

## A functional model for characterizing long-distance movement behaviour

Frances E. Buderman<sup>1\*</sup>, Mevin B. Hooten<sup>1,2,3,4</sup>, Jacob S. Ivan<sup>5</sup> and Tanya M. Shenk<sup>6</sup>

<sup>1</sup>Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO, 80523-1484, USA; <sup>2</sup>U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO, 80523-1484, USA; <sup>3</sup>Department of Statistics, Colorado State University, Fort Collins, CO, 80523-1484, USA; <sup>4</sup>Graduate Degree Program in Ecology, Colorado State University, Fort Collins, CO, 80523-1484, USA; <sup>5</sup>Colorado Parks and Wildlife, Fort Collins, CO, 80526, USA; and <sup>6</sup>National Park Service, Fort Collins, CO, 80525, USA

### Summary

1. Advancements in wildlife telemetry techniques have made it possible to collect large data sets of highly accurate animal locations at a fine temporal resolution. These data sets have prompted the development of a number of statistical methodologies for modelling animal movement.
2. Telemetry data sets are often collected for purposes other than fine-scale movement analysis. These data sets may differ substantially from those that are collected with technologies suitable for fine-scale movement modelling and may consist of locations that are irregular in time, are temporally coarse or have large measurement error. These data sets are time-consuming and costly to collect but may still provide valuable information about movement behaviour.
3. We developed a Bayesian movement model that accounts for error from multiple data sources as well as movement behaviour at different temporal scales. The Bayesian framework allows us to calculate derived quantities that describe temporally varying movement behaviour, such as residence time, speed and persistence in direction. The model is flexible, easy to implement and computationally efficient.
4. We apply this model to data from Colorado Canada lynx (*Lynx canadensis*) and use derived quantities to identify changes in movement behaviour.

**Key-words:** Argos, Bayesian model, Canada lynx, functional data analysis, movement modelling, splines, telemetry

### Introduction

Data sets consisting of animal locations are often collected for purposes other than movement analysis (e.g., survival analysis, demographic studies; White & Shenk 2001; Winterstein, Pollock & Bunck 2001) or with technology that prohibits long-term fine-scale movement modelling (Yasuda & Arai 2005). For example, radiotelemetry may be used to estimate survival (Cowen & Schwarz 2005), but the locations may not be used in the analysis (e.g., Buderman *et al.* 2014; Hightower, Jackson & Pollock 2001). These data sets are costly and time-consuming to collect, but often contain a wealth of unused spatial information. The ability to spatially characterize movement behaviours using data sets that are insufficient for fine-scale movement modelling may help management and conservation agencies identify critical areas for wildlife movement (Berger 2004). In addition, with appropriate temporal data, researchers can also better understand mechanisms that regulate movement behaviour (Hays *et al.* 2014; Scott, Marsh & Hays 2014).

Runge *et al.* (2014) divide long-distance movements into four categories: irruption (dispersal), migration, nomadism and intergenerational relays (which we do not address). Such movement behaviour can vary among individuals and over an individual's lifetime, though some species may be more inclined to exhibit one kind of long-distance movement behaviour (LDMB) over another (Jonzén *et al.* 2011; Mueller *et al.* 2011; Singh *et al.* 2012). For most organisms, the causes and costs of dispersal will vary by individual and in space and time (Bowler & Benton 2005), resulting in a continuum of movement behaviours (Jonzén *et al.* 2011). LDMB may contribute substantially to population dynamics because it is the main determinant of population spread and colonization rates (Greenwood & Harvey 1982; Shigesada & Kawasaki 2002). Thus, LDMB is an important life-history trait for many processes such as species invasions, range shifts and local extinctions, reintroduction programmes, metapopulation dynamics, connectivity and gene flow (Trakhtenbrot *et al.* 2005).

The spatial location of these behaviours could inform conservation efforts for species capable of long-distance movements, as some behaviours may be more important than others for population persistence (Runge *et al.* 2014). In

\*Corresponding author. E-mail: franny.buderman@colostate.edu

addition, comparing contemporary movement data with properly analysed historical data may identify changes in movement behaviour resulting from natural and anthropogenic disturbances. Changes in migratory behaviour could have wide-ranging consequences in cases where the species contributes significantly to the biological assemblage (Robinson *et al.* 2009). Species are usually limited in their range by dispersal ability, foraging ecology or available habitat (Hays & Scott 2013; Wood & Pullin 2002), and as habitat fragmentation and climate variability increase, the ability of species to traverse long distance will become critical (Bowler & Benton 2005). Species that have the capability for long-distance movement may be able to track habitat as environmental conditions change. However, individuals usually depend on a network of suitable habitats for different behaviours (e.g., breeding or migration; Robinson *et al.* 2009). Long-term survival of the species can be reduced when the distance between patches exceeds dispersal ability (Trakhtenbrot *et al.* 2005), or when suitable habitat is not available for all of the behaviours that occur during an annual cycle (Robinson *et al.* 2009).

Although dispersal, migration and nomadism are all LDMBs, they may differ in characteristics that can be quantitatively measured, such as residence time, speed and persistence in direction (described using the turning angle). For example, areas where individuals are foraging or maintaining a home range may be identified by longer residence times or slower speeds (Schofield *et al.* 2013) and undirected motion (Morales *et al.* 2004). In contrast, movement may be faster (Dickson, Jenness & Beier 2005) and more directed (Haddad 1999) within corridors. Nomadic individuals may exhibit similar speeds as migrators and dispersers, but they would appear to be perpetually dispersing, with no consistent activity centre and a turning angle independent of previous movements (Lidicker & Stenseth 1992). Dispersal and migration may have similar speed and directional characteristics, but migration is a seasonally repeated movement between the same areas (Berger 2004) by individuals within a population (Sawyer, Lindzey & McWhirter 2005), whereas dispersal is a one-way movement (Lidicker & Stenseth 1992).

Movement behaviour is typically monitored using very high frequency (VHF) or satellite telemetry devices. These monitoring devices are more effective at detecting LDMB than plot-based studies, which may underestimate long-distance movement (Koenig *et al.* 1996). The frequency of VHF data is determined by how often an individual can be located and are spatially restricted to the actively searched area. Aerial location accuracy associated with VHF data may be affected by antenna type, altitude and observer skill, while ground triangulation accuracy may be additionally impacted by terrain, vegetation, power lines, and weather (Mech 1983). In contrast, the intended fix rate for a satellite telemetry device is pre-programmed and often regularly spaced in time. Fix success rates and accuracy can be influenced by animal behaviour, such as diving behaviour, canopy cover, terrain and climatic conditions (e.g., Di Orio, Callas & Schaefer 2003; Dujon, Lindstrom & Hays 2014; Heard, Ciarniello & Seip 2008; Mattisson *et al.* 2010). The device's satellite system (GPS or Argos Satellite

maintained by Service Argos) can also influence accuracy of the location observations (Costa *et al.* 2010; Dujon, Lindstrom & Hays 2014; Heard, Ciarniello & Seip 2008; Patterson *et al.* 2010; Vincent *et al.* 2002). In addition, fix success rate, battery life and accuracy may all depend on transmitter manufacturer and model. Both VHF and satellite components can be placed into the same device or individuals can be outfitted with two separate devices, resulting in data sets consisting of multiple data types.

Movement modelling often seeks to spatially characterize an individual's location as a function of time; however, this function may be highly complex and non-stationary. In addition, measurement error varies among monitoring methods and can be large enough to overwhelm small-scale movement patterns (Breed *et al.* 2011; Kuhn *et al.* 2009). Coupled with temporal irregularity and missing data, these attributes may prohibit the use of contemporary movement models. We have found that many available methods do not readily accommodate multiple sources of data and must impute missing data to obtain locations at regular intervals (e.g., Hanks *et al.* 2011; Hanks, Hooten & Alldredge 2015; Hooten *et al.* 2010; Johnson, London & Kuhn 2011). For example, the continuous-time correlated random walk model presented by Johnson *et al.* (2008) only accounts for elliptical error distributions. Breed *et al.* (2012) incorporated an augmented particle smoother into a CRW process model to allow for time-varying parameters; however, their method does not account for multiple data sources and its effectiveness was only demonstrated on highly accurate GPS data at a fine temporal scale (10–30 locations day<sup>-1</sup>). Winship *et al.* (2012) incorporated multiple data sources (Argos, GPS and geolocation data) into a state-space model, but the method performed poorly when there were data gaps, relied heavily on the estimates of Argos precision presented in Jonsen, Flemming & Myers (2005) and treated the GPS data as equivalent to the best Argos location class. Change-point models require specifying or estimating the number of change points, and the change points are discrete in time (Gurarie, Andrews & Laidre 2009; Hanks *et al.* 2011; Jonsen, Flemming & Myers 2005; Jonsen, Myers & James 2007); modelling smooth transitions in the change-point framework is more difficult. Given that one individual may exhibit many different LDMBs, we seek a model that is flexible enough to detect different types and degrees of movement behaviour, without specifying or estimating the number of change points. Brownian bridge movement models, a method commonly used with high-resolution telemetry data, have been shown to work well only when the measurement error is negligible (Pozdnyakov *et al.* 2014), making them unsuitable for data sets obtained with VHF or Argos technology, which can be subject to substantial error. Recent applications of wavelet analyses also do not account for location error or uncertainty in the change-point identification and are not feasible with sparse and irregular data sets (Lavielle 1999; Sur *et al.* 2014).

Basis functions are a useful set of tools for approximating continuous functions, such as movement paths, when ordinary polynomials are inadequate to describe the behaviour of the function (Rice 1969). Commonly used basis functions include

wavelets, Fourier series and splines. Approximating a function with splines is computationally easy because the function is just a weighted sum of simpler functions (Wold 1974), and such tools have been incorporated into standard statistical software. Wold (1974) recognized that splines may be most useful in low information settings where the ultimate goal is to compare individual estimates of a few characteristic parameters that describe the curve. Basis functions have been used extensively in fields such as physics (e.g., Sapirstein & Johnson 1996), medicine (e.g., Gray 1992) and medical imaging (e.g., Carr, Fright & Beatson 1997), and climate science (e.g., Sáenz-Romero et al. 2010). However, basis functions and associated statistical methods are less commonly used in ecology. Most applications focus on modelling species distributions (e.g., Lawler *et al.* 2006; Leathwick *et al.* 2005) and population dynamics (e.g., Bjørnstad *et al.* 1999), though splines have broad applicability in generalized additive models (Hastie & Tibshirani 1990; Wood & Augustin 2002). For example, Hanks, Hooten & Aldredge (2015) used B-splines to model spatial transition rates as a function of location and direction-based covariates and time-varying coefficients. In addition, recent efforts have used B-splines to estimate density functions associated with movement-related behavioural states (Langrock *et al.* 2014). Tremblay *et al.* (2006) used Bezier, hermite and cubic splines as strict interpolators of irregular telemetry data from ocean-obligate species; however, they assumed the filtered Argos locations were the true locations. There is also precedent in the statistical literature for the equivalence between stochastic movement processes, such as the Wiener process, and smoothing polynomial splines (Wahba 1978; Wecker & Ansley 1983).

We describe a functional approach to movement modelling using basis functions within a Bayesian model that accounts for multiple data types and their associated error, recognizing that the observed locations are not the true location. The basis functions allow us to account for temporal variation in the continuous underlying movement path without specifying movement mechanisms. We then use derived quantities, such as residence time, speed and persistence in direction, to characterize movement behaviour. In addition, the model is multi-scale, allowing for movement behaviour at multiple biologically relevant temporal scales. We use this model to describe how reintroduced Canada lynx (*Lynx canadensis*) moved throughout Colorado. The two data collection methods, along with their measurement error and the sampling irregularity, make this an ideal data set to demonstrate the utility of our model.

## Methods

Conventional functional data analysis (FDA) assumes that there is a continuous underlying process, but the observations are temporally discrete, may be subject to error and are temporally irregular (Ramsay & Dalzell 1991; Ramsay & Silverman 2002; Ramsay & Silverman 2005). Unlike traditional time series analysis, FDA does not assume stationarity or regularity of time intervals (Levitin *et al.* 2007). The continuous function of interest is approximated using basis functions, which are a set of patterns that capture the main shape of the curve (Ferraty & Vieu

2006; Hastie, Tibshirani & Friedman 2009; Ramsay & Silverman 2005). In our case, different sets of basis functions account for complexity in the process at different temporal scales, allowing us to detect both large- and small-scale movement. In addition, FDA is useful when the objectives of an analysis are to estimate the derivatives of a function (Levitin *et al.* 2007; Ramsay & Dalzell 1991). In our framework, functions of temporal derivatives, such as residence time, speed and persistence in direction, are derived quantities that can characterize the movement path. The Bayesian framework allows for inference concerning these derived quantities and their associated uncertainty while incorporating multiple data sources; for our purposes, we incorporated VHF and Argos data into a single model.

## DATA MODEL

We consider each observed (centred and scaled) location,  $\mathbf{s}_j(t)$  for a time  $t \in \mathcal{T}$  associated with data type  $j$  ( $j = 1, \dots, 6$  are Argos error classes and  $j = 7$  denotes VHF), to arise from a multivariate normal mixture model with mean,  $\mathbf{z}(t)$ , representing the true location at time  $t$  and a covariance matrix  $\Sigma_j$  such that

$$\mathbf{s}_j(t) \sim \begin{cases} \mathbf{N}(\mathbf{z}(t), \Sigma_j), & \text{if } w_j(t) = 1 \\ \mathbf{N}(\mathbf{z}(t), \tilde{\Sigma}_j), & \text{if } w_j(t) = 0 \end{cases} \quad \text{eqn 1}$$

The covariance matrix,  $\Sigma_j \equiv \sigma_j^2 \mathbf{R}_j$ , represents the error variance associated with each data type where the correlation matrix is

$$\mathbf{R} \equiv \begin{bmatrix} 1 & \sqrt{c} \rho \\ \sqrt{c} \rho & c \end{bmatrix} \quad \text{eqn 2}$$

for  $j = 1, \dots, 6$  and  $\mathbf{R} \equiv \mathbf{I}$  for  $j = 7$ . The prior distribution for the measurement error variance,  $\sigma_j^2$ , was modelled as an inverse gamma,  $\text{IG}(q, r)$ , where  $q$  is the shape parameter and  $r$  is the rate parameter. Argos error for all error classes has been shown to be larger than reported by Argos and greater in the longitudinal direction (Boyd & Brightsmith 2013; Costa *et al.* 2010; Hoenner *et al.* 2012); therefore, we use the parameter  $c$ , where  $c \sim \text{Beta}(\alpha_c, \beta_c)$ , to scale the error variance to be less in latitude than longitude. The  $\rho$  parameter scales the degree of covariance between latitude and longitude and is modelled as  $\text{Beta}(\alpha_\rho, \beta_\rho)$ .

The indicator  $w_j(t)$  determines which mixture component gives rise to the observed location and is modelled as  $\text{Bern}(0.5)$ . The covariance matrix of the rotated distribution,  $\tilde{\Sigma}_j$ , is calculated as  $\mathbf{H}_j \Sigma_j \mathbf{H}_j'$  where  $\mathbf{H}_j$  is a transformation matrix equal to

$$\mathbf{H} \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{eqn 3}$$

for  $j = 1, \dots, 6$ , and  $\mathbf{H} \equiv \mathbf{I}$  (the identity matrix) for  $j = 7$ . The mixture model accounts for the fact that Argos error locations do not follow a symmetric distribution around the true location, but are more likely to be found in an X-pattern, due to the polar orbit of the satellites (Costa *et al.* 2010; Douglas *et al.* 2012). In preliminary analyses not presented here, the multivariate normal mixture model fit the data better than a multivariate normal non-mixture model. Argos locations are commonly modelled with a  $t$ -distribution to account for extreme outliers (following Jonsen, Flemming & Myers 2005), however, the mixture model allows us to model anisotropic outliers. Though the aforementioned studies have modelled or estimated Argos error, the information is not directly applicable in the form of priors because the mixture model is a novel method for modelling Argos error and there is significant variability in reported estimates of Argos error (Costa *et al.* 2010). Beginning in 2011, the Argos system implemented a new algorithm that provides an error ellipse, as opposed to a radius, for each location (Lopez *et al.* 2014). Recent work by McClintock *et al.* (2014b) used the

ellipse parameters provided by the Argos system and a bivariate normal distribution to model the data.

#### PROCESS MODEL

In the FDA paradigm, a continuous process for a set of times  $t$  ( $t \in \mathcal{T}$ ) is written as an expansion of  $M$  basis functions of order  $k$ :

$$z(t) = \sum_{m=1}^M c_m \phi_{m,k}(t) \quad \text{eqn 4}$$

where  $z(t)$  is the curve of interest,  $c_m$  is a coefficient that determines the weight of each basis function in the construction of the curve, and  $\phi_m(t)$  is a particular basis function (Levitin *et al.* 2007). The type of pattern present in the data dictates the best choice of basis function; for example, splines are often used for non-periodic data, Fourier series for periodic data, and wavelet bases for data with sharp localized patterns. We employed the B-spline basis, which is commonly used in semiparametric regression, because it has local support and stable numerical properties when the number of knots (the points at which the basis functions connect) is large (Keele 2008; Ruppert, Wand & Carroll 2003). However, the model we present is general enough to accommodate any type of basis functions. B-spline basis functions are defined recursively according to the Cox-de Boor formula (see De Boor 1978). Let  $x_{m,k}$  denote the  $m$ th B-spline basis function of order  $k$  (cubic B-splines are 4th order and 3rd degree) for the knot sequence  $\tau$ , where  $k \leq K$ . Then for  $m=1, \dots, N+2K-k$ ,

$$x_{m,k}(t) = \frac{t - \tau_m}{\tau_{m+k-1} - \tau_m} B_{m,k-1}(t) + \frac{\tau_{m+k} - t}{\tau_{m+k} - \tau_{m+1}} B_{m+1,k-1}(t), \quad \text{eqn 5}$$

where  $N$  is the number of interior knots (Hastie, Tibshirani & Friedman 2009).

In the spatial statistics and signal processing framework, a continuous stochastic process is often written as a convolution, or a moving average, of a smoothing kernel function,  $k(\tau-t)$  and a latent process (e.g., white noise),  $\eta(\tau)$ :

$$z(t) = \int_{\mathcal{T}} k(\tau - t) \eta(\tau) d\tau \quad \text{eqn 6}$$

for  $\tau \in \mathcal{T}$  (Calder 2007; Higdon 2002; Lee *et al.* 2002). When discretized, (6) takes on a general formulation (4) (Calder 2007; Higdon 2002; Lee *et al.* 2002). Non-stationary processes can be modelled by allowing the kernel to be a function of time (or space) and not just distance (Cressie & Wikle 2011; Higdon 2002; Higdon, Swall & Kern 1999). In the context of animal movement, one can consider the smoothing kernel as some function that imposes temporal dependence on the observed locations (the latent process) to create a continuous and smooth movement path.

In our case, the location of an individual at time  $t$  in each direction,  $z(t)$ , is a function of an individual's geographic mean in that direction,  $\beta_0$  and the summation of  $M$  cubic B-splines evaluated at time  $t$ ,  $x_{m,4}(t)$ , and the regularized, direction-specific coefficient,  $\beta_m$ , for that B-spline. The location in longitude and latitude is

$$z_{lon}(t) = \beta_{0_{lon}} + \sum_{m=1}^M x_{m,4}(t) \beta_{m_{lon}} \quad \text{eqn 7}$$

$$z_{lat}(t) = \beta_{0_{lat}} + \sum_{m=1}^M x_{m,4}(t) \beta_{m_{lat}} \quad \text{eqn 8}$$

Using matrix notation, we can write (7) and (8) jointly as

$$\mathbf{z}(t) = \beta_0 + \mathbf{X}(t)\boldsymbol{\beta} \quad \text{eqn 9}$$

where  $\mathbf{z}(t)$  is a vector describing the location in space at time  $t$ . The matrix  $\mathbf{X}(t)$  is a 2-by- $2M$  matrix where  $\mathbf{x}(t)'$  is a row vector containing all of the B-splines evaluated at time  $t$ , such that

$$\mathbf{X}(t) \equiv \begin{pmatrix} \mathbf{x}(t)' & \mathbf{0}' \\ \mathbf{0}' & \mathbf{x}(t)' \end{pmatrix} \quad \text{eqn 10}$$

As such, it can be multiplied by a single  $2M$ -by-1 vector of regularized coefficients

$$\boldsymbol{\beta} \equiv \begin{pmatrix} \beta_{lon} \\ \beta_{lat} \end{pmatrix} \quad \text{eqn 11}$$

The regularized coefficients for higher-order splines are not generally interpreted (Weisberg 2014), but can be thought of as the contribution, or the directional forcing, of that basis function to the process at that time. The intercept,  $\beta_0$ , can be interpreted as the geographic centre of mass for each individual, for which we specified a relatively uninformative 2-dimensional normal prior (Appendix S1). We specified a normal prior with mean  $\mathbf{0}$  and covariance matrix  $\Sigma_{\beta}$  for the coefficients such that

$$\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\beta}) \quad \text{eqn 12}$$

We selected three sets of B-splines and varied the number of knots to align with temporal scales we believe are biologically important for lynx movement: year, season (3 months) and month. Including multiple sets of basis functions allows the continuous function to capture behaviour at different temporal scales without losing predictive capability when there is an absence of fine-scale temporal data. However, the required number of knots results in a large design matrix of coefficients that is difficult to visualize; for example, there were 36 and 41 basis functions for the two Canada lynx presented in the case study. The number of basis functions will increase as the length of the time series increases. We used the covariance matrix

$$\Sigma_{\beta} \equiv \begin{pmatrix} \sigma_{\beta_{lon}}^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_{\beta_{lat}}^2 \mathbf{I} \end{pmatrix} \quad \text{eqn 13}$$

as a regulator in the ridge regression framework to shrink the  $\boldsymbol{\beta}$  coefficients. The variance terms,  $\sigma_{\beta_{lon}}^2$  and  $\sigma_{\beta_{lat}}^2$ , control the smoothing in each dimension; a very small variance leads to underfitting, whereas a large variance can lead to overfitting (Eilers & Marx 1996). We selected the variance components by calculating the Deviance Information Criterion (DIC; Spiegelhalter *et al.* 2002) over 10 000 MCMC iterations and optimizing the DIC over 400 pairs of variance components (Appendix S2). In simulation, we found that DIC and K-fold cross-validation methods performed similarly. The details of regularization and ridge regression are beyond the scope of this paper and are explored in more detail in (Hastie, Tibshirani & Friedman 2009) and Hooten & Hobbs (2015).

The model described above yields the posterior distribution

$$[\beta_0, \boldsymbol{\beta}, \sigma^2, \rho, c, \mathbf{w} | \mathbf{S}] \propto \prod_{j=1}^J \prod_{t \in \mathcal{T}} [s_j(t) | \beta_0, \boldsymbol{\beta}, \sigma_j^2, \rho, c, w_j(t)] [\beta_0] [\boldsymbol{\beta}] [\sigma^2] [\rho] [c] [\mathbf{w}] \quad \text{eqn 14}$$

where  $\sigma^2 \equiv [\sigma_1^2, \dots, \sigma_J^2]$ ,  $\mathbf{w}$  is a vector of the indicators  $w_j(t)$ , and  $\mathbf{S}$  is a matrix of observed locations. This is the form of a typical 'integrated' model where multiple data sources provide information about the same underlying processes. Similar multidata source models have become popular in demographic studies (e.g., Barker 1997; Burnham 1993; Nasution *et al.* 2001; Schaub & Abadi 2011), but have not been as common in movement studies (but see Winship *et al.* 2012). If inference for multiple individuals is desired, the data model can be shared among individuals while the process model parameters ( $\beta_0$ ,  $\boldsymbol{\beta}$ ) and regulator

( $\Sigma_{\beta}$ ) can be allowed to vary by individual. This model can be extended to account for additional stochasticity using a first-order Gaussian process, such that  $\mathbf{z}(t) \sim N(\beta_0 + \mathbf{X}(t)\beta, \sigma_{\mu}^2 \mathbf{R})$ , where  $\sigma_{\mu}^2 \mathbf{R}$  accounts for process error separately. Such Gaussian processes are commonly used as statistical emulators of complicated nonlinear mechanistic models (Hooten *et al.* 2011; O’Hagan & Kingman 1978).

See Appendix S1 for prior specifications. The model was fit using Markov chain Monte Carlo (MCMC), and a Gibbs sampler was constructed to sample from the posterior using the full-conditional distributions for all parameters except  $\rho$  and  $c$ , because they were not conjugate. Metropolis-Hastings was used to sample  $\rho$  and  $c$ . See Appendix S3 for R code (R Core Team 2013).

CHARACTERIZING MOVEMENT

We are interested in quantities derived from  $\mathbf{z}(t)$  that can be used as movement descriptors. We describe three relevant derived quantities; however, our framework can be extended to other systems and conservation questions by modifying these quantities. These derived quantities represent the physical outcome in the movement path from various movement behaviours. The Bayesian framework allows us to obtain inference for derived quantities through Monte Carlo integration. We can visualize these quantities both temporally and spatially. All quantities are calculated in the MCMC algorithm using techniques described in Appendix S4 (spatial quantities) and Appendix S5 (temporal quantities).

To describe the quantities of interest spatially, we define a grid of equally sized regions,  $\mathcal{A}_l$  for  $l = 1, \dots, L$ , that comprise the area for which we desire inference. This method is similar to that used by Johnson, London & Kuhn (2011) to describe diving behaviour of northern fur seals (*Callorhinus ursinus*). The first derived quantity we describe is residence time,  $r_l$  and is calculated on each MCMC iteration as a per area frequency of locations in region  $\mathcal{A}_l$ :

$$r_l = \lim_{\Delta t \rightarrow 0} \sum_{t \in \mathcal{T}} \Delta t I_{\{\mathbf{z}(t) \in \mathcal{A}_l\}} \tag{eqn 15}$$

where the indicator  $I$  identifies whether location  $\mathbf{z}(t)$  was in region  $\mathcal{A}_l$ .

The second derived quantity of interest is speed. To calculate the average speed per unit of area, we first need the velocity between the location at time  $\mathbf{z}(t)$  and the location at time  $\mathbf{z}(t - \Delta t)$ . When  $\Delta t$  is sufficiently small, the first derivative of  $\mathbf{z}(t)$  with respect to  $t$  can be approximated by

$$\frac{d\mathbf{z}(t)}{dt} \approx \delta(t) \tag{eqn 16}$$

where

$$\delta(t) = \frac{\mathbf{z}(t) - \mathbf{z}(t - \Delta t)}{\Delta t} \tag{eqn 17}$$

In practice,  $\Delta t$  is constant for the entire time series, and velocity is related to speed  $v(t)$  such that

$$v(t) = \sqrt{\delta(t)' \delta(t)} \tag{eqn 18}$$

The average speed in  $\mathcal{A}_l$ , given a positive residence time, is

$$\bar{v}_l = \frac{\lim_{\Delta t \rightarrow 0} \sum_{t \in \mathcal{T}} \Delta t v(t) I_{\{\mathbf{z}(t) \in \mathcal{A}_l\}}}{r_l} \tag{eqn 19}$$

A large average speed describes areas where the individual was moving quickly and spending little time. Therefore, large average speeds (19) identify areas that individuals may use to travel.

Persistence in direction is the third metric of interest and may be useful for describing directed, as opposed to nomadic, movement. We can describe persistence in direction by deriving the turning angle,  $\theta$ , using the velocity calculated in (17),

$$\theta(t) = \left| \arccos \left( \frac{\delta(t+1)\delta(t)}{\sqrt{\delta(t)\delta(t)}\sqrt{\delta(t+1)\delta(t+1)}} \right) \right| \tag{eqn 20}$$

Given that residence time is positive, the average turning angle,  $\theta(t)$ , in region  $\mathcal{A}_l$  is

$$\bar{\theta}_l = \frac{\lim_{\Delta t \rightarrow 0} \sum_{t \in \mathcal{T}} \Delta t \theta(t) I_{\{\mathbf{z}(t) \in \mathcal{A}_l\}}}{r_l} \tag{eqn 21}$$

Alternatively, we can describe these quantities temporally, negating the need for a spatially defined grid. This decreases computation time and allows the quantities to be visualized temporally and spatially. Speed and persistence in direction can be calculated as they were in (18) and (20) and residence time can be calculated as the inverse of speed:

$$r(t) = \frac{1}{v(t)} \tag{eqn 22}$$

CASE STUDY: CANADA LYNX REINTRODUCTION IN COLORADO

Colorado Division of Wildlife (now Colorado Parks and Wildlife) initiated a reintroduction programme for Canada lynx (*Lynx canadensis*) in 1997. Between 1999 and 2006, 218 wild-caught lynx from Alaska, Yukon Territory, British Columbia, Manitoba and Quebec were released in the San Juan Mountains within 40 km of the Rio Grande Reservoir (Devineau *et al.* 2010). Individuals were fitted with either VHF collars (Telonics<sup>TM</sup>, Mesa, AZ, USA) that were active for 12 h per day or satellite/VHF collars (Sirtrack<sup>TM</sup>, Havelock North, New Zealand) that were active for 12 h per week with locations obtained using the Argos system (Devineau *et al.* 2010). Weekly airplane flights were conducted over a 20 684 km<sup>2</sup> area, which included the reintroduction area and surrounding high-elevation sites (>2591 m; Devineau *et al.* 2010); attempts were made to locate each VHF-collared individual in the study area once every 2 weeks. Additional flights outside of the study area were conducted when feasible and during denning season (Devineau *et al.* 2010). Accuracy of VHF locations was self-reported as 50–500 m (Devineau *et al.* 2010). Irregular location data were obtained from 1999 to 2011 due to one or both of the transmitter components failing, logistical constraints or movement out of the study area precluding VHF data collection. Therefore, data for each individual vary in the length of the time series, the temporal regularity of locations and the number of locations from each data type and error class. We have analysed the telemetry data from two Canada lynx (Appendix S6).

We obtained 10 000 MCMC iterations, with a burn-in period of 1000 iterations. All data used in this paper are available in Appendix S7. Additional results from fitting the model to simulated data are available in Appendix S8.

Results

To visualize the fit of the model to the data, we calculated standard posterior quantities, such as means and 95% credible intervals for the marginal location in each direction (Fig. 1a, b). Increasing uncertainty is evident during long periods of missing data (Fig. 1a, b). The derived quantities were scaled relative to the maximum value for that quantity over the indi-

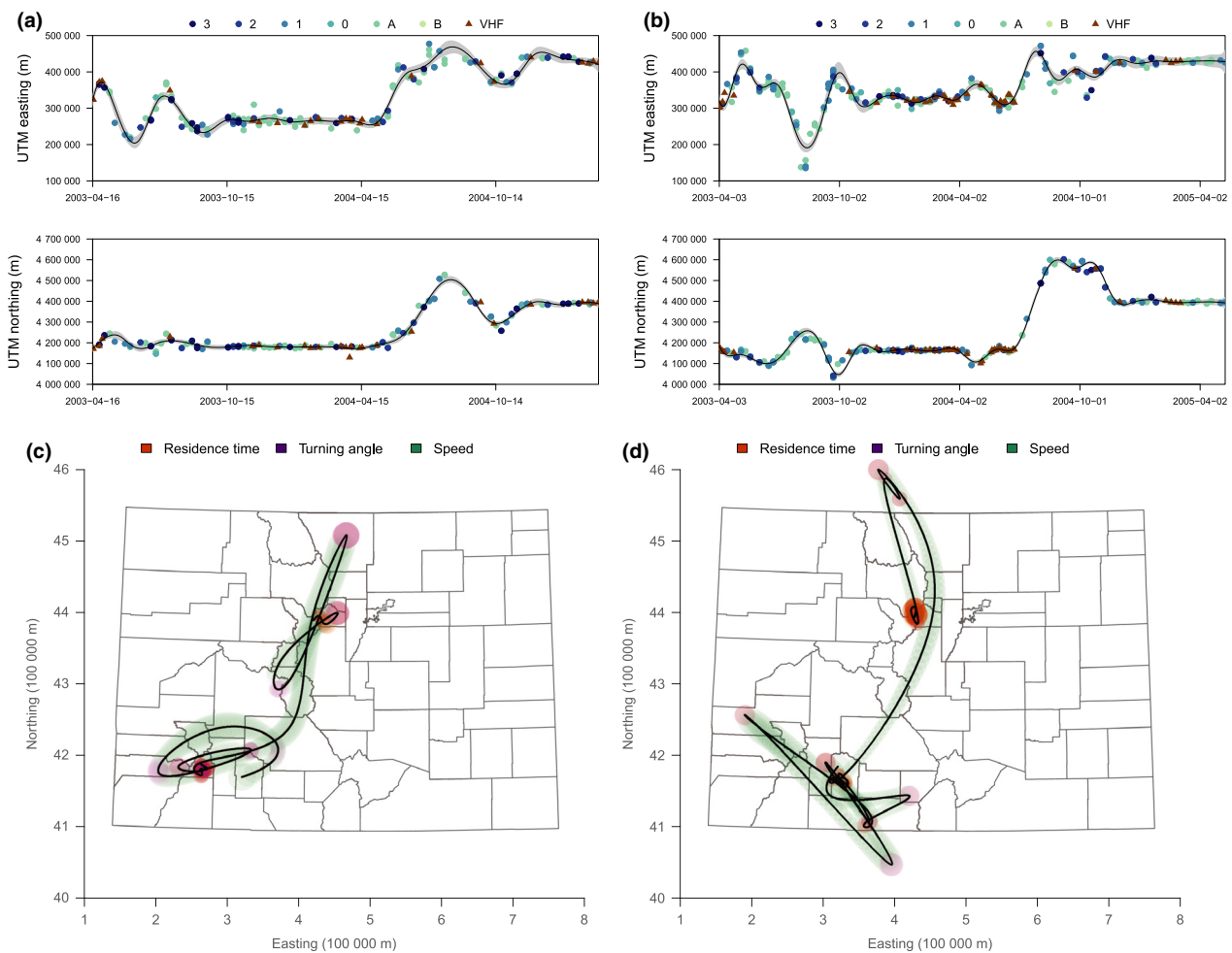
vidual's lifetime and plotted both spatially, on a map of Colorado (Figs 1c, d, and 2), and temporally (Fig. 2). These relative values are useful for visualizing the degree of each behaviour at a given time point, despite the quantities having different units; the degree of shading represents the strength of that behaviour, with the size corresponding to the spatial uncertainty (Figs 1c, d, and 2). The optimal variance terms for the regulator matrix (13) and mean and 95% credible intervals for the covariance matrix (2) are presented in Appendix S6.

Both individuals had multiple periods of fast speeds, large turning angles and high residence times (Figs 1c, d and 2). For these individuals, high residence time often indicated a corresponding large turning angle; however, these behavioural quantities were not always concurrent (Fig. 2). For example, individual BC03M04 displayed periods early in the time series where the turning angle was the strongest quantity, while speed and residence time were fairly low, suggesting a searching or nomadic behaviour (Fig. 2a). Both time series culminated with the individuals residing in two specific counties (Clear Creek and Summit), which includes an area that is considered impor-

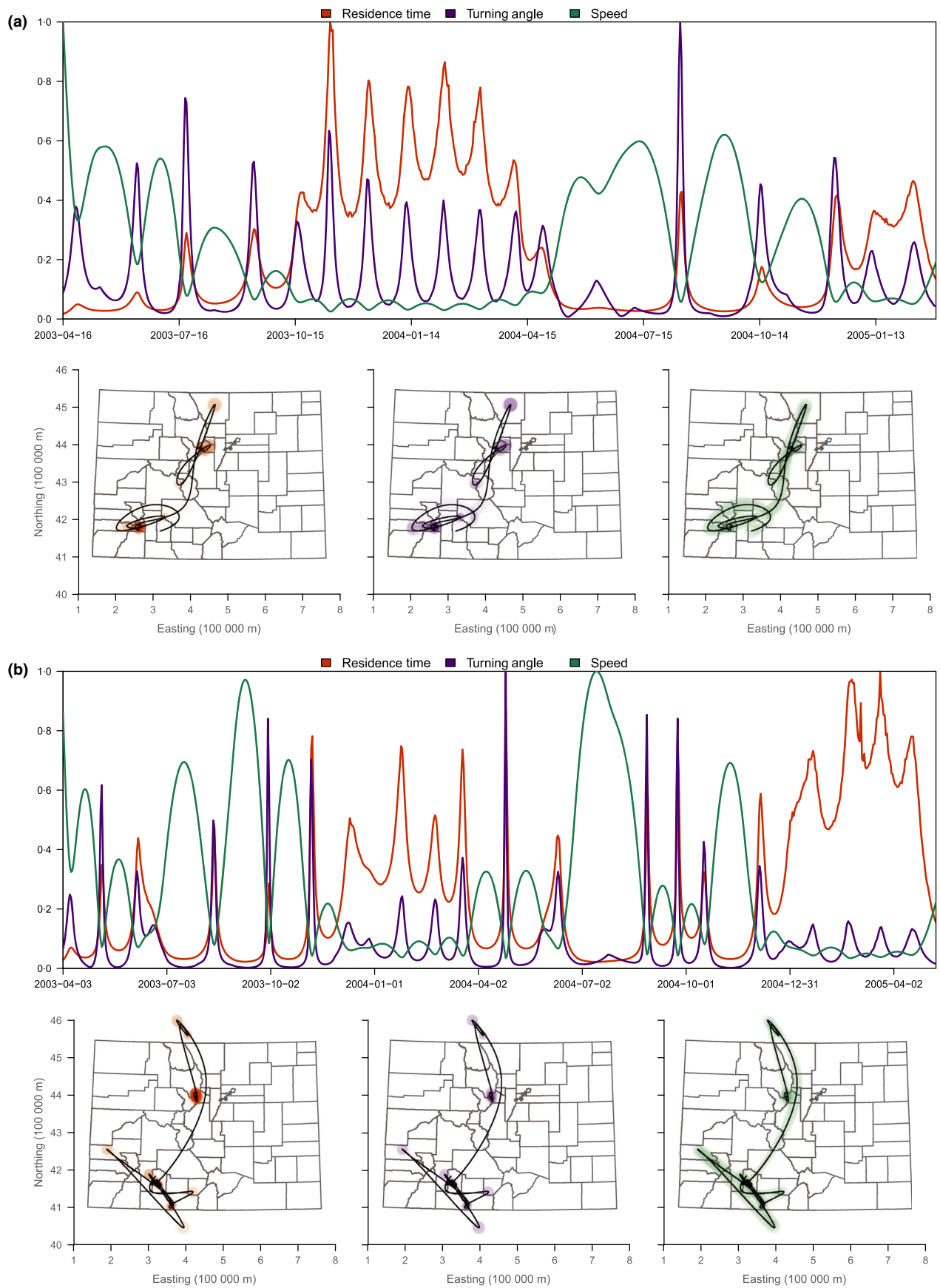
tant lynx habitat (Loveland Pass; Colorado Parks and Wildlife, pers. commun.). These results also indicate that lynx are capable of consistent long-term movement across large distances without establishing an area of high residence time. For example, within a period of two months, individual BC03F03 travelled approximately 480 km (posterior mean), from the southern portion of Colorado (Mineral County) to southern Wyoming (Medicine Bow National Forest, specifically the area located within Carbon and Albany counties; Fig. 2b).

**Discussion**

The process model we propose falls within the same class of models as statistical emulators, functional data models and process convolutions, and we showed that it can be written in much the same way. The model presented could be written as a hierarchical model, by allowing the latent process to be stochastic. However, it is well known that hierarchical models with two sources of unstructured error and lacking replication will have identifiability issues (Hobbs & Hooten 2015). Given a



**Fig. 1.** Mean and 95% credible intervals of the marginal locations for two Canada lynx [BC03M04 (a) and BC03F03 (b)], with the observed locations. The posterior mean of each movement descriptor, shown with the counties of Colorado, for individuals BC03M04 (c) and BC03F03 (d). The size of the point corresponds to spatial uncertainty, and the transparency indicates the strength of the behaviour at that location; for visualization purposes, any value below 25% of the maximum value for that behaviour is not shown. Coordinates correspond to Universal Transverse Mercator zone 13N.



**Fig. 2.** Mean relative movement descriptors through time and space for two Canada lynx reintroduced to Colorado [BC03M04's (a) and BC03F03's (b)]. Coordinates correspond to Universal Transverse Mercator zone 13N.

situation where there are strong constraints on the movement process, it may be possible to separate data and process error. For example, Brost *et al.* (In Press) were able to separately estimate data and process error in a resource selection framework by constraining the spatial domain of the process. However, their study focused on a marine mammal, and therefore, the process can be constrained to the marine environment (Brost *et al.* In Press). Constraining movement in a terrestrial environment may be possible but is less intuitive and will impose strong assumptions.

In addition, one of the benefits of using a functional data approach is its flexibility, in contrast to more constrained mechanistic models. The simplicity of the process model results in greater computational efficiency than other available methods for movement modelling. For example, our model can be fit on the order of minutes for each individual, compared to other models that require on the order of days (e.g., Hooten *et al.* 2010; McClintock *et al.* 2014a). Although small-scale movement patterns may be difficult to detect given the coarse temporal resolution and the large amount of measurement error associated with Argos locations, large-scale movement patterns are easily discernible and informative. However, researchers analysing data at a finer temporal scale could discern small-scale movement with properly scaled basis functions (e.g., daily or weekly).

The model can be used to estimate an animal's movement path alone, but is especially useful for learning about movement behaviours that describe how individuals are utilizing the landscape. For example, persistence in direction may be used to infer when and where an individual is migrating or dispersing, whereas variation in direction may indicate habitat suitability for a home range (Haddad 1999; Morales *et al.* 2004). In the data we analysed the movement descriptors corresponded with anecdotal evidence of lynx movement behaviour. Many existing methods for analysing location data explicitly model the quantities that give rise to the movement path (e.g., speed, turning angle, residence time, velocity), such that the quantities must be estimated while fitting the model (mechanistic models; e.g., Breed *et al.* 2012; Johnson *et al.* 2008; Jonsen, Flemming & Myers 2005; McClintock *et al.* 2012; Morales *et al.* 2004; Winship *et al.* 2012). In contrast, we use the equivariance property of MCMC to calculate derived quantities as well as the proper uncertainty associated with each behaviour (Hobbs & Hooten 2015). Alternative *ad hoc* methods could be used, such as calculating derived quantities based on the mean predicted path, but to ensure the validity of those quantities as estimators with proper uncertainty, a procedure like the one we describe is necessary. Quantities of interest beyond those presented can be derived, such as bearing or tortuosity, or summarized with respect to temporal and spatial features. However, our model would need to be adjusted to accommodate other sources of measurement error (e.g., GPS data).

The model that we developed may be particularly well suited for analysing data sets that have not been collected explicitly for movement analysis. These data sets may contain multiple data types, have large amounts of error and have been collected at a coarse temporal resolution. As such,

they may not be conducive for fine-scale mechanistic movement modelling. We used a data set that embodied these characteristics, the telemetry data from the Canada lynx reintroduction to Colorado, to demonstrate that the FDA approach can be used to estimate movement paths and associated movement descriptors. The biological inference from the derived movement descriptors can also be extended beyond what we show here. For example, our framework could be extended to incorporate spatial and temporal covariates into the process model, similar to the approach described by Hanks, Hooten & Alldredge (2015). In addition, the spatial distribution of the movement descriptors can be used to summarize movement behaviour across linear landscape features such as roads. Likewise, movement behaviour through nonlinear landscape features, such as National Parks, can be described with the average posterior mean of a movement descriptor within a spatial boundary. Our model can also be generalized for use with multiple individuals. In this case, the derived quantities can be aggregated to describe population-level movement. This type of population movement model allows the Argos and VHF covariance matrices to borrow strength across individuals, potentially improving parameter estimates. Such extensions are the subject of ongoing research.

## Acknowledgements

Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the US Government. Funding was provided by Colorado Parks and Wildlife (1304) and the National Park Service (P12AC11099). Data were provided by Colorado Parks and Wildlife. The authors would like to thank the anonymous reviewers for their constructive commentary that helped improve the manuscript.

## Data accessibility

Case Study Data: uploaded as online supporting information in Appendix S6: Case study tables.

## References

- Barker R.J. (1997) Joint modeling of live-recapture, tag-resight, and tag-recovery data. *Biometrics*, **53**, 666–677.
- Berger, J. (2004) The last mile: how to sustain long-distance migration in mammals. *Conservation Biology*, **18**, 320–331.
- Bjørnstad, O.N., Fromentin, J.M., Stenseth, N.C. & Gjøsæter, J. (1999) A new test for density-dependent survival: the case of coastal cod populations. *Ecology*, **80**, 1278–1288.
- Bowler, D.E. & Benton, T.G. (2005) Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, **80**, 205–225.
- Boyd, J.D. & Brightsmith, D.J. (2013) Error properties of Argos satellite telemetry locations using least squares and kalman filtering. *Plos One*, **8**, e63051
- Breed, G.A., Costa, D.P., Goebel, M.E. & Robinson, P.W. (2011) Electronic tracking tag programming is critical to data collection for behavioral time-series analysis. *Ecosphere*, **2**, art10
- Breed, G.A., Costa, D.P., Jonsen, I.D., Robinson, P.W. & Mills-Flemming, J. (2012) State-space methods for more completely capturing behavioral dynamics from animal tracks. *Ecological Modelling*, **235**, 49–58.
- Brost, B.M., Hooten, M.B., Hanks, E.M. & Small, R.J. (In Press) Animal movement constraints improve resource selection inference in the presence of telemetry error. *Ecology*. doi: 10.1890/15-0472.1.
- Buderman, F.E., Diefenbach, D.R., Casalena, M.J., Rosenberry, C.S. & Wallingford, B.D. (2014) Accounting for tagging-to-harvest mortality in a brownie



- tag-recovery model by incorporating radio-telemetry data. *Ecology and Evolution*, **4**, 1439–1450.
- Burnham, K.P. (1993) A theory for combined analysis of ring recovery and recapture data. *Marked Individuals in the Study of Bird Population* (eds J. Lebreton & P. North), pp. 199–213. Birkhauser-Verlag, Basel.
- Calder, C.A. (2007) Dynamic factor process convolution models for multivariate space-time data with application to air quality assessment. *Environmental and Ecological Statistics*, **14**, 229–247.
- Carr, J.C., Fright, W. & Beatson, R.K. (1997) Surface interpolation with radial basis functions for medical imaging. *IEEE Transactions on Medical Imaging*, **16**, 96–107.
- Costa, D.P., Robinson, P.W., Arnould, J.P., Harrison, A.L., Simmons, S.E., Hassrick, J.L., Hoskins, A.J., Kirkman, S.P., Oosthuizen, H., Villegas-Amtmann, S. & Crocker, D.E. (2010) Accuracy of ARGOS locations of pinnipeds at-sea estimated using Fastloc GPS. *PLoS One*, **5**, e8677
- Cowen, L. & Schwarz, C.J. (2005) Capture–recapture studies using radio telemetry with premature radio-tag failure. *Biometrics*, **61**, 657–664.
- Cressie, N. & Wikle, C.K. (2011) *Statistics for Spatio-Temporal Data*. John Wiley & Sons, Hoboken, New Jersey.
- De Boor, C. (1978) *A Practical Guide to Splines*. Springer, New York.
- Devineau, O., Shenk, T.M., White, G.C., Doherty Jr, P.F., Lukacs, P.M. & Kahn, R.H. (2010) Evaluating the Canada lynx reintroduction programme in Colorado: patterns in mortality. *Journal of Applied Ecology*, **47**, 524–531.
- Di Oriò, A.P., Callas, R. & Schaefer, R.J. (2003) Performance of two GPS telemetry collars under different habitat conditions. *Wildlife Society Bulletin*, **31**, 372–379.
- Dickson, B.G., Jenness, J.S. & Beier, P. (2005) Influence of vegetation, topography, and roads on cougar movement in southern California. *Journal of Wildlife Management*, **69**, 264–276.
- Douglas, D.C., Weinzierl, R., Davidson, S., Kays, R., Wikelski, M. & Bohrer, G. (2012) Moderating Argos location errors in animal tracking data. *Methods in Ecology and Evolution*, **3**, 999–1007.
- Dujon, A.M., Lindstrom, R.T. & Hays, G.C. (2014) The accuracy of Fastloc-GPS locations and implications for animal tracking. *Methods in Ecology and Evolution*, **5**, 1162–1169.
- Eilers, P.H. & Marx, B.D. (1996) Flexible smoothing with B-splines and penalties. *Statistical Science*, **11**, 89–102.
- Ferraty, F. & Vieu, P. (2006) *Nonparametric Functional Data Analysis: Theory and Practice*. Springer, New York.
- Gray, R.J. (1992) Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association*, **87**, 942–951.
- Greenwood, P.J. & Harvey, P.H. (1982) The natal and breeding dispersal of birds. *Annual Review of Ecology and Systematics*, **13**, 1–21.
- Gurarie, E., Andrews, R.D. & Laidre, K.L. (2009) A novel method for identifying behavioural changes in animal movement data. *Ecology Letters*, **12**, 395–408.
- Haddad, N.M. (1999) Corridor use predicted from behaviors at habitat boundaries. *The American Naturalist*, **153**, 215–227.
- Hanks, E.M., Hooten, M.B., Johnson, D.S. & Sterling, J.T. (2011) Velocity-based movement modeling for individual and population level inference. *PLoS One*, **6**, e22795.
- Hanks, E.M., Hooten, M.B. & Allredge, M. (2015) Continuous-time discrete-space models of animal movement. *Annals of Applied Statistics*, **9**, 145–165.
- Hastie, T.J. & Tibshirani, R.J. (1990) *Generalized Additive Models*. Chapman & Hall, London.
- Hastie, T., Tibshirani, R. & Friedman, J. (2009) *The Elements of Statistical Learning*, vol 2. Springer, New York.
- Hays, G.C. & Scott, R. (2013) Global patterns for upper ceilings on migration distance in sea turtles and comparisons with fish, birds and mammals. *Functional Ecology*, **27**, 748–756.
- Hays, G.C., Christensen, A., Fossette, S., Schofield, G., Talbot, J. & Mariani, P. (2014) Route optimisation and solving zermelo's navigation problem during long distance migration in cross flows. *Ecology Letters*, **17**, 137–143.
- Heard, D.C., Ciarniello, L.M. & Seip, D.R. (2008) Grizzly bear behavior and global positioning system collar fix rates. *The Journal of Wildlife Management*, **72**, 596–602.
- Higdon, D. (2002) Space and space-time modeling using process convolutions. *Quantitative Methods for Current Environmental Issues* (eds C. Anderson, V. Barnett, P. Chatwin & A. El-Shaarawi), pp. 37–56. Springer, New York.
- Higdon, D., Swall, J. & Kern, J. (1999) Non-stationary spatial modeling. *Bayesian Statistics*, **6**, 761–768.
- Hightower, J.E., Jackson, J.R. & Pollock, K.H. (2001) Use of telemetry methods to estimate natural and fishing mortality of striped bass in Lake Gaston, North Carolina. *Transactions of the American Fisheries society*, **130**, 557–567.
- Hobbs, N.T. & Hooten, M.B. (2015) *Bayesian Models: A Statistical Primer for Ecologists*. Princeton University Press, Princeton, New Jersey.
- Hoerner, X., Whiting, S.D., Hindell, M.A. & McMahon, C.R. (2012) Enhancing the use of Argos satellite data for home range and long distance migration studies of marine animals. *PLoS One*, **7**, e40713.
- Hooten, M.B. & Hobbs, N.T. (2015) A guide to Bayesian model selection for ecologists. *Ecological Monographs*, **85**, 3–28.
- Hooten, M.B., Johnson, D.S., Hanks, E.M. & Lowry, J.H. (2010) Agent-based inference for animal movement and selection. *Journal of Agricultural, Biological, and Environmental Statistics*, **15**, 523–538.
- Hooten, M.B., Leeds, W.B., Fiechter, J. & Wikle, C.K. (2011) Assessing first-order emulator inference for physical parameters in nonlinear mechanistic models. *Journal of Agricultural, Biological, and Environmental Statistics*, **16**, 475–494.
- Johnson, D.S., London, J.M., Lea, M.A. & Durban, J.W. (2008) Continuous-time correlated random walk model for animal telemetry data. *Ecology*, **89**, 1208–1215.
- Johnson, D.S., London, J.M. & Kuhn, C.E. (2011) Bayesian inference for animal space use and other movement metrics. *Journal of Agricultural, Biological, and Environmental Statistics*, **16**, 357–370.
- Jonsen, I.D., Flemming, J.M. & Myers, R.A. (2005) Robust state-space modeling of animal movement data. *Ecology*, **86**, 2874–2880.
- Jonsen, I.D., Myers, R.A. & James, M.C. (2007) Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Marine Ecology Progress Series*, **337**, 255–264.
- Jonzén, N., Knudsen, E., Holt, R.D. & Sæther, B.E. (2011) Uncertainty and predictability: the niches of migrants and nomads. *Animal Migration: A Synthesis* (ed E. Milner-Gulland, J.M. Fryxell & A.R.E. Sinclair), pp. 91–109. Oxford University Press, Oxford.
- Keele, L.J. (2008) *Semiparametric Regression for the Social Sciences*. John Wiley & Sons, Hoboken, New Jersey.
- Koenig, W.D., Van Vuren, D. & Hooge, P.N. (1996) Detectability, philopatry, and the distribution of dispersal distances in vertebrates. *Trends in Ecology & Evolution*, **11**, 514–517.
- Kuhn, C.E., Johnson, D.S., Ream, R.R. & Gelatt, T.S. (2009) Advances in the tracking of marine species: using gps locations to evaluate satellite track data and a continuous-time movement model. *Marine Ecology Progress Series*, **393**, 97–109.
- Langrock, R., Kneib, T., Sohn, A. & DeRuiter, S. (2014) Nonparametric inference in hidden Markov models using P-splines. arXiv:13090423v2.
- Lavielle, M. (1999) Detection of multiple changes in a sequence of dependent variables. *Stochastic Processes and their Applications*, **83**, 79–102.
- Lawler, J.J., White, D., Neilson, R.P. & Blaustein, A.R. (2006) Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biology*, **12**, 1568–1584.
- Leathwick, J., Rowe, D., Richardson, J., Elith, J. & Hastie, T. (2005) Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology*, **50**, 2034–2052.
- Lee, H.K., Holloman, C.H., Calder, C.A. & Higdon, D.M. (2002) Flexible gaussian processes via convolution. Technical Report 02-09, Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina.
- Levitin, D.J., Nuzzo, R.L., Vines, B.W. & Ramsay, J. (2007) Introduction to functional data analysis. *Canadian Psychology*, **48**, 135.
- Lidicker Jr, W. & Stenseth, N. (1992) To disperse or not to disperse: who does it and why? *Animal Dispersal: Small Mammals as a Model* (eds N.C. Stenseth & W.Z. Lidicker), pp. 21–36. Chapman & Hall, London.
- Lopez, R., Malarde, J.P., Royer, F. & Gaspar, P. (2014) Improving Argos doppler location using multiple-model kalman filtering. *IEEE Transactions on Geoscience and Remote Sensing*, **52**, 4744–4755.
- Mattisson, J., Andrén, H., Persson, J. & Segerström, P. (2010) Effects of species behavior on global positioning system collar fix rates. *The Journal of Wildlife Management*, **74**, 557–563.
- McClintock, B.T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B.J. & Morales, J.M. (2012) A general discrete-time modeling framework for animal movement using multistate random walks. *Ecological Monographs*, **82**, 335–349.
- McClintock, B.T., Johnson, D.S., Hooten, M.B., Ver Hoef, J.M. & Morales, J.M. (2014a) When to be discrete: the importance of time formulation in understanding animal movement. *Movement Ecology*, **2**, 2–21.
- McClintock, B.T., London, J.M., Cameron, M.F. & Boveng, P.L. (2014b) Modelling animal movement using the Argos satellite telemetry location error ellipse. *Methods in Ecology and Evolution*, **6**, 266–277.

- Mech, L.D. (1983) *Handbook of Animal Radio-Tracking*. University of Minnesota Press, Minneapolis.
- Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E. & Fryxell, J.M. (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, **85**, 2436–2445.
- Mueller, T., Olson, K.A., Dressler, G., Leimgruber, P., Fuller, T.K., Nicolson, C., Novaro, A.J., Bolgeri, M.J., Wattles, D., DeStefano, S., Calabrese, J.M. & Fagan, W.F. (2011) How landscape dynamics link individual-to population-level movement patterns: a multispecies comparison of ungulate relocation data. *Global Ecology and Biogeography*, **20**, 683–694.
- Nasution, M.D., Brownie, C., Pollock, K.H. & Bennetts, R.E. (2001) Estimating survival from joint analysis of resighting and radiotelemetry capture-recapture data for wild animals. *Journal of Agricultural, Biological, and Environmental Statistics*, **6**, 461–478.
- O'Hagan, A. & Kingman, J. (1978) Curve fitting and optimal design for prediction. *Journal of the Royal Statistical Society Series B (Methodological)*, **40**, 1–42.
- Patterson, T.A., McConnell, B.J., Fedak, M.A., Bravington, M.V. & Hindell, M.A. (2010) Using GPS data to evaluate the accuracy of state-space methods for correction of Argos satellite telemetry error. *Ecology*, **91**, 273–285.
- Pozdnyakov, V., Meyer, T.H., Wang, Y.B. & Yan, J. (2014) On modeling animal movements using Brownian motion with measurement error. *Ecology*, **95**, 247–253.
- R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramsay, J.O. & Dalzell, C. (1991) Some tools for functional data analysis. *Journal of the Royal Statistical Society Series B (Methodological)*, **53**, 539–572.
- Ramsay, J.O. & Silverman, B.W. (2002) *Applied Functional Data Analysis: Methods and Case Studies*. Springer-Verlag, New York.
- Ramsay, J.O. & Silverman, B.W. (2005) *Functional Data Analysis*, 2nd edn. Springer, New York.
- Rice, J.R. (1969) *The Approximation of Functions*, vol 2. Addison-Wesley, Reading.
- Robinson, A., Crick, H.Q., Learmonth, J.A., Maclean, I.M., Thomas, C.D., Bairlein, F., et al. (2009) Travelling through a warming world: climate change and migratory species. *Endangered Species Research*, **7**, 87–99.
- Runge, C.A., Martin, T.G., Possingham, H.P., Willis, S.G. & Fuller, R.A. (2014) Conserving mobile species. *Frontiers in Ecology and the Environment*, **12**, 395–402.
- Ruppert, D., Wand, M.P. & Carroll, R.J. (2003) *Semiparametric Regression*. 12. Cambridge University Press, New York.
- Sáenz-Romero, C., Rehfeldt, G.E., Crookston, N.L., Duval, P., St-Amant, R., Beaulieu, J. & Richardson, B.A. (2010) Spline models of contemporary, 2030, 2060, and 2090 climates for Mexico and their use in understanding climate-change impacts on the vegetation. *Climatic Change*, **102**, 595–623.
- Sapirstein, J. & Johnson, W. (1996) The use of basis splines in theoretical atomic physics. *Journal of Physics B: Atomic, Molecular and Optical Physics*, **29**, 5213.
- Sawyer, H., Lindzey, F. & McWhirter, D. (2005) Mule deer and pronghorn migration in western Wyoming. *Wildlife Society Bulletin*, **33**, 1266–1273.
- Schaub, M. & Abadi, F. (2011) Integrated population models: a novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, **152**, 227–237.
- Schofield, G., Dimadi, A., Fossette, S., Katselidis, K.A., Koutsoubas, D., Lilley, M.K., Luckman, A., Pantis, J.D., Karagouni, A.D. & Hays, G.C. (2013) Satellite tracking large numbers of individuals to infer population level dispersal and core areas for the protection of an endangered species. *Diversity and Distributions*, **19**, 834–844.
- Scott, R., Marsh, R. & Hays, G.C. (2014) Ontogeny of long distance migration. *Ecology*, **95**, 2840–2850.
- Shigesada, N. & Kawasaki, K. (2002) Invasion and the range expansion of species: effects of long-distance dispersal. *Dispersal Ecology* (eds J. Bullock, R. Kenward & S. Hails), pp. 350–373. Blackwell Science, Malden, Massachusetts.
- Singh, N.J., Börger, L., Dettki, H., Bunnefeld, N. & Ericsson, G. (2012) From migration to nomadism: movement variability in a northern ungulate across its latitudinal range. *Ecological Applications*, **22**, 2007–2020.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & Van Der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **64**, 583–639.
- Sur, M., Skidmore, A.K., Exo, K.M., Wang, T., J Ens, B. & Toxopeus, A. (2014) Change detection in animal movement using discrete wavelet analysis. *Ecological Informatics*, **20**, 47–57.
- Trakhtenbrot, A., Nathan, R., Perry, G. & Richardson, D.M. (2005) The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions*, **11**, 173–181.
- Tremblay, Y., Shaffer, S.A., Fowler, S.L., Kuhn, C.E., McDonald, B.I., Weise, M.J., Bost, C.A., Weimerskirch, H., Crocker, D.E., Goebel, M.E. & Costa, D.P. (2006) Interpolation of animal tracking data in a fluid environment. *Journal of Experimental Biology*, **209**, 128–140.
- Vincent, C., McConnell, B.J., Ridoux, V. & Fedak, M.A. (2002) Assessment of Argos location accuracy from satellite tags deployed on captive gray seals. *Marine Mammal Science*, **18**, 156–166.
- Wahba, G. (1978) Improper priors, spline smoothing and the problem of guarding against model errors in regression. *Journal of the Royal Statistical Society Series B (Methodological)*, **40**, 364–372.
- Wecker, W.E. & Ansley, C.F. (1983) The signal extraction approach to nonlinear regression and spline smoothing. *Journal of the American Statistical Association*, **78**, 81–89.
- Weisberg, S. (2014) *Applied Linear Regression*, 4th edn. John Wiley & Sons, Hoboken, New Jersey.
- White, G.C. & Shenk, T.M. (2001) Population estimation with radio-marked animals. *Radio Tracking and Animal Populations* (eds J.J. Millsbaugh & J.M. Marzluff), pp. 329–350. Academic Press, San Diego, California.
- Winship, A.J., Jorgensen, S.J., Shaffer, S.A., Jonsen, I.D., Robinson, P.W., Costa, D.P. & Block, B.A. (2012) State-space framework for estimating measurement error from double-tagging telemetry experiments. *Methods in Ecology and Evolution*, **3**, 291–302.
- Winterstein, S.R., Pollock, K.H. & Bunck, C.M. (2001) Analysis of survival date from radiotelemetry studies. *Radio Tracking and Animal Populations* (eds J.J. Millsbaugh & J.M. Marzluff), pp. 351–380. Academic Press, San Diego, California.
- Wold, S. (1974) Spline functions in data analysis. *Technometrics*, **16**, 1–11.
- Wood, S.N. & Augustin, N.H. (2002) GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological Modelling*, **157**, 157–177.
- Wood, B.C. & Pullin, A.S. (2002) Persistence of species in a fragmented urban landscape: the importance of dispersal ability and habitat availability for grassland butterflies. *Biodiversity & Conservation*, **11**, 1451–1468.
- Yasuda, T. & Arai, N. (2005) Fine-scale tracking of marine turtles using GPS-Argos ppts. *Zoological Science*, **22**, 547–553.

Received 23 January 2015; accepted 10 August 2015

Handling Editor: Jason Matthiopoulos

## Supporting Information

Additional Supporting Information may be found in the online version of this article.

**Appendix S1.** Prior specifications.

**Appendix S2.** Deviance information criterion calculation.

**Appendix S3.** MCMC algorithm for fitting the spline-based movement model and calculating derived behavioral.

**Appendix S4.** Spatial quantities.

**Appendix S5.** Temporal quantities.

**Appendix S6.** Case study tables.

**Appendix S7.** Centered and scaled location data for two Canada lynx, BC03M04 and BC03F03, released in Colorado. Data type and Argos error classification is listed for each location.

**Appendix S8.** Simulations.