally travels with the prevailing winds, although the strength of this effect varies with time. There is also a generally positive effect of geostrophic current on the rate of movement, although for only a small fraction of the deployment is it significantly positive according to the 95% credible interval. This small interval is not enough to make the covariate overall significant, as the wind only model has  $\text{PMP} = 0.99$ . SST has a negative but insignificant effect meaning that the animal slows in warmer seas, but the conclusion is highly uncertain. The kinetic nuisance effects, north and east are also illustrated and describe the general trend in movement, as expected—first, southerly movement followed by generally easterly movement. Closer ex-

amination of Figure 2 shows the variation in covariate effect inference for each of the 20 path imputations,  $P^{(k)}$ . Use of any single imputation would have given us, heuristically, the same inference with respect to the effect of each covariate on movement. But there is noticeable variation in estimates caused by path uncertainty.

We now turn our attention to the base rate of transition,  $b_0(t|P_t)$ . We wanted to assess whether a constant base rate was appropriate for this data, that is, without covariate effects is the residence time approximately exponential distributed. Here, we examine the full model (model 8) again to see, if we account for all covariates, do the data support a constant base rate function,  $b_0(t|P_t) = \beta_0$ . For this analysis the fitted base rate model is

$$b_0(t|P_t) = \beta_0 + (\beta_1 + \beta_3 \text{previous}_{ij}) \cdot \log r_t.$$

For residence times to be exponential distributed,  $\tilde{\beta}_{1,ij} = (\beta_1 + \beta_3 \text{previous}_{ij}) \approx 0$  for a given  $\text{previous}_{ij}$ . For animal 355, when traveling in the same direction as the previous move ( $\text{previous}_{ij} = 1$ ),  $\tilde{\beta}_{1,ij} = 1.49$  [95% CI: 1.28, 1.71], significantly  $> 0$ . This implies that transitions are significantly more regular than would be expected for an exponential residence time. When moving in the opposite direction,  $\tilde{\beta}_{1,ij} = -1.00$  [95% CI:  $-1.07, -0.94$ ], significantly  $< 0$ , which means that transitions will be more clustered than an exponential distribution. Diagonal moves are slightly moderated toward zero, but all are significantly different from zero in the respective relative directions. Table A.2 in Supplement A shows that this same pattern holds for all 15 animals. These results make sense because an animal traveling in the same direction for multiple moves is probably traveling at relatively the same speed as it is transitioning. Thus, it would transition out of cells at a more regular schedule. The opposite is true for reversing direction. Speed and direction are more variable, indicative of an area restricted search pattern. Therefore, a constant exponential base rate is inappropriate for these data.

In addition to the CTSMC analysis here, for the focal animal 355 we also fitted CTMC version of the models. This was done to examine the level of improvement that is obtained by dropping the Markov movements assumption. To accomplish this, the covariates were altered such that the conditions the animal experiences upon first entering a cell are maintained till the animal leaves the cell, regardless of how it changes in the meantime. Also,  $\log r_t$  is not part of the base model, as this is not part of the CTMC model. Finally, the time index is also held constant when the animal first enters a cell for the basis function modeling of the coefficients. The results of this extra analysis are presented in Section A.3 of Supplementary Material A. In summary, the PMP of the CTMC models were virtually zero compared to the CTSMC models.

**5. Discussion.** Herein, we have presented an extension to the CTMC model of Hanks, Hooten and Alldredge (2015) for animal movement modeling. In addition to the ability to handle dynamic habitat covariates, the semi-Markov version (CTSMC) relaxes the assumption of exponentially distributed residence times for each cell. However, even after the extensions, the Poisson GLM method for likelihood calculation can still be used; however, it is now an approximation of the true likelihood. The only difference is that for quadrature times when no movement is made, all  $z_{nj} = 0$ . It can be thought of as the absence of movement at that time. An important distinction we would like to emphasize, is that while changes in rate parameters and covariates are permissible in CTMC model over the course of deployment, they are only permissible at times of cell transitions. As we have done in Supplementary Material A.3, one can approximate the CTSMC model with a CTMC model by only allowing changes at cell transitions, but the degree of approximation of the continuous-time dynamics is controlled by the animal and how often it moves. On the other hand, the level of approximation in the CTSMC model is entirely under the control of the researcher. The level of

approximation can be made as small as desired by simply creating a finer scale quadrature time grid. In addition, the CTSMC models can also capture clustered (or regular) residence time distributions.

In the fur seal analysis the CTSMC model proved to be a better fit to the data. For all 15 animals the parameters associated with the base rate function of  $\log r_t$  proved to be significantly different from zero, implying clustered or regular times of transition depending on the movement of the animal. For animal 355 none of the CTMC approximations to the CTSMC model possessed any posterior model weight. This may be due to the more accurate representation of the covariate effects in time or the nonexponential residence times of the CTSMC model. However, either one illustrates the additional benefits obtainable without increased effort on the part of the researcher for fitting models.

To fully account for path uncertainty in the analysis, we used a very similar two-stage imputation approach used by [Hanks, Hooten and Alldredge \(2015\)](#). For the data from animal 355, we would not come to different ecological conclusions using only one of these imputations, but there is nonzero variation due to complete path uncertainty that should be accommodated for more accurate effect inference. Because of the large spatial extent of this analysis, combined with the amount of data collected relative to the cell size, our results changed little, but for higher resolution spatial covariates and courser scale telemetry observations there may be more notable differences between imputations. An alternative to the imputation method used here is to use a “stacked” likelihood (e.g., [Hanks and Hughes \(2016\)](#), [White, Royston and Wood \(2011\)](#)). One can make inference using this weighted likelihood and avoid having to recombine separate model inferences. However, the large model data sets built using the methodology of Section 4.1 might preclude this, as the complete data set might become too large to deal with efficiently. 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 can be parallelized for increased computational speed, as we have done for the analysis in Section 4. Similarly, the two-stage method of [Hooten et al. \(2016\)](#) and [Hooten, Johnson and Brost \(2019\)](#) can be used to recombine separately fitted models for each animal if population-level inference is desired.

In addition to extension of the CTMC model to a semi-Markov version, we have also explored model selection in the context of a two-stage imputation approach. This method is not only applicable to this analysis but also to any multiple imputation studies. 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  $[M_l|P^{(k)}]$ . 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  $< 5d_l$  are worrisome, while those  $> 20d_l$  should be sufficiently accurate. [Volinsky et al. \(1997\)](#) note that the appropriate way to quantify sample size for proportional hazard models is the number of observed events in the data. For CTSMC models this would be the number of observed cell transitions  $\sum z_{nj}$  for each  $P^{(k)}$ . In the fur seal example the largest model (model 8) contains  $d_l = 30$  parameters, thus sample sizes  $> 600$  should cause little worry for any of the models fitted here. For the fur seal data, only one animal fell below this threshold, and its minimum was 560 transitions over all 20 imputed paths. So, we believe that sample size was sufficient in all animals for this analysis. However, this is something to keep in mind for other data.

The CTSMC model presents a computationally feasible approach for making inference with respect to the influence of dynamically changing habitat on animal movement. Based on the Markov (CTMC) analysis of animal 355, we would not have a drastic difference in the overall effect of wind on movement, but here we have a long deployment with undoubtedly spatially correlated observational covariates. Other areas of movement analysis, where these models might be very useful, is the analysis of dose-response in marine mammal disturbance studies (Booth, Sinclair and Harwood (2020)). In that instance, timing of changes in directional movement rates immediately following a known noise disturbance are the quantity of interest.

From a biological perspective, the ratio  $\lambda_{ij}/\Lambda_i$  represents a discrete choice resource selection model (McCracken, Manly and Vander Heyden (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 authors would like to thank M. Hooten and E. Hanks for review and discussion of early versions of the manuscript. 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.

## SUPPLEMENTARY MATERIAL

**Supplement A: Additional details and results** (DOI: [10.1214/20-AOAS1408SUPPA](https://doi.org/10.1214/20-AOAS1408SUPPA); .pdf). This supplement provides additional details of the likelihood derivation and Poisson approximation as well as additional results in the fur seal analysis.

**Supplement B: R code** (DOI: [10.1214/20-AOAS1408SUPPB](https://doi.org/10.1214/20-AOAS1408SUPPB); .zip). This supplement provides all R code used in the fur seal migration analysis.

**Supplement C: Data** (DOI: [10.1214/20-AOAS1408SUPPC](https://doi.org/10.1214/20-AOAS1408SUPPC); .zip). This supplement provides all data used in the fur seal migration analysis.

## 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 Ecol. Evol.* **7** 619–630.
- BAYARRI, M. J., BERGER, J. O., FORTE, A. and GARCÍA-DONATO, G. (2012). Criteria for Bayesian model choice with application to variable selection. *Ann. Statist.* **40** 1550–1577. [MR3015035 https://doi.org/10.1214/12-AOS1013](https://doi.org/10.1214/12-AOS1013)
- BENGTSSON, H. (2017). doFuture: A universal foreach parallel adaptor using the future API of the ‘future’ package. R package version 0.6.0.
- BENGTSSON, H. (2018). future: Unified parallel and distributed processing in R for everyone. R package version 1.8.1.
- BOOTH, C. G., SINCLAIR, R. R. and HARWOOD, J. (2020). Methods for monitoring for the population consequences of disturbance in marine mammals: A review. *Front. Mar. Sci.* **7** 115.
- BUDERMAN, F. E., HOOTEN, M. B., ALLDREDGE, M., HANKS, E. M. and IVAN, J. S. (2018). Predatory behavior is primary predictor of movement of wildland-urban cougars. [bioRxiv 257295](https://doi.org/10.1101/257295).
- COX, D. R. and OAKES, D. (1984). *Analysis of Survival Data* **21**. CRC Press.
- HANKS, E. M., HOOTEN, M. B. and ALLDREDGE, M. W. (2015). Continuous-time discrete-space models for animal movement. *Ann. Appl. Stat.* **9** 145–165. [MR3341111 https://doi.org/10.1214/14-AOAS803](https://doi.org/10.1214/14-AOAS803)
- HANKS, E. M. and HUGHES, D. A. (2016). Flexible discrete space models of animal movement. [arXiv:1606.07986](https://arxiv.org/abs/1606.07986).
- HOLFORD, T. R. (1980). The analysis of rates and of survivorship using log-linear models. *Biometrics* **36** 299–305.

- HOOPER, C. L. (1895). Report of Captain Hooper, dated November 21, 1892. In *US Senate, 53rd Congress, 2nd Session. Fur-seal Arbitration, Proceedings of the Tribunal of Arbitration, Volume VII, Senate Executive Document 177 (serial Set 3166)* 228–233. Government Printing Office, Washington, D.C.
- HOOTEN, M. B., JOHNSON, D. S. and BROST, B. M. (2019). Making recursive Bayesian inference accessible. *Amer. Statist.* <https://doi.org/10.1080/00031305.2019.1665584>
- HOOTEN, M. B., JOHNSON, D. S., HANKS, E. M. and LOWRY, J. H. (2010). Agent-based inference for animal movement and selection. *J. Agric. Biol. Environ. Stat.* **15** 523–538. MR2788638 <https://doi.org/10.1007/s13253-010-0038-2>
- HOOTEN, M. B., BUDERMAN, F. E., BROST, B. M., HANKS, E. M. and IVAN, J. S. (2016). Hierarchical animal movement models for population-level inference. *Environmetrics* **27** 322–333. MR3545640 <https://doi.org/10.1002/env.2402>
- HOOTEN, M. B., JOHNSON, D. S., MCCLINTOCK, B. T. and MORALES, J. M. (2017). *Animal Movement: Statistical Models for Telemetry Data*. Chapman and Hall/CRC, Boca Raton, FL.
- HOOVER, A. L., LIANG, D., ALFARO-SHIGUETO, J., MANGEL, J. C., MILLER, P. I., MORREALE, S. J., BAILEY, H. and SHILLINGER, G. L. (2019). Predicting residence time using a continuous-time discrete-space model of leatherback turtle satellite telemetry data. *Ecosphere* **10** e02644.
- JOHNSON, D. S., HOOTEN, M. B. and KUHN, C. E. (2013). Estimating animal resource selection from telemetry data using point process models. *J. Anim. Ecol.* **82** 1155–1164.
- JOHNSON, D. S. and LONDON, J. M. (2018). crawl: An R package for fitting continuous-time correlated random walk models to animal movement data. <https://doi.org/10.5281/zenodo.596464>
- JOHNSON, D. S., LONDON, J. M. and KUHN, C. E. (2011). Bayesian inference for animal space use and other movement metrics. *J. Agric. Biol. Environ. Stat.* **16** 357–370. MR2843131 <https://doi.org/10.1007/s13253-011-0056-8>
- JOHNSON, D., PELLAND, N. and STERLING, J. (2021). Supplement to “A continuous-time semi-markov model for animal movement in a dynamic environment.” <https://doi.org/10.1214/20-AOAS1408SUPPA>, <https://doi.org/10.1214/20-AOAS1408SUPPB>, <https://doi.org/10.1214/20-AOAS1408SUPPC>
- JOHNSON, D. S., LONDON, J. M., LEA, M. A. and DURBAN, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology* **89** 1208–1215.
- KASS, R. E. and RAFTERY, A. E. (1995). Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–795. MR3363402 <https://doi.org/10.1080/01621459.1995.10476572>
- LAIRD, N. and OLIVIER, D. (1981). Covariance analysis of censored survival data using log-linear analysis techniques. *J. Amer. Statist. Assoc.* **76** 231–240. MR0624329
- LEA, M. A., JOHNSON, D., REAM, R., STERLING, J., MELIN, S. and GELATT, T. (2009). Extreme weather events influence dispersal of naive northern fur seals. *Biol. Lett.* **5** 252–257.
- LIU, Y., BATTAILLE, B. C., ZIDEK, J. V. and TRITES, A. (2014). Bayesian melding of the dead-reckoned path and GPS measurements for an accurate and high-resolution path of marine mammals. [arXiv:1411.6683](https://arxiv.org/abs/1411.6683).
- MCCRACKEN, M. L., MANLY, B. F. J. and VANDER HEYDEN, M. (1998). The use of discrete-choice models for evaluating resource selection. *J. Agric. Biol. Environ. Stat.* **3** 268–279. MR1817038 <https://doi.org/10.2307/1400582>
- MICROSOFT and WESTON, S. (2017). foreach: Provides foreach looping construct for R. R package version 1.4.4.
- MILLER, D. L. (2019). Bayesian views of generalized additive modelling. [arXiv:1902.01330](https://arxiv.org/abs/1902.01330).
- NAKAGAWA, S. and FRECKLETON, R. P. (2011). Model averaging, missing data and multiple imputation: A case study for behavioural ecology. *Behav. Ecol. Sociobiol.* **65** 103–116.
- NORRIS, J. R. (1998). *Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics 2*. Cambridge Univ. Press, Cambridge. Reprint of 1997 original. MR1600720
- PELLAND, N. A., STERLING, J. T., LEA, M. A., BOND, N. A., REAM, R. R., LEE, C. M. and ERIKSEN, C. C. (2014). Fortuitous encounters between seagliders and adult female northern fur seals (*Callorhinus ursinus*) off the Washington (USA) coast: Upper ocean variability and links to top predator behavior. *PLoS ONE* **9** e101268.
- R CORE TEAM (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- REAM, R. R., STERLING, J. T. and LOUGHLIN, T. R. (2005). Oceanographic features related to northern fur seal migratory movements. *Deep-Sea Res., Part 2, Top. Stud. Oceanogr.* **52** 823–843.
- RUBIN, D. B. and SCHENKER, N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *J. Amer. Statist. Assoc.* **81** 366–374. MR0845877
- SCHARF, H., HOOTEN, M. B. and JOHNSON, D. S. (2017). Imputation approaches for animal movement modeling. *J. Agric. Biol. Environ. Stat.* **22** 335–352. MR3692468 <https://doi.org/10.1007/s13253-017-0294-5>

- STERLING, J. T., SPRING, A. M., IVERSON, S. J., JOHNSON, S. P., PELLAND, N. A., JOHNSON, D. S., LEA, M. A. and BOND, N. A. (2014). The sun, moon, wind, and biological imperative—shaping contrasting wintertime migration and foraging strategies of adult male and female northern fur seals (*Callorhinus ursinus*). *PLoS ONE* **9** e93068.
- VALLIS, G. K. (2006). *Atmospheric and Oceanic Fluid Dynamics*. Cambridge Univ. Press.
- VOLINSKY, C. T., MADIGAN, D., RAFTERY, A. E. and KRONMAL, R. A. (1997). Bayesian model averaging in proportional hazard models: Assessing the risk of a stroke. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **46** 433–448.
- WHITE, I. R., ROYSTON, P. and WOOD, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Stat. Med.* **30** 377–399. MR2758870 <https://doi.org/10.1002/sim.4067>
- WOOD, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 3–36. MR2797734 <https://doi.org/10.1111/j.1467-9868.2010.00749.x>