



## Basis Function Models for Animal Movement

Mevin B. Hooten<sup>a</sup> and Devin S. Johnson<sup>b</sup>

<sup>a</sup>U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife, and Conservation Biology, Department of Statistics, Colorado State University, Fort Collins, CO; <sup>b</sup>Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA

### ABSTRACT

Advances in satellite-based data collection techniques have served as a catalyst for new statistical methodology to analyze these data. In wildlife ecological studies, satellite-based data and methodology have provided a wealth of information about animal space use and the investigation of individual-based animal–environment relationships. With the technology for data collection improving dramatically over time, we are left with massive archives of historical animal telemetry data of varying quality. While many contemporary statistical approaches for inferring movement behavior are specified in discrete time, we develop a flexible continuous-time stochastic integral equation framework that is amenable to reduced-rank second-order covariance parameterizations. We demonstrate how the associated first-order basis functions can be constructed to mimic behavioral characteristics in realistic trajectory processes using telemetry data from mule deer and mountain lion individuals in western North America. Our approach is parallelizable and provides inference for heterogeneous trajectories using nonstationary spatial modeling techniques that are feasible for large telemetry datasets. Supplementary materials for this article are available online.

### ARTICLE HISTORY

Received January 2016  
Revised July 2016

### KEYWORDS

Bayesian model averaging;  
Continuous-time model;  
Process convolution;  
Stochastic differential equation;  
Telemetry data

## 1. Introduction

Advancements in satellite data collection techniques have stimulated the development of dynamic statistical models for individual-based movement processes (Kays et al. 2015). Individual-based statistical movement models have been used in a variety of recent applications including: vehicle (e.g., Gloaguen et al. 2015), cellular phone (e.g., Calabrese et al. 2011), and wildlife (e.g., Hooten et al. 2010) tracking. In particular, new inferential tools are crucial for improving the understanding of wildlife behavior and the response of individual animals to changing landscapes and environmental conditions. Modern telemetry technology allows for remote data collection via “on board” devices (e.g., often using satellite-based observations of geographic position) and has provided massive repositories of information (e.g., Wikelski and Kays 2015).

For example, our application is focused on inferring the movement dynamics of two species of animals, mule deer (*Odocoileus hemionus*) and mountain lion (*Puma concolor*), using archival data sources. Both species occur in western North America and our satellite telemetry data arose from two separate studies to investigate animal spatial ecology where tracking devices were fitted to individuals and recovered later.

Despite numerous improvements to data collection methodology for remote tracking of individual animals, several remaining features of contemporary telemetry data must be addressed when making statistical inference. For example, all forms of remotely collected telemetry data (i.e., measured geographic locations or positions) are susceptible to measurement error that

can depend on the device, satellite system, terrain, land cover, weather, and behavior. Recent advances have led to improved data modeling techniques that properly incorporate (and sometimes estimate) the uncertainty associated with telemetry measurement error (e.g., Brost et al. 2015; McClintock et al. 2015; Buderman et al. 2016).

Irregular temporal measurement is another important feature to consider when modeling telemetry data. Telemetry devices are often programmed (i.e., duty cycled) to record position data at a prespecified set of times. However, the frequency and regularity of these times are not consistent across studies. Furthermore, despite the deterministic programming of satellite telemetry devices, missing data can occur stochastically due to instrumental difficulties as well as environmental and behavioral influences (e.g., terrain and weather).

Dynamic statistical models for animal movement that formally incorporate measurement error typically assume a hierarchical structure (Berliner 1996), where  $\mathbf{s}(t_i)$  are the measured positions at time  $t_i$  (for observation  $i = 1, \dots, n$ ) and depend on the true positions  $\boldsymbol{\mu}(t_i)$  that arise as a dynamical process. Various process models have been proposed for the true underlying individual positions  $\boldsymbol{\mu}(t_i)$ , depending on the desired form of inference (Hooten et al. 2017). Primarily, statistical models for position processes have fallen into three main categories: (1) point process models (e.g., Johnson et al. 2008b; Forester, Im, and Rathouz 2009; Johnson, Hooten, and Kuhn 2013; Brost et al. 2015), (2) discrete-time dynamic models (e.g., Morales et al. 2004; Jonsen, Flemming, and Myers 2005; McClintock

et al. 2012), and (3) continuous-time dynamic models (e.g., Dunn and Gipson 1977; Blackwell 1997; Brillinger et al. 2001; Johnson et al. 2008a; Brillinger 2010). While there have been similar advancements in each of these classes of movement models, we focus on the continuous-time formulations in what follows.

Despite the popularity of point process models and discrete-time dynamic models, both present computational difficulties that prohibit widespread use for all but the simplest forms. Point process models require numerical integration to calculate the likelihood (Berman and Turner 1992; Cressie 1993; Warton and Shepherd 2010; Aarts, Fieberg, and Matthiopoulos 2012). Discrete-time dynamic models often lack a joint model specification and require iterative calculation (Morales et al. 2004; Jonsen, Flemming, and Myers 2005; McClintock et al. 2012). Furthermore, discrete-time models provide inference relative to the scale of temporal discretization and must contain some mechanism to reconcile the times at which data are available with those of the latent discrete-time process (McClintock et al. 2012).

Discrete-time dynamic models for telemetry data are popular because they are heuristically straightforward to understand. They have a long history of use and there is a large body of existing literature associated with modeling discrete time series (e.g., Anderson-Sprecher and Ledolter 1991). Discrete-time models can also be extended to incorporate change-points and hidden Markov processes that allow for time-varying changes in the dynamics, hence better accommodating heterogenous animal behavior through time (e.g., Morales et al. 2004).

Given the attractive properties of discrete-time formulations for model building and the continuous-time nature of the true underlying trajectory, we present a general framework for constructing continuous-time models for animal movement based on limiting processes involving discrete-time models. Our approach provides an intuitive connection to previously existing continuous-time stochastic process models for telemetry data (e.g., Dunn and Gipson 1977; Blackwell 1997; Brillinger et al. 2001; Johnson et al. 2008a; Brillinger 2010). Borrowing techniques commonly used in spatial and spatio-temporal statistics (Cressie and Wikle 2011), we show how continuous-time models based on first-order (i.e., mean) dynamic structure can be implemented using equivalent second-order (i.e., covariance) specifications. Second-order model specifications for animal movement applications have appeared only recently in the literature (e.g., Fleming et al. 2014, 2016).

We present a natural basis function approach to constructing appropriate covariance models for movement processes. Our basis function specification is amenable to rank reduction and, thus, is computationally feasible to implement for large telemetry datasets. We also demonstrate how the choice of basis function may correspond to an explicit representation of animal cognitive processes that may involve memory and perception. Finally, to allow for heterogenous dynamics in movement, we induce a nonstationary continuous-time temporal process by appropriately warping the temporal domain (i.e., temporal deformation, Sampson and Guttorp 1992). Our warping approach allows for temporal clustering of movement behavior in continuous time similar to popular state-switching approaches in discrete-time models (e.g., Morales et al. 2004;

Hanks et al. 2011; McClintock et al. 2012). We also describe a parallelization strategy that can reduce required computational time substantially.

## 2. Methods

### 2.1. Continuous-Time Stochastic Trajectory Models

We focus on process models for continuous-time trajectories (e.g., individual animal movement) and begin by describing dynamic model specifications for trajectories in multiple dimensions. We return to measurement error models for observed trajectories in the next section.

Many hierarchical statistical models involving continuous-time processes have relied on direct connections to Eulerian differential equations (e.g., Cangelosi and Hooten 2009). In contrast, we begin with a discrete-time representation and describe the trajectory of a moving particle in terms of its position  $\boldsymbol{\mu}(t_i)$ , at time  $t_i$ , as

$$\boldsymbol{\mu}(t_i) = \boldsymbol{\mu}(t_0) + \lim_{\Delta t \rightarrow 0} \sum_{j=1}^i \boldsymbol{\varepsilon}(t_j) \tag{1}$$

$$= \boldsymbol{\mu}(t_0) + \mathbf{b}(t_i), \tag{2}$$

where  $\boldsymbol{\mu}(t_0)$  represents the beginning position at time  $t_0$ , and the limiting sum in (1) accumulates a sequence of discrete steps (i.e., displacement vectors,  $\boldsymbol{\varepsilon}(t_j)$ ) up to time  $t_i$ . If the displacement vectors,  $\boldsymbol{\varepsilon}(t_j) \equiv \boldsymbol{\mu}(t_j) - \boldsymbol{\mu}(t_{j-1})$ , are independent multivariate Gaussian, such that  $\boldsymbol{\varepsilon}(t_j) \sim N(\mathbf{0}, \sigma^2 \Delta t \mathbf{I})$  (i.e., white noise), we arrive at a continuous-time stochastic integral representation for  $\boldsymbol{\mu}(t_i)$  where  $\sigma$  controls the magnitude of displacement and  $\mathbf{b}(t_i)$  is scaled multivariate Brownian motion (i.e., a multivariate Wiener process). In the limiting trajectory model (1), the number of displacement steps in the sum increases as  $\Delta t \rightarrow 0$ , resulting in an infinite series that is often written as the integral

$$\mathbf{b}(t_i) = \int_{t_0}^{t_i} d\mathbf{b}(\tau), \tag{3}$$

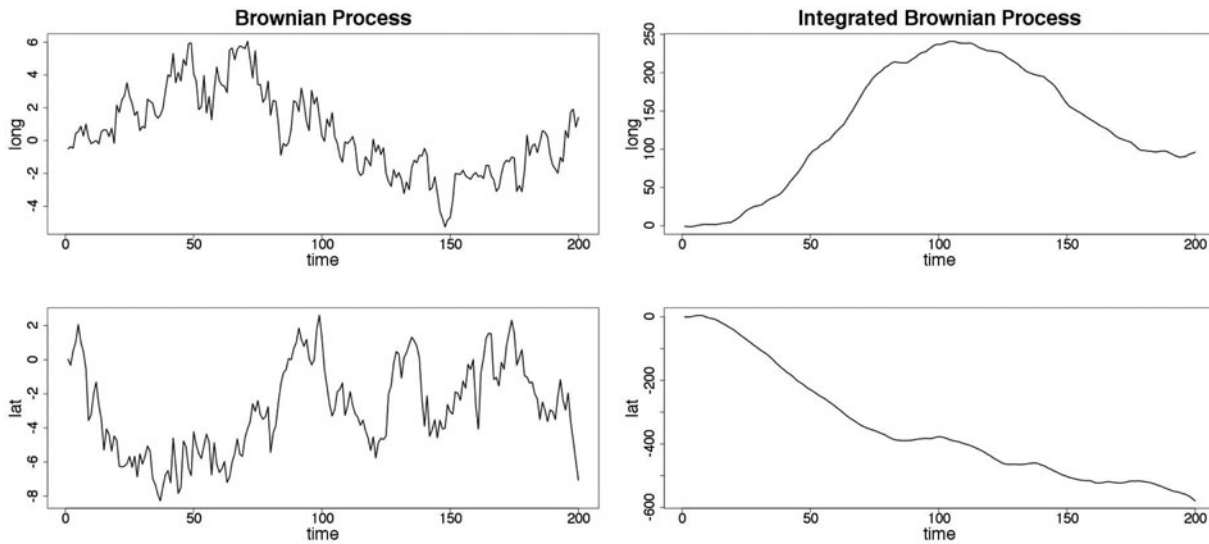
using Ito notation (Protter 2004).

While the Brownian motion (BM) process model for an individual animal trajectory in (3) is a stochastic integral equation (SIE), it is also common to see it expressed as a stochastic differential equation (SDE) by differentiating both sides of (1) using Ito calculus (e.g., Dunn and Gipson 1977; Blackwell 1997; Brillinger et al. 2001; Preisler et al. 2004).

### 2.2. Smoothness in Trajectory Models

Stochastic trajectory models based on Brownian motion (e.g., (2)) are inherently noisy (Figure 1, left panels) as continuous-time processes. Johnson et al. (2008a) presented a stochastic trajectory model for the velocity associated with animal movement that was integrated over time to yield a smoother position process (Figure 1, right panels).

We show that the velocity model by Johnson et al. (2008a) fits into a larger class of stochastic trajectory models by reparameterizing the simple Brownian SIE (2). Recall that Brownian motion ( $\mathbf{b}(t)$ ), at time  $t$ , can be expressed as an integral of white



**Figure 1.** Left panels: Two-dimensional Brownian motion process,  $\mathbf{b}(t)$ . Right panels: Two-dimensional integrated Brownian motion,  $\boldsymbol{\eta}(t)$ . Processes are displayed marginally in longitude and latitude.

noise. Thus, if we integrate Brownian motion itself, with respect to time, we have

$$\boldsymbol{\eta}(t) = \int_{t_0}^t \mathbf{b}(\tau) d\tau, \quad (4)$$

where  $\boldsymbol{\eta}(t)$  is a version of the integrated stochastic process proposed by Johnson et al. (2008a). This integrated Brownian motion (IBM, e.g., Shepp 1966; Wecker and Ansley 1983; Rue and Held 2005) model (4) can be likened to that of Johnson et al. (2008a) by substituting  $\boldsymbol{\eta}(t)$  into the position process (2) to yield  $\boldsymbol{\mu}(t) = \boldsymbol{\mu}(t_0) + \boldsymbol{\eta}(t)$ . Jonsen, Flemming, and Myers (2005) set an earlier precedent for modeling velocity, but strictly in a discrete-time framework.

The approach proposed by Johnson et al. (2008a) suggests a more general framework for modeling movement that can be obtained by reparameterizing the velocity model using the  $2 \times 2$  matrix  $\mathbf{H}(t, \tau)$  with diagonal elements equal to the function

$$h(t, \tau) = \begin{cases} 1 & \text{if } t_0 < \tau \leq t \\ 0 & \text{if } t < \tau \leq t_n \end{cases}, \quad (5)$$

where  $t_n$  is the last time at which data are observed, and off-diagonal elements equal zero. Substituting  $\mathbf{H}(t, \tau)$  into (4), the velocity-based Brownian motion model appears as the convolution

$$\boldsymbol{\eta}(t) = \int_{t_0}^{t_n} \mathbf{H}(t, \tau) \mathbf{b}(\tau) d\tau. \quad (6)$$

The convolution in (6) is the key feature in a more general class of stochastic process models for animal movement trajectories. For example, if  $h(t, \tau)$  is a continuous function, such that  $t_0 \leq t \leq t_n$ ,  $t_0 \leq \tau \leq t_n$  and with finite positive integral  $0 < \int_{t_0}^{t_n} h(t, \tau) d\tau < \infty$ , then a new general class of continuous-time animal movement models arises. We refer to this class of models as “functional movement models” (FMMs; Buderman et al. 2016).

The ability to specify continuous-time movement models as a convolution (6) has two major advantages. First, it clearly identifies the connections among animal movement models and similar models used in spatial statistics and time series. Second, for the same reasons convolution specifications are popular in spatial statistics and time series, we show that FMMs share similar advantageous properties.

In Appendix A (supplementary material), we show that the FMM in (6) can be rewritten as

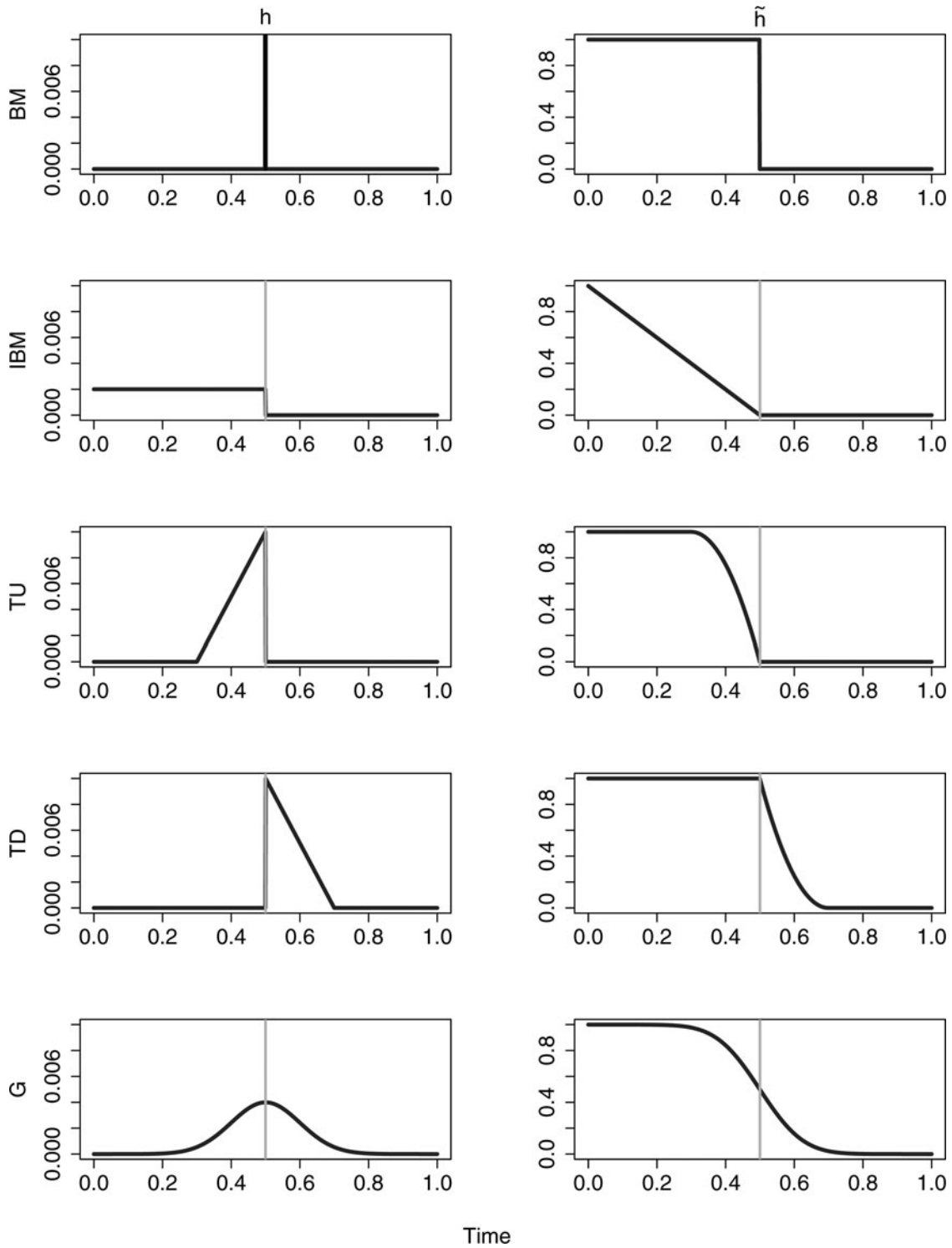
$$\boldsymbol{\eta}(t) = \int_{t_0}^{t_n} \tilde{\mathbf{H}}(t, \tau) d\mathbf{b}(\tau), \quad (7)$$

where  $\tilde{\mathbf{H}}(t, \tau)$  has diagonal elements

$$\tilde{h}(t, \tau) = \int_{\tau}^{t_n} h(t, \tilde{\tau}) d\tilde{\tau}. \quad (8)$$

The FMM in (7) has the same form described in spatial statistics as a “process convolution” (or kernel convolution, e.g., Barry and Ver Hoef 1996; Higdon 1998; Lee et al. 2005; Calder 2007). The process convolution has been instrumental in many fields, but especially in spatial statistics for allowing both complicated and efficient representations of covariance structure.

It is clear from (7) that we need not simulate Brownian motion, rather, we can operate on it implicitly by transforming the matrix function  $\mathbf{H}(t, \tau)$  to  $\tilde{\mathbf{H}}(t, \tau)$  via integration and convolving  $\tilde{\mathbf{H}}(t, \tau)$  with white noise directly. For example, consider the Gaussian kernel as the function  $h(t, \tau)$ . The Gaussian kernel is one of the most commonly used functions in kernel convolution methods (e.g., Barry and Ver Hoef 1996; Higdon 1998). We convert  $h(t, \tau)$  to the required function  $\tilde{h}(t, \tau)$  using (8) and arrive at a numerical solution for the new kernel function by subtracting the normal cumulative distribution function (CDF) from one, a trivial calculation in most statistical software. With respect to the time domain, the kernel  $\tilde{h}(t, \tau)$  appears different than most kernels used in time series or spatial statistics



**Figure 2.** Example kernels  $h(t, \tau)$  (left column) and resulting integrated kernels  $\tilde{h}(t, \tau)$  (right column). The first row results in the regular Brownian motion (BM), row 2 shows integrated Brownian motion (IBM), rows 3 and 4 show tail-up (TU) and tail-down (TD) kernel functions, and row 5 shows the Gaussian (G) kernel functions. Rows 3 through 5 (TU, TD, and G) are more common in time series and spatial statistics. The vertical gray line indicates time  $t$  for the particular kernel shown; in this case  $t = 0.5$ .

(Figure 2, G, row 5). Rather than being unimodal and symmetric, it has a sigmoidal shape equal to one at  $t = t_0$  and nonlinearly decreasing to zero at  $t = t_n$ , resembling an I-spline (Ramsay 1988). In effect, this new kernel is accumulating the white noise up to near time  $t$  and including a discounted amount of white noise ahead of time  $t$ .

There are many options for kernel functions and each will result in different stochastic process models for animal

movement. In fact, we have already shown that the FMM class of movement models is general enough to include an integrated Brownian motion model similar to that developed by Johnson et al. (2008a). The FMM class also includes the original unsmoothed Brownian motion process if we let  $h(t, \tau)$  be a point mass function at  $\tau = t$  and zero elsewhere (Figure 2). An approximation to the point mass kernel function can also be achieved by taking the limit as  $\phi \rightarrow 0$  of our Gaussian kernel

and results in an integrated kernel function of

$$\tilde{h}(t, \tau) = \begin{cases} 1 & \text{if } t_0 < \tau \leq t \\ 0 & \text{if } t < \tau \leq t_n \end{cases} \quad (9)$$

The integrated kernel sums all past velocities to obtain the current position. The steep drop at  $\tau = t$  is what induces roughness in the original Brownian motion process (Figure 2, BM, row 1). Whereas, when we use a nonpointmass function for  $h(t, \tau)$ , we arrive at a smoother stochastic process model for movement.

We highlight a few different kernel functions to examine their implications for animal movement behavior. In doing so, it is simplest to interpret the  $h(t, \tau)$  and  $\tilde{h}(t, \tau)$  functions directly. For example, using the direct integration of velocity as proposed by Johnson et al. (2008a) results in the integrated Brownian motion (IBM) kernel ( $\tilde{h}(t, \tau)$ ) on the second row in Figure 2. In this case, the individual's position accumulates its past steps, which are noisy themselves in direction and length, but have some general momentum. The "tail up" (TU) kernel shown in the third row of Figure 2 models the current position based on past steps that decay linearly with time (we borrow the tail up and tail down terminology from Ver Hoef and Peterson 2010). In this case, the individual's position is more strongly a function of recent steps than steps in the distant past. The opposite is true with the "tail down" (TD) kernel shown in row 4 of Figure 2, where only future steps influence position. Heuristically, we interpret the resulting movement as perception driven. That is, the individual may have an awareness of a distant destination that affects their movement. Finally, the Gaussian kernel discussed earlier and shown in the bottom row of Figure 2 indicates a symmetric mixture of previous and future velocities, suggesting an equal perception of former and future events by the individual. Appendix B (supplementary material) contains the functional forms and a visualization of the basis functions in  $\tilde{\mathbf{H}}$ .

### 2.3. Functional Movement Models and Covariance

In the statistical setting, we seek inference for model parameters as well as predictions (interpolation) of the position process in time. Therefore, we need to connect telemetry data at a finite set of observation times  $\{t_1, \dots, t_n\}$  to the dependence structure imposed by the FMM. A common approach for expressing dependence structure in correlated processes is using covariance. Translating dynamically induced temporal structure into covariance yields several computational advantages for FMMs. Ultimately, we construct a hierarchical statistical model for telemetry data  $\mathbf{s}(t_i)$ , for  $i = 1, \dots, n$ , that relies on a latent dynamical process governed by an FMM. In what follows, we describe the covariance properties of FMMs and then specify a hierarchical model that exploits those properties in the next sections.

Using the basic FMM in (6), we can choose from a large set of possible smoothing kernels ( $h(t, \tau)$ ) for Brownian motion and arrive at the appropriate form for the integrated kernel ( $\tilde{h}(t, \tau)$ ) that is convolved with white noise. We can then use  $\tilde{h}(t, \tau)$  directly to construct the proper covariance function for the joint process. In fact, for a one-dimensional movement process  $\eta(t)$ , the covariance function can be calculated (e.g., Paciorek and

Schervish 2006) as the convolution of the kernels

$$\text{cov}(\eta(t_1), \eta(t_2)) = \int_{t_0}^{t_n} \sigma^2 \tilde{h}(t_1, \tau) \tilde{h}(t_2, \tau) d\tau, \quad (10)$$

for any two times, say  $t_1$  and  $t_2$ .

One benefit of the covariance function (10) is that, for a finite subset of  $n$  times  $\{t_1, \dots, t_n\}$  and process  $\boldsymbol{\eta} \equiv (\eta_1, \dots, \eta_n)'$ , the joint probability model can be expressed as

$$\boldsymbol{\eta} \sim \mathbf{N}(\mathbf{0}, \sigma^2 \Delta t \tilde{\mathbf{H}} \tilde{\mathbf{H}}'), \quad (11)$$

where  $\mathbf{0}$  is an  $n \times 1$  vector of zeros and  $\tilde{\mathbf{H}}$  is a matrix of basis functions with the  $i$ th row equal to  $\tilde{h}(t_i, \tau)$  for all  $\tau$ . This procedure for constructing the covariance matrix and defining a correlated process is similar to that recommended in spatial statistics (e.g., Paciorek and Schervish 2006; Ver Hoef and Peterson 2010). For simplicity, the resulting model for the joint one-dimensional position process ( $\boldsymbol{\mu}$ , an  $n \times 1$  vector) at the observation times for smooth Brownian motion is

$$\boldsymbol{\mu} \sim \mathbf{N}(\boldsymbol{\mu}(0) \mathbf{1}, \sigma^2 \Delta t \tilde{\mathbf{H}} \tilde{\mathbf{H}}'). \quad (12)$$

Despite the fact that it is often more intuitive to model the process from the first-moment (i.e., mean dynamical structure) rather than the second-moment (Wikle and Hooten 2010), the joint form of (12) with dependence imposed through the  $n \times n$  correlation matrix  $\tilde{\mathbf{H}} \tilde{\mathbf{H}}'$  can be useful computationally (Sampson 2010). When the integral in (10) cannot be used to analytically compute the necessary covariance matrix, we can still use the outer product of the matrices explicitly (i.e.,  $\tilde{\mathbf{H}} \tilde{\mathbf{H}}'$ ). However, the true covariance requires the number of columns of  $\tilde{\mathbf{H}}$  to approach infinity, which, under approximation, can lead to computational difficulties. Higdon (2002) suggested a finite process convolution as an approximation. In the finite approximation, the number of columns of  $\tilde{\mathbf{H}}$  could be reduced to say,  $m$  columns. This rank reduction implies that there are  $m$  knots in the temporal domain that anchor the basis functions (i.e., kernels) and, thus, only  $m$  white-noise terms are required so that  $\boldsymbol{\eta} \approx \tilde{\mathbf{H}} \boldsymbol{\varepsilon}$ , where  $\tilde{\mathbf{H}}$  is an  $n \times m$  matrix and  $\boldsymbol{\varepsilon} \equiv (\varepsilon(t_1), \dots, \varepsilon(t_j), \dots, \varepsilon(t_m))'$  is an  $m \times 1$  vector. The use of a finite approximation to the convolution is also sometimes referred to as a reduced-rank method (Wikle 2010). Rank reduction can improve computational efficiency and has become very popular in spatial and spatio-temporal statistics for large datasets (Nychka and Saltzman 1998; Nychka, Wikle, and Royle 2002; Banerjee et al. 2008; Cressie and Wikle 2011).

To illustrate how kernel functions relate to covariance thus far, we simplified the movement process so that it is one-dimensional in space. The same approach generalizes to higher dimensions. In the more typical two-dimensional case, we stack the vectors in each dimension to form a single  $2n \times 1$  vector  $\boldsymbol{\eta}$ . Then the joint model can be written as

$$\boldsymbol{\eta} \sim \mathbf{N}(\mathbf{0}, \sigma^2 \Delta t (\mathbf{I} \otimes \tilde{\mathbf{H}} \tilde{\mathbf{H}}')), \quad (13)$$

where  $\mathbf{0}$  is a  $2n \times 1$  vector of zeros and  $\mathbf{I}$  is a  $2 \times 2$  identity matrix (assuming that  $\tilde{\mathbf{H}}$  is the appropriate set of basis functions for both directions, i.e., latitude and longitude).

### 2.4. Heterogenous Dynamics

Discrete-time trajectory models for animal movement (e.g., Morales et al. 2004) are commonly specified with time-varying dynamics. In spatial statistics, heterogeneity in dependence structure is often treated as nonstationarity. A variety of approaches have been suggested for modeling nonstationary continuous processes (e.g., Sampson and Guttorp 1992; Higdon 2002). We describe a temporal warping approach to modeling nonstationary movement.

Our warping approach (also known as “deformation”) was described in a continuous spatial modeling context by Sampson and Guttorp (1992) and later extended (e.g., Damian, Sampson, and Guttorp 2001; Schmidt and O’Hagan 2003). In the warping approach, we map the times  $\mathbf{t} \equiv (t_1, \dots, t_n)'$  to a new set of times  $\mathbf{w} \equiv (w_1, \dots, w_n)'$  using a smooth function so that no “folding” occurs (i.e., transformed times retain order). A model-based method to perform the warping allows  $\mathbf{w}$  to arise as a correlated random field anchored by  $\mathbf{t}$ . For example, a simple additive warping can be induced through the Gaussian process model  $\mathbf{w} \sim N(\mathbf{t}, \Sigma_w)$ , where a smooth covariance function, such as the Gaussian  $\Sigma_{w,ij} \equiv \sigma_w^2 \exp(-(t_i - t_j)^2/\phi_w^2)$ , is chosen. Conventionally, the warping parameters  $\sigma_w^2$  and  $\phi_w$  are constrained to yield nonfolding temporal warp fields.

When fitting the FMM (12), we construct the basis functions so that they depend on the warped times. For example, using a Gaussian kernel, we now have

$$h(t, \tau) \propto \exp\left(-\frac{(w(t) - \tau)^2}{\phi^2}\right). \tag{14}$$

Notice that the kernel  $h(t, \tau)$  only depends on a single range parameter  $\phi$  that does not need to vary in time because the heterogeneity enters through the norm in the warped temporal domain  $(w(t) - \tau)^2$ . Thus, the warp field  $\mathbf{w}$  is a latent process to be estimated in the model and can provide inference associated with changes in movement dynamics. For example, the derivative  $dw(t)/dt$  of the warp field provides insight about changes in the individual’s velocity over time. When  $dw(t)/dt$  is large, the warp field expands the temporal domain to allow for smoother paths, either from faster, directed movement, or lack of movement. Expanding the temporal domain can account for behavior indicative of long-distance migration or sedentariness. By contrast, when the  $dw(t)/dt$  is small, the temporal domain compresses, resulting in more tortuous paths and accounting for periods of time with sharper turning angles. Temporal compression leads to behavior more typical of an area restricted search (e.g., Knell and Codling 2012), or an individual performing routine activities within its home range (e.g., foraging).

### 2.5. Full Model Specification

Assuming Gaussian measurement error associated with the telemetry observations and a full-rank position process  $\boldsymbol{\mu}$ , we specify the full model in a hierarchical Bayesian framework:

$$\mathbf{s} \sim N(\mathbf{K}\boldsymbol{\mu}, \mathbf{I} \otimes \Sigma_s), \tag{15}$$

$$\boldsymbol{\mu} = \boldsymbol{\mu}(0) \otimes \mathbf{1} + (\mathbf{I} \otimes \tilde{\mathbf{H}})\boldsymbol{\epsilon}, \tag{16}$$

$$\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \Delta t \mathbf{I}), \tag{17}$$

where  $\mathbf{K}$  is a  $2n \times 2m$  matrix that maps the data to the underlying process at the appropriate times and the error covariance matrix could be left general or simplified as  $\Sigma_s \equiv \sigma_s^2 \mathbf{I}$  depending on the data source. As before, the initial position is denoted as  $\boldsymbol{\mu}(0)$  and the process variance is  $\sigma^2$ . In the model formulation that allows for temporal heterogeneity, the matrix of basis functions depends on the range parameter  $\phi$ , as well as the warped time process  $\mathbf{w}$ . Thus, we augment the model specification by letting  $\mathbf{w} \sim N(\mathbf{t}, \Sigma_w)$ .

In principle, the choice of priors is study specific, but any proper prior distributions with the correct support may be used for the parameters in the FMM. In Gaussian state-space models, conjugate priors for  $\sigma_s^2$  and  $\sigma^2$  are inverse gamma, where strong prior information helps to separate the measurement error variance  $\sigma_s^2$ . Thus, similar to geostatistics, inference can be sensitive to the chosen prior (e.g., Finley, Banerjee, and Gelfand 2015) when available information is used to inform priors. Alternative priors for the latent process variance are also available. For example, we used a proper uniform prior for  $\sigma$ , as recommended by Gelman (2006) for latent Gaussian processes. A variety of options are available for the temporal range parameter  $\phi$ , however, a discrete uniform prior ( $\phi \sim \text{DiscUnif}(\Phi)$ ) is computationally advantageous (Diggle and Ribeiro 2002). We elaborate on the discrete uniform prior for the temporal range parameter ( $\phi$ ) in the next section.

### 2.6. Model Implementation

The full model, described in the previous section, can be implemented using an MCMC algorithm to sample from all full-conditional distributions sequentially. However, such an algorithm would be prohibitively slow for all but the smallest telemetry datasets. Thus, we provide a model reparameterization that results in several computational improvements. Our approach involves three critical features: (1) Discrete uniform prior for the range parameter, (2) integrated likelihood, and (3) a mixture model to accommodate temporal heterogeneity. We describe each of these features in what follows.

When specifying the prior  $\phi \sim \text{DiscUnif}(\Phi)$ , for each value of  $\phi$  in the support  $\Phi$ , the entire matrix of basis functions  $\tilde{\mathbf{H}}$  can be precomputed and accessed as needed in a Markov chain Monte Carlo (MCMC) algorithm. The discrete uniform prior permits an MCMC algorithm that requires only minimal matrix calculations (Diggle and Ribeiro 2002).

We used Rao–Blackwellization to derive an integrated likelihood that results in a more stable MCMC algorithm and allows us to avoid sampling the latent process directly. The integrated likelihood provides better mixing for covariance parameters and the latent process can be recovered post hoc using saved MCMC samples and composition sampling in a secondary algorithm. This approach is a common computational strategy used in spatial statistics (e.g., Finley, Banerjee, and Gelfand 2015). In our case, multivariate normal properties allow us to integrate  $\boldsymbol{\epsilon}$  out of our hierarchical model, resulting in the integrated model formulation

$$\mathbf{s} \sim N(\mathbf{K}(\boldsymbol{\mu}(0) \otimes \mathbf{1}), \mathbf{I} \otimes \Sigma_s + \sigma^2 \mathbf{K}(\mathbf{I} \otimes \tilde{\mathbf{H}})(\mathbf{I} \otimes \tilde{\mathbf{H}})' \mathbf{K}'). \tag{18}$$

Note that we omit the time step  $\Delta t$  from the white-noise specification because the variance term  $\sigma^2$  can account for it in the discrete-time implementation. The resulting combined covariance matrix in (18) includes both the measurement error variance and the process covariance and has the form  $\mathbf{A} + \mathbf{BCD}$ , which can be inverted efficiently using the Sherman–Morrison–Woodbury identity if necessary (Appendix C, supplementary material). In addition to the precalculation of terms involving the matrix of basis functions  $\tilde{\mathbf{H}}$ , fast matrix calculations are essential for fitting the FMM efficiently to real datasets.

The two steps described above facilitate the fitting of a temporally homogenous model that is conditioned on a known warping ( $\mathbf{w}$ ). However, because the optimal warping is unknown and occurs in a nonlinear covariance function for the data, we would need to sample it using Metropolis–Hastings within the broader MCMC algorithm for fitting the model. The resulting computational requirements are prohibitively large. Thus, we propose a mixture model framework that can be implemented using reversible-jump MCMC (RJCMCMC; Green 1995). In practice, the mixture model can be implemented using a finite set of candidate temporal warp fields, say  $\mathbf{w}_j$ , for  $j = 1, \dots, J$  and  $L$  basis function types (Figure 2) resulting in  $N = L \cdot J$  total models. We formulate the mixture using models  $\mathcal{M}_{lj}$  (for  $l = 1, \dots, L$  and  $j = 1, \dots, J$ ) and latent indicator variables  $z_{lj}$  as

$$\mathbf{s} \sim \begin{cases} \mathcal{M}_{11}, & z_{11} = 1 \\ \mathcal{M}_{21}, & z_{12} = 1 \\ \vdots \\ \mathcal{M}_{LJ}, & z_{LJ} = 1 \end{cases}, \quad (19)$$

where  $\mathcal{M}_{lj} \equiv \mathbf{N}(\mathbf{K}(\boldsymbol{\mu}(0) \otimes \mathbf{1}), \mathbf{I} \otimes \boldsymbol{\Sigma}_s + \sigma^2 \mathbf{K} \mathbf{R}_{lj}(\boldsymbol{\phi}, \mathbf{w}_j) \mathbf{K}')$  for  $l = 1, \dots, L$  and  $j = 1, \dots, J$ , and each correlation matrix is defined as  $\mathbf{R}_{lj} \equiv (\mathbf{I} \otimes \tilde{\mathbf{H}}_{lj})(\mathbf{I} \otimes \tilde{\mathbf{H}}_{lj})'$  (with implicit dependence on the appropriate basis function type, range parameter, and temporal warping). The indicator variables  $\mathbf{z} \equiv (z_{11}, z_{12}, \dots, z_{LJ})'$  are multinomial ( $\mathbf{z} \sim \text{MN}(1, \mathbf{p})$ ) with  $\mathbf{p}$  serving as prior model probabilities. Using the mixture model specification in (19), and conditioning on a large set of potential warp fields, the resulting MCMC algorithm involves Gibbs updates for the indicator variables,  $z_{lj}$ . The marginal posterior mean of the indicator variable  $z_{lj}$ ,  $E(z_{lj}|\mathbf{s})$ , corresponds to the posterior model probability for model  $\mathcal{M}_{lj}$ ,  $P(\mathcal{M}_{lj}|\mathbf{s})$ .

An alternative computational strategy involves the two-stage RJCMCMC approach by Barker and Link (2013) to find the posterior model probability  $P(\mathcal{M}_{lj}|\mathbf{s})$  and model averaged posterior inference (Appendix D, supplementary material). The advantage of the RJCMCMC approach by Barker and Link (2013) is that each model can be fit independently (and in parallel), and a secondary MCMC algorithm can be used to sample from the averaged posterior distributions. Parallel model fits provide appreciable gains in computational efficiency. Finally, after fitting each of the models, and applying Bayesian model averaging via RJCMCMC, we use a tertiary algorithm to obtain realizations of the latent position process  $\boldsymbol{\mu}(t)$  for any time of interest  $t$ . To fit the models we propose, the tertiary algorithm can also be parallelized, but is usually fast enough to implement sequentially.

The two-stage RJCMCMC facilitates the ability to compare models with different basis function specifications. Thus, we include both, a set of candidate temporal warp fields and all

of the covariance models constructed from the basis functions shown in Figure 2. We accumulate the posterior model probability resulting from the RJCMCMC algorithm for each form of basis function, resulting in  $\sum_{j=1}^J P(\mathcal{M}_{lj}|\mathbf{s})$  for each basis function type  $l$ , which may provide insight about potential mechanisms (i.e., perception and memory) that influence animal behavior and movement.

Finally, to obtain a set of candidate warp fields, we simulate warp fields as Gaussian processes with Gaussian covariance structure to provide appropriate smoothness in the resulting temporal deformation. High variance and low temporal correlation in the warp distribution results in a higher proportion of folding warp fields. Thus, we use a rejection sampling algorithm to retain only those warp fields that obey the ordering of the original time domain (i.e., do not fold). We simulate warp fields based on a range of parameter values to yield a candidate set that do not fold, yet contain a variety of multiscale patterns. The appropriate temporal scale for the warp distribution is informed by the data through the model averaging procedure.

In simulation, our method for fitting the FMM to data was approximately an order of magnitude faster than conventional Bayesian computational methods. We demonstrate that our Bayesian FMM recovers model parameters for simulated datasets with varying amounts of data missingness and temporal heterogeneity in Appendix E (supplementary material). In what follows, we apply the temporally heterogenous FMM to satellite telemetry data corresponding to the movement of large mammals in western North America and describe the implications for our proposed methods.

### 3. Data Analysis: Animal Movement

We apply our temporally heterogenous FMM to satellite telemetry datasets consisting of recorded mule deer (*Odocoileus hemionus*) and mountain lion (*Puma concolor*) positions in Southeastern Utah and Northern Colorado, USA (originally analyzed for other purposes by Hooten et al. 2010 and Hanks, Hooten, and Alldredge 2015). Mule deer and mountain lion are important species in the western North America that often exhibit seasonally varying movement behavior (Pierce et al. 1999; Ager et al. 2003). Many mule deer individuals split their time between winter and summer ranges and migrate from one to the other in the spring and autumn. Mountain lion movement behavior also varies temporally, with individuals roaming regionally and often returning to locations where prey has been captured. In both mule deer and mountain lion individuals, there is a clear heterogeneity in movement dynamics. Our FMM approach accommodates such movement heterogeneity by treating it as temporal nonstationarity.

For the mule deer example, we focus on a five-day period during autumn of 2005 that spans the migratory behavior of an individual near Castle Valley, Utah, USA. These data were collected as part of a larger study of the spread of chronic wasting disease (McFarlane 2007), a contagious prion disease that occurs in North American ungulates (e.g., Farnsworth et al. 2006; Miller, Hobbs, and Tavener 2006; Garlick et al. 2011; Garlick et al. 2014; Evans et al. 2016). Migratory routes are a critical element of the natural history of mule deer in western North America

(Nicholson, Bowyer, and Kie 1997) and an improved understanding of them is essential for habitat management.

For the mountain lion, we focus on a 33-day period during summer of 2011 in the Front Range of Colorado, USA. These data were collected as part of a larger study to assess the movement and survival of mountain lions in the Front Range of Colorado. The mountain lion is an important native carnivore that was once ubiquitous throughout much of North America (Young and Goldman 1946; Laundré 2013). Mountain lions play an essential role in ecosystem function by attenuating populations of prey species, which often include mule deer (Knopff et al. 2010). Inference concerning mountain lion movement is also important in the Front Range because of rapid development of exurban areas in existing habitat (Blecha 2015).

The resulting telemetry data (Figure 3) for both species were obtained using global positioning system (GPS) telemetry devices affixed to the individuals and are comprised of 226 measured geographic positions spaced approximately 30 min apart for the mule deer individual and 221 observations spaced approximately 3 hr apart for the mountain lion individual. For analysis and plotting purposes, we standardized the position data (i.e., subtracted the sample mean and divided by the pooled sample standard deviation) and converted the time scale to the [0,1] interval. In principle, data transformation is not necessary for model implementation, but it allows us to use the same software to simulate warp fields, fit the models, as well as use similar prior specification among species, leading to fewer numerical overflow issues.

To implement an FMM for the mule deer and mountain lion data using the specification presented in the previous section (18), we used all of the kernel forms,  $h_l(t, \tau)$  for  $l = 1, \dots, 5$ , described in Figure 2 (and Appendix B). Following Diggle and Ribeiro (2002), we reparameterized the covariance matrix for a given basis function type  $l$  in the FMM (18) such that

$$\mathbf{I} \otimes \Sigma_s + \sigma^2 (\mathbf{I} \otimes \tilde{\mathbf{H}}_l) (\mathbf{I} \otimes \tilde{\mathbf{H}}_l)' = \sigma_s^2 (\mathbf{I} + \sigma_{\mu/s}^2 (\mathbf{I} \otimes \tilde{\mathbf{H}}_l) (\mathbf{I} \otimes \tilde{\mathbf{H}}_l)'), \quad (20)$$

where,  $\sigma_{\mu/s}^2 = \sigma^2 / \sigma_s^2$  and  $\tilde{\mathbf{H}}_l(\phi)$  is a function of the temporal range parameter  $\phi$  for the Gaussian, tail-up, and tail-down kernel forms. For the standardized data, we used a discrete uniform prior for  $\phi$  on 100 evenly spaced values from 0.001 to 0.1, a uniform prior for  $\sigma_{\mu/s}$  on support (0, 20), and an informative inverse gamma prior for  $\sigma_s^2 \sim \text{IG}(12, 0.01)$ , representing strong prior information that GPS telemetry error is often less than 150 m.

In implementing the FMM, we used 400 equally spaced temporal knots for the kernels and we sampled approximately 4000 candidate warp fields using a Latin hypercube design from a Gaussian process with Gaussian covariance constrained to avoid temporal folding. The warp distribution was based on 100 parameter combinations for  $\sigma_w$  and  $\phi_w$  (10 equally spaced values for each parameter, ranging from 0.001 to 1). The resulting multiscale candidate warp fields represent a wide range of temporal deformation patterns that are combined optimally using Bayesian model averaging. Trace plots for model parameters indicated excellent MCMC mixing and convergence, based on 10,000 MCMC samples for each model fit. Each model fit

required only 3 min on a 16-core workstation with 3 Ghz processors and 64 GB of memory. The total time required to fit the model for each of the warp fields in parallel was approximately 3 hr, which is substantially less than the 50 hr it would require to compute in sequence. The secondary and tertiary algorithms to perform the BMA and posterior predictive sampling required only four additional minutes. We used empirical variograms based on the posterior predictive residuals to check the models (Appendix F, supplementary material). For both the mule deer and mountain lion data, the variograms showed no evidence of remaining temporal autocorrelation unaccounted for by the model.

The resulting inference is summarized in Figure 3. The top row of panels in Figure 3 presents the data and estimated latent process for the mule deer and mountain lion individuals in two-dimensional geographic space. The gray lines are posterior predictive realizations of the latent process  $\mu(t)$ . The middle two rows of panels in Figure 3 display the marginal latent process in longitude and latitude. The bottom row of panels in Figure 3 shows the inferred derivative of the temporal warp at the posterior mean of the position process.

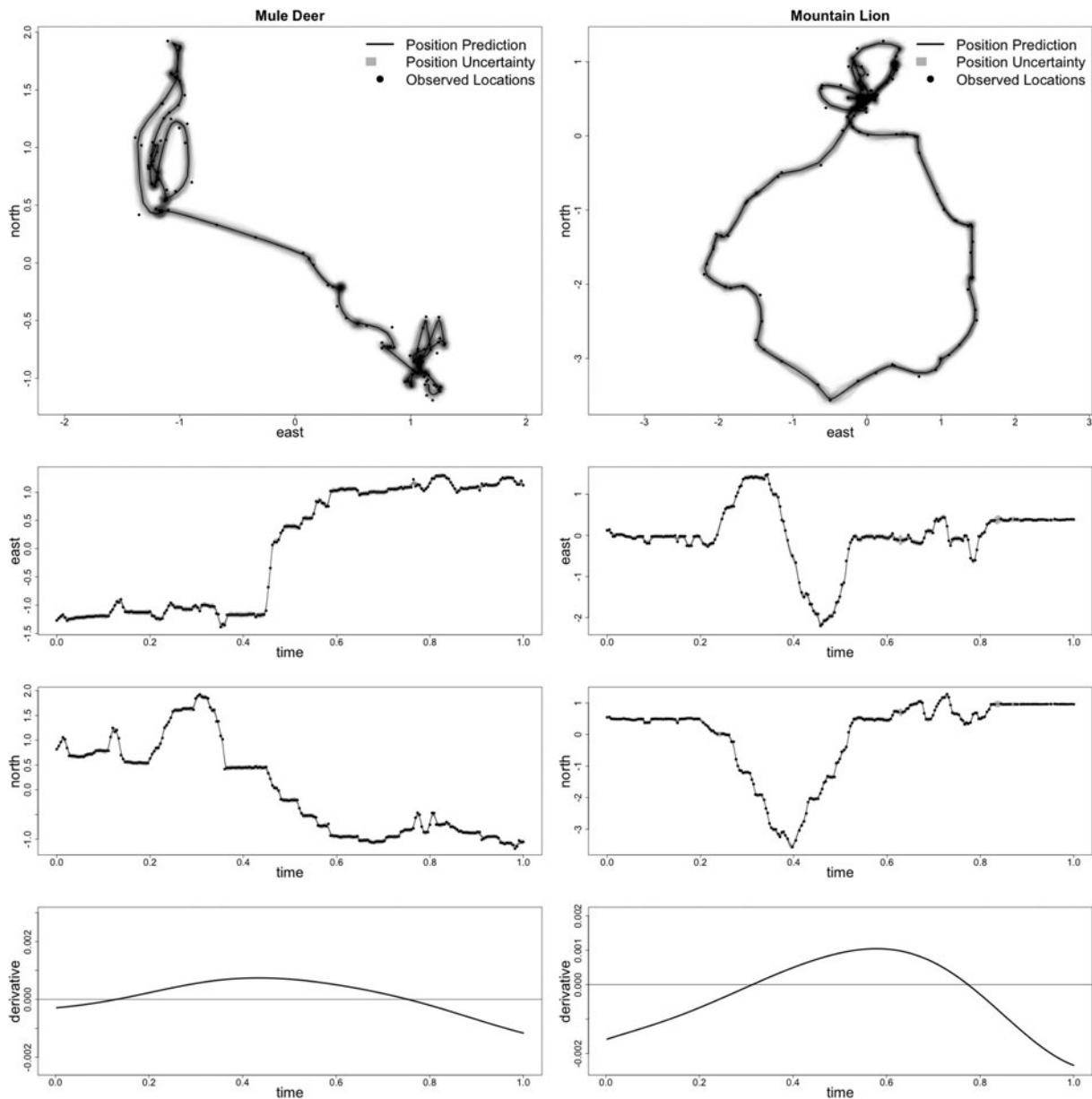
The mule deer individual migrated during the time interval 0.3–0.6 (Figure 3). As discussed in the Heterogenous Dynamics section, the warp derivative can provide insight on the movement dynamics. The migratory movement behavior caused the BMA warp derivative to be positive, indicating that a temporal expansion is needed to accommodate the trajectory during that time period. In contrast, the warp derivative was negative during the time interval 0.6–0.75 suggesting a temporal contraction was necessary to capture the different movement behavior for the individual directly after migration.

The accumulated posterior model probabilities for the kernel forms suggested that Brownian motion and integrated Brownian motion were not helpful in describing mule deer movement based on these data (i.e., posterior model probability equal to zero). However, the Gaussian, tail-up, and tail-down kernels accounted for 0.059%, 0.920%, and 0.021%, respectively. The distribution of posterior model probabilities for the kernel forms suggest that, while Gaussian, tail-up, and tail-down kernels have some nonzero contribution, the tail-up kernel dominates in characterizing mule deer movement. Recall that the tail-up kernel acts as a local filter for Brownian motion, smoothing the individual trajectory using past Brownian steps. A potential interpretation of the posterior model probabilities suggest that memory may play the largest role in mule deer movement during migration.

The warp derivative for the mountain lion individual (Figure 3) suggests a temporal contraction at the beginning and end of the temporal domain (i.e.,  $t < 0.2$  and  $t > 0.8$ ) with a temporal expansion between (i.e.,  $0.4 < t < 0.8$ ). The FMM inference confirms that the mountain lion individual occupies a location with prey before traversing a large loop, only to return to the same location a few weeks later. Mountain lion movement is thought to be strongly influenced by a knowledge of kill sites (Hanks, Hooten, and Alldredge 2015) and our analysis provides additional quantitative evidence to support that.

Similar to the mule deer results, the accumulated posterior model probabilities for the kernel forms suggested that Brownian motion and integrated Brownian motion were not





**Figure 3.** Posterior predicted mule deer (left column) and mountain lion (right column) paths  $\mu(t)$  using BMA. Top panels: Paths in two-dimensional geographic space. Middle panels: Marginal paths for easting and northing. Position uncertainty is represented as posterior realizations in the top panel and 95% marginal credible intervals in the middle two panels. Bottom panel: Warp derivative obtained via BMA.

helpful in describing mountain lion movement based on these data (i.e., posterior model probability equal to zero). However, the Gaussian, tail-up, and tail-down kernels accounted for 0.27%, 0.7295%, and 0.0005%, respectively. The distribution of posterior model probabilities for the kernel forms suggest that, while both Gaussian and tail-up kernels are useful, the tail-up kernel is twice as important for characterizing mountain lion movement. The combination of smoothing based on previous and future Brownian steps suggest that both memory and perception may influence mountain lion movement based on the data we analyzed.

#### 4. Discussion

The rapidly expanding literature on statistical models for animal movement has provided a wealth of useful tools for

analyzing telemetry data. In particular, individual-based models for trajectories have become a popular means for obtaining inference concerning the patterns and processes involved in animal movement. Despite the development of approaches for modeling animal trajectories in continuous-time, discrete-time movement models have dominated the recent literature. Heuristically, discrete-time animal movement models (e.g., Morales et al. 2004; Jonsen, Flemming, and Myers 2005; McClintock et al. 2012) are straightforward and easy to understand. They also provide a direct means to incorporate time-varying heterogeneity. In particular, temporal clustering of animal movement trajectories has proven useful for inferring animal behavior and for better understanding the spatial patterning of behavior. However, movement processes are necessarily continuous in time, but traditional continuous-time trajectory models have been oversimplified or

computationally challenging to implement for large datasets (McClintock et al. 2014).

We presented a continuous-time stochastic trajectory model that retains the natural intuition of discrete-time models, is feasible to implement for large datasets, accommodates time-varying heterogeneity in movement, and is general enough to allow for cognitive realism (e.g., memory and perception). We showed that existing stochastic process models for movement can be generalized to facilitate basis function representations of the dynamical process. Basis function representations are commonly used in spatial and spatio-temporal statistics, semiparametric modeling, and functional data analysis. We show how the same machinery used in spatial statistics to model complicated dependence structures in data can also be used in animal movement modeling. For example, temporal heterogeneity can be expressed as nonstationarity in a second-order Gaussian process model for individual trajectories using existing methods.

In previous implementations of both continuous- and discrete-time animal movement models, computation has been a bottleneck for inference based on statistical modeling. Computationally feasible model structures were often over-simplified and lacked appropriate realism. Historically, more complicated models required custom software written in compiled languages that do not facilitate user modifications. We combined multiple computational approaches to develop a strategy for fitting hierarchical stochastic process models to telemetry data. Our method exploits the substantially improved storage and multi-core processing capabilities of modern computers so that the models we presented herein can be fit on a standard laptop computer.

We showed that our modeling framework appropriately accommodates temporally heterogeneous structure in realistic animal trajectories as well as providing a natural way to handle irregular fix rates common in telemetry data. Most discrete-time trajectory models for telemetry data rely on either imputation or a linear interpolation to accommodate irregular acquisition times. For mule deer and mountain lion individuals, we showed that our model is able to accommodate mechanistic forms of temporal heterogeneity due to natural seasonality and environmental cues, providing inference concerning where and when changes in movement behavior occur.

The basis function approach we present is both familiar and further generalizable. For example, the Gaussian data and process model components in the FMM are important for computational tractability, however they can easily be generalized. In particular, other forms of satellite telemetry data (e.g., Argos) have distinctly non-Gaussian telemetry error patterns (Brost et al. 2015; McClintock et al. 2015; Buderman et al. 2016). Incorporating non-Gaussian data models in the FMM framework so that the procedure remains computationally efficient is a high priority for future research.

Furthermore, because the computational approach we presented is feasible for large datasets, the individual-level models could be nested within a population-level framework to provide inference for groups of individuals simultaneously (e.g., Hooten et al. 2016). Population-level FMMs with interacting individuals are the subject of ongoing research.

Our study system was somewhat devoid of hard boundaries to movement, but future research to formally incorporate

physical constraints to the true position process would be useful (e.g., Tracey, Zhu, and Crooks 2005; Brost et al. 2015). Statistically rigorous methods for accommodating constraints in multivariate continuous-time dynamic processes have been proposed in other fields (e.g., Cangelosi and Hooten 2009), but are still nascent in statistical approaches for telemetry data.

## Supplementary Material

Appendix A: Derivation of reparameterization of convolved multivariate Brownian motion.

Appendix B: Specification and visualization of basis functions in  $\tilde{\mathbf{H}}$ .

Appendix C: Sherman–Morrison–Woodbury identity for FMM.

Appendix D: RJMCMC algorithm for fitting FMM.

Appendix E: Simulation examples involving homogenous and heterogenous dynamics.

Appendix F: Empirical variograms for posterior predictive residuals of mule deer and mountain lion models.

## Acknowledgments

The authors thank Mat Alldredge, Frances Buderman, Marti Garlick, Ephraim Hanks, Bill Link, Brett McClintock, Leslie McFarlane, Juan Morales, Jim Powell, Henry Scharf, and Jay Ver Hoef for help for insightful discussions and feedback on the data and research. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

## Funding

Funding for this research was provided by NOAA (RWO 103), CPW (TO 1304), and NSF (DMS 1614392).

## References

- Aarts, G., Fieberg, J., and Matthiopoulos, J. (2012), “Comparative Interpretation of Count, Presence-Absence, and Point Methods for Species Distribution Models,” *Methods in Ecology and Evolution*, 3, 177–187. [579]
- Ager, A., Johnson, B., Kern, J., and Kie, J. (2003), “Daily and Seasonal Movements and Habitat use by Female Rocky Mountain Elk and Mule Deer,” *Journal of Mammalogy*, 84, 1076–1088. [584]
- Anderson-Sprecher, R., and Ledolter, J. (1991), “State-Space Analysis of Wildlife Telemetry Data,” *Journal of the American Statistical Association*, 86, 596–602. [579]
- Banerjee, S., Gelfand, A., Finley, A., and Sang, H. (2008), “Gaussian Predictive Process Models for Large Spatial Datasets,” *Journal of the Royal Statistical Society, Series B*, 70, 825–848. [582]
- Barker, R., and Link, W. (2013), “Bayesian Multimodel Inference by RJMCMC: A Gibbs Sampling Approach,” *The American Statistician*, 67, 150–156. [584]
- Barry, R., and Ver Hoef, J. (1996), “Blackbox Kriging: Spatial Prediction Without Specifying Variogram Models,” *Journal of Agricultural, Biological and Environmental Statistics*, 1, 297–322. [580]
- Berliner, L. (1996), “Hierarchical Bayesian Time Series Models,” in *Maximum Entropy and Bayesian Methods*, eds. K. Hanson, and R. Silver, Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 15–22. [578]
- Berman, M., and Turner, T. (1992), “Approximating Point Process Likelihoods with GLIM,” *Applied Statistics*, 41, 31–38. [579]
- Blackwell, P. (1997), “Random Diffusion Models for Animal Movement,” *Ecological Modelling*, 100, 87–102. [579]
- Blecha, K. (2015), “Risk-Reward Tradeoffs in the Foraging Strategy of Cougar (*Puma concolor*): Prey Distribution, Anthropogenic Development, and Patch-Selection,” M.S. Thesis, Colorado State University, Graduate Degree Program in Ecology. [585]

- Brillinger, D. (2010), "Modeling Spatial Trajectories," in *Handbook of Spatial Statistics*, eds. A. Gelfand, P. Diggle, M. Fuentes, and P. Guttorp, Boca Raton, FL: Chapman & Hall/CRC, pp. 463–475. [579]
- Brillinger, D., Preisler, H., Ager, A., and Kie, J. (2001), "The Use of Potential Functions in Modeling Animal Movement," in *Data Analysis from Statistical Foundations*, ed. E. Saleh, Huntington, NY: Nova Science Publishers, pp. 369–386. [579]
- Brost, B., Hooten, M., Hanks, E., and Small, R. (2015), "Animal Movement Constraints Improve Resource Selection Inference in the Presence of Telemetry Error," *Ecology*, 96, 2590–2597. [578,587]
- Buderman, F., Hooten, M., Ivan, J., and Shenk, T. (2016), "A Functional Model for Characterizing Long Distance Movement Behavior," *Methods in Ecology and Evolution*, 7, 264–273. [578,580,587]
- Calabrese, F., Colonna, M., Lovisolo, P., Parata, D., and Ratti, C. (2011), "Real-Time Urban Monitoring using Cell Phones: A Case Study in Rome," *IEEE Transactions on Intelligent Transportation Systems*, 12, 141–151. [578]
- Calder, C. (2007), "Dynamic Factor Process Convolution Models for Multivariate Space-Time Data with Application to Air Quality Assessment," *Environmental and Ecological Statistics*, 14, 229–247. [580]
- Cangelosi, A., and Hooten, M. (2009), "Models for Bounded Systems with Continuous Dynamics," *Biometrics*, 65, 850–856. [579,587]
- Cressie, N. (1993), *Statistics for Spatial Data: Revised Edition*, New York: Wiley. [579]
- Cressie, N., and Wikle, C. (2011), *Statistics for Spatio-Temporal Data*, New York: Wiley. [579,582]
- Damian, D., Sampson, P., and Guttorp, P. (2001), "Bayesian Estimation of Semi-Parametric Non-Stationary Spatial Covariance Structures," *Environmetrics*, 12, 161–178. [583]
- Diggle, P., and Ribeiro, P. (2002), "Bayesian Inference in Gaussian Model-Based Geostatistics," *Geographical and Environmental Modelling*, 6, 129–146. [583,585]
- Dunn, J., and Gipson, P. (1977), "Analysis of Radio-Telemetry Data in Studies of Home Range," *Biometrics*, 33, 85–101. [579]
- Evans, T., Kirchgessner, M., Eyler, B., Ryan, C., and Walter, W. (2016), "Habitat Influences Distribution of Chronic Wasting Disease in White-Tailed Deer," *The Journal of Wildlife Management*, 80, 284–291. [584]
- Farnsworth, M., Hoeting, J., Hobbs, N., and Miller, M. (2006), "Linking Chronic Wasting Disease to Mule Deer Movement Scales: A Hierarchical Bayesian Approach," *Ecological Applications*, 16, 1026–1036. [584]
- Finley, A., Banerjee, S., and Gelfand, A. (2015), "spBayes for Large Univariate and Multivariate Point-Referenced Spatio-Temporal Data Models," *Journal of Statistical Software*, 63, 1–28. [583]
- Fleming, C., Calabrese, J., Mueller, T., Olson, K., Leimgruber, P., and Fagan, W. (2014), "Non-Markovian Maximum Likelihood Estimation of Autocorrelated Movement Processes," *Methods in Ecology and Evolution*, 5, 462–472. [579]
- Fleming, C., Fagan, W., Mueller, T., Olson, K., Leimgruber, P., and Calabrese, J. (2016), "Estimating Where and How Animals Travel: An Optimal Framework for Path Reconstruction from Autocorrelated Tracking Data," *Ecology*, 97, 576–582. [579]
- Forester, J., Im, H., and Rathouz, P. (2009), "Accounting for Animal Movement in Estimation of Resource Selection Functions: Sampling and Data Analysis," *Ecology*, 90, 3554–3565. [578]
- Garlick, M., Powell, J., Hooten, M., and McFarlane, L. (2011), "Homogenization of Large-Scale Movement Models in Ecology," *Bulletin of Mathematical Biology*, 73, 2088–2108. [584]
- (2014), "Homogenization, Sex, and Differential Motility Predict Spread of Chronic Wasting Disease in Mule Deer in Southern Utah," *Journal of Mathematical Biology*, 69, 369–399. [584]
- Gelman, A. (2006), "Prior Distributions for Variance Parameters in Hierarchical Models," *Bayesian Analysis*, 1, 1–19. [583]
- Gloaguen, P., Mahevas, S., Rivot, E., Woillez, M., Guitton, J., Vermard, Y., and Etienne, M. (2015), "An Autoregressive Model to Describe Fishing Vessel Movement and Activity," *Environmetrics*, 26, 17–28. [578]
- Green, P. (1995), "Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination," *Biometrika*, 82, 711–732. [584]
- Hanks, E., Hooten, M., and Alldredge, M. (2015), "Continuous-Time Discrete-Space Models for Animal Movement," *Annals of Applied Statistics*, 9, 145–165. [584,585]
- Hanks, E., Hooten, M., Johnson, D., and Sterling, J. (2011), "Velocity-Based Movement Modeling for Individual and Population Level Inference," *PLoS One*, 6, e22795. [579]
- Higdon, D. (1998), "A Process-Convolution Approach to Modeling Temperatures in the North Atlantic Ocean," *Environmental and Ecological Statistics*, 5, 173–190. [580]
- (2002), "Space and Space-Time Modeling using Process Convolutions," in *Quantitative Methods for Current Environmental Issues*, eds. C. Anderson, V. Barnett, P. Chatwin, and A. El-Shaarawi, London: Springer-Verlag, pp. 37–56. [582]
- Hooten, M., Buderman, F., Brost, B., Hanks, E., and Ivan, J. (2016), "Hierarchical Animal Movement Models for Population-Level Inference," *Environmetrics*, 27, 322–333. [587]
- Hooten, M., Johnson, D., Hanks, E., and Lowry, J. (2010), "Agent-Based Inference for Animal Movement and Selection," *Journal of Agricultural, Biological and Environmental Statistics*, 15, 523–538. [578,584]
- Hooten, M., Johnson, D., McClintock, B., and Morales, J. (2017), *Animal Movement: Statistical Models for Telemetry Data*, Boca Raton, FL: Chapman & Hall/CRC. [578]
- Johnson, D., Hooten, M., and Kuhn, C. (2013), "Estimating Animal Resource Selection from Telemetry Data using Point Process Models," *Journal of Animal Ecology*, 82, 1155–1164. [578]
- Johnson, D., London, J., Lea, M., and Durban, J. (2008a), "Continuous-Time Correlated Random Walk Model for Animal Telemetry Data," *Ecology*, 89, 1208–1215. [579,580,581,582]
- Johnson, D., Thomas, D., Ver Hoef, J., and Christ, A. (2008b), "A General Framework for the Analysis of Animal Resource Selection from Telemetry Data," *Biometrics*, 64, 968–976. [578]
- Jonsen, I., Flemming, J., and Myers, R. (2005), "Robust State-Space Modeling of Animal Movement data," *Ecology*, 45, 589–598. [578,579,580,586]
- Kays, R., Crofoot, M., Jetz, W., and Wikelski, M. (2015), "Terrestrial Animal Tracking as an Eye on Life and Planet," *Science*, 348, aaa2478. [578]
- Knell, A., and Codling, E. (2012), "Classifying Area-Restricted Search (ARS) Using a Partial Sum Approach," *Theoretical Ecology*, 5, 325–339. [583]
- Knopff, K., Knopff, A., Kortello, A., and Boyce, M. (2010), "Cougar Kill Rate and Prey Composition in a Multiprey System," *The Journal of Wildlife Management*, 74, 1435–1447. [585]
- Laundré, J. (2013), "The Feasibility of the North-Eastern USA Supporting the Return of the Cougar *Puma concolor*," *Oryx*, 47, 96–104. [585]
- Lee, H., Higdon, D., Calder, C., and Holloman, C. (2005), "Efficient Models for Correlated Data Via Convolutions of Intrinsic Processes," *Statistical Modelling*, 5, 53–74. [580]
- McClintock, B., Johnson, D., Hooten, M., Ver Hoef, J., and Morales, J. (2014), "When to be Discrete: The Importance of Time Formulation in Understanding Animal Movement," *Movement Ecology*, 2, 21. [587]
- McClintock, B., King, R., Thomas, L., Matthiopoulos, J., McConnell, B., and Morales, J. (2012), "A General Discrete-Time Modeling Framework for Animal Movement using Multistate Random Walks," *Ecological Monographs*, 82, 335–349. [579,586]
- McClintock, B., London, J., Cameron, M., and Boveng, P. (2015), "Modelling Animal Movement using the Argos Satellite Telemetry Location Error Ellipse," *Methods in Ecology and Evolution*, 6, 266–277. [578,587]
- McFarlane, L. (2007), "Breeding Behavior and Space Use of Male and Female Mule Deer: An Examination of Potential Risk Differences for Chronic Wasting Disease Infection," M.S. Thesis, Utah State University, Department of Wildland Resources. [584]
- Miller, M., Hobbs, N., and Taverer, S. (2006), "Dynamics of Prion Disease Transmission in Mule Deer," *Ecological Applications*, 16, 2208–2214. [584]
- Morales, J., Haydon, D., Friar, J., Holsinger, K., and Fryxell, J. (2004), "Extracting More Out of Relocation Data: Building Movement Models as Mixtures of Random Walks," *Ecology*, 85, 2436–2445. [578,579,583,586]

- Nicholson, M., Bowyer, R., and Kie, J. (1997), "Habitat Selection and Survival of Mule Deer: Tradeoffs Associated With Migration," *Journal of Mammalogy*, 78, 483–504. [585]
- Nychka, D., and Saltzman, N. (1998), "Design of Air-Quality Monitoring Designs," in *Case Studies in Environmental Statistics*, eds. D. Nychka, W. Piegorsch, and L. Cox, New York: Springer, pp. 51–76. [582]
- Nychka, D., Wikle, C., and Royle, J. (2002), "Multiresolution Models for Nonstationary Spatial Covariance Functions," *Statistical Modelling*, 2, 315–331. [582]
- Paciorek, C., and Schervish, M. (2006), "Spatial Modelling Using a New Class of Nonstationary Covariance Functions," *Environmetrics*, 17, 483–506. [582]
- Pierce, B., Bleich, V., Wehausen, J., and Bowyer, R. (1999), "Migratory Patterns of Mountain Lions: Implications for Social Regulation and Conservation," *Journal of Mammalogy*, 80, 986–992. [584]
- Preisler, H., Ager, A., Johnson, B., and Kie, J. (2004), "Modeling Animal Movements Using Stochastic Differential Equations," *Environmetrics*, 15, 643–657. [579]
- Protter, P. (2004), *Stochastic Integration and Differential Equations* (2nd ed.), Berlin, Germany: Springer-Verlag. [579]
- Ramsay, J. (1988), "Monotone Regression Splines in Action," *Statistical Science*, 3, 425–441. [581]
- Rue, H., and Held, L. (2005), *Gaussian Markov Random Fields: Theory and Applications*, Boca Raton, FL: Chapman & Hall/CRC. [580]
- Sampson, P. (2010), "Constructions for Nonstationary Spatial Processes," in *Handbook of Spatial Statistics*, eds. A. Gelfand, P. Diggle, M. Fuentes, and P. Guttorp, Boca Raton, FL: Chapman & Hall/CRC. [582]
- Sampson, P., and Guttorp, P. (1992), "Nonparametric Estimation of Nonstationary Spatial Covariance Structure," *Journal of the American Statistical Association*, 87, 108–119. [579,583]
- Schmidt, A., and O'Hagan, A. (2003), "Bayesian Inference for Non-Stationary Spatial Covariance Structure via Spatial Deformations," *Journal of the Royal Statistical Society, Series B*, 65, 743–758. [583]
- Shepp, L. (1966), "Radon-Nikodym Derivatives of Gaussian Measures," *The Annals of Mathematical Statistics*, 37, 321–354. [580]
- Tracey, J., Zhu, J., and Crooks, K. (2005), "A Set of Nonlinear Regression Models for Animal Movement in Response to a Single Landscape Feature," *Journal of Agricultural, Biological and Environmental Statistics*, 10, 1–18. [587]
- Ver Hoef, J., and Peterson, E. (2010), "A Moving Average Approach for Spatial Statistical Models of Stream Networks," *Journal of the American Statistical Association*, 105, 6–18. [582]
- Warton, D., and Shepherd, L. (2010), "Poisson Point Process Models Solve the "Pseudo-Absence Problem" for Presence-Only Data in Ecology," *Annals of Applied Statistics*, 4, 1383–1402. [579]
- Wecker, W., and Ansley, C. (1983), "The Signal Extraction Approach to Nonlinear Regression and Spline Smoothing," *Journal of the American Statistical Association*, 78, 81–89. [580]
- Wikelski, M., and Kays, R. (2015), "Movebank: Archive, Analysis and Sharing of Animal movement Data," available at <http://www.movebank.org>. [578]
- Wikle, C. (2010), "Low-Rank Representations for Spatial Processes," in *Handbook of Spatial Statistics*, eds. A. Gelfand, P. Diggle, M. Fuentes, and P. Guttorp, Boca Raton, FL: Chapman & Hall/CRC, pp. 107–118. [582]
- Wikle, C., and Hooten, M. (2010), "A General Science-Based Framework for Nonlinear Spatio-Temporal Dynamical Models," *Test*, 19, 417–451. [582]
- Young, S., and Goldman, E. (1946), *The Puma: Mysterious American Cat*, New York: Dover Publications. [585]