Model 1G

pp. 1-24 (col. fig: NIL)

STATISTICS

Spatial Statistics xxx (xxxx) xxx

ELSEVIER

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

A flexible movement model for partially migrating species

Elizabeth Eisenhauer^{a,*}, Ephraim Hanks^a, Matthew Beckman^a, Robert Murphy^b, Tricia Miller^c, Todd Katzner^d

^a Department of Statistics, The Pennsylvania State University, University Park, PA, USA

^b Eagle Environmental, Inc., Santa Fe, NM, USA

^c Conservation Science Global, Inc., West Cape May, NJ, USA

^d U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Boise, ID, USA

ARTICLE INFO

Article history: Received 1 November 2021 Received in revised form 2 February 2022 Accepted 3 February 2022 Available online xxxx

Keywords: Animal movement Stochastic differential equations Animal tracking Ecology Spatial statistics

ABSTRACT

We propose a flexible model for a partially migrating species, which we demonstrate using yearly paths for golden eagles (Aquila chrysaetos). Our model relies on a smoothly time-varying potential surface defined by a number of attractors. We compare our proposed approach using varying coefficients to a latentstate model, which we define differently for migrating, dispersing, and local individuals. While latent-state models are more common in the existing animal movement literature, varying coefficient models have various benefits including the ability to fit a wide range of movement strategies without the need for major model adjustments. We compare simulations from the models for three individuals to illustrate the ability of our model to better describe movement behavior for specific movement strategies. We also demonstrate the flexibility of our model by fitting several individuals whose movement behavior is less stereotypical.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Movement behavior within species is often highly variable across individuals and years. While some animal populations follow similar migratory trajectories or travel in groups, many display

⁶ Corresponding author. *E-mail address:* eisenhauer@psu.edu (E. Eisenhauer).

https://doi.org/10.1016/j.spasta.2022.100637 2211-6753/© 2022 Elsevier B.V. All rights reserved.

Spatial Statistics xxx (xxxx) xxx

partial migration, meaning seasonal migration is observed only in a fraction of individuals in the population (Chapman et al., 2011). Non-migratory strategies include residential (i.e., sedentary), nomadic, and dispersal behaviors (Mueller and Fagan, 2008). Current inference frameworks for partially migrating species require researchers to first define explicit movement strategies exhibited by the species (e.g., Fullman et al., 2021; Poessel et al., 2016). Researchers then classify individual paths as one of the defined strategies using methods based on spatially explicit measures or model selection (Cagnacci et al., 2016). Classification is often followed by descriptive statistics for each movement strategy or interpretation of statistical models formulated for each movement strategy (Fullman et al., 2021).

Golden eagles (*Aquila chrysaetos*) display partial migration (Poessel et al., 2016). Understanding the yearly movement strategies of golden eagles is important for conservation and management of the species. Golden eagles' high mobility, for example, carries individuals across political boundaries, forcing management efforts for the same individuals to be shared by multiple governing bodies (Brown et al., 2017).

Morales et al. (2010) argue for the importance of understanding the links between movement and population dynamics. Population-level inference using a hierarchical structure depends on individual-level models (e.g., Hooten et al., 2016), so it is essential to develop individual-level models that describe behavior well. The task of developing realistic individual-level models becomes more difficult the more heterogeneous the population.

In Fig. 1, we display year-long paths for three individuals that used three movement strategies, which could be described as residence, migration, and dispersal. We define residence as attraction to a single location throughout the year. We define migration as a path where the individual spends a portion or all of the summer season in a single location and a portion or all of winter in a more southern location. We define dispersal as a path where the individual is attracted to one location for a period starting in the beginning of the year and switches to a new location for the remainder of the year.

Partitioning groups of golden eagles based on movement strategy can be a challenging task due to the presence of 'less-stereotyped' or 'mixed' cases (Cagnacci et al., 2016). Some authors have suggested movement strategies in partially migrating populations would be better described as existing on a continuum, which would better accommodate those less-stereotyped cases (Ball et al., 2001; Cagnacci et al., 2016). Thus there is a need for flexible models that are capable of fitting multiple movement strategies, without predefining those strategies.

Varying coefficient models that allow behavior to transition smoothly in time have recently received attention in the animal movement literature for being a more flexible and realistic alternative to the latent-state model (Michelot et al., 2020; Russell et al., 2018, 2017). In this work, we describe a single varying coefficient model which utilizes a stochastic differential equation (SDE) framework similar to that of Eisenhauer and Hanks (2020). We fit the varying coefficient model for a variety of movement paths including those displaying residential, migratory, and dispersal behavior. The advantage of this varying coefficient framework is that the same model can easily be used to provide insight into movement behavior that fits one of these three categories, as well as behavior that does not clearly fit into only one of these categories. Our proposed model can produce realistic simulated paths for a range of movement strategies.

We compare our approach to a latent-state model within the same SDE framework, and we show that our varying coefficient model better describes movement behavior. Latent-state models are commonly used in animal movement modeling (Pirotta et al., 2018; Patterson et al., 2017), and there exist popular R packages that can be used to easily fit these types of models for animal telemetry data (Michelot et al., 2016; McClintock and Michelot, 2018). The latent-state models are not as flexible as our varying coefficient model and need to be specified differently depending on the movement strategy. We defined different sets of states for residential, dispersal, and migratory movement strategies.

In Section 2, we describe the golden eagle data and motivate the selection of the subset we focus on in this paper. In Section 3, we describe the SDE model framework which is common to all models we consider in this paper. In Section 4, we present our varying coefficient model. In Sections 5–7, we fit the varying coefficient and alternative models to three paths we selected to illustrate three



Spatial Statistics xxx (xxxx) xxx



Fig. 1. Year-long movement paths for 3 golden eagles in the western United States, where path color differentiates between individuals. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

stereotypic movement strategies: residence, dispersal, and migration. We then illustrate how the varying coefficient model can be used to fit a wider range of movement behaviors in Section 8. Lastly, we summarize the results and suggest areas for future work in Section 9.

2. Golden eagle telemetry data

We obtained satellite telemetry data for 68 individuals, each of which was tracked for at least 1 year in the western United States. Tagging of eagles and collection of data was funded by the National Raptor Program of the U.S. Fish and Wildlife Service (USFWS), and we accessed the data through collaboration with the USFWS and one of its contractors, Eagle Environmental, Inc. Movement paths for all individuals, based on hourly GPS locations accurate to within 19 m, are shown in Fig. 2. Most of the eagles in this dataset were fledgelings when tagged. The eagles utilized a wide range of habitats in the western United States from desert, semi-arid plains, shrub-steppe, and mountains to arctic tundra. We subdivided individual movement paths by year (Jan 1–Dec 31) and removed paths with a span of observations shorter than 290 days. We were left with 194 yearly paths for a total of 67 unique individuals. An exploratory analysis of each yearly path using the migrateR R package (Spitz et al., 2017) identified 18 dispersers or nomads, 161 migrants or mixed migrants, and 15 residents using net squared displacement. The meaning of these terms is shown graphically in Fig. 3 (Spitz et al., 2017). Model selection using AIC favors the more complex model, and inspection revealed that many of the paths classified as migration or mixed migration appeared closer to the dispersal, residential, or nomadic strategies. Our varying coefficient modeling approach removes the need to classify each path into only one category (e.g., migrant, resident, disperser) and allows for a more nuanced and realistic representation of bird behavior.



Spatial Statistics xxx (xxxx) xxx



Fig. 2. Paths for all eagles in the dataset. Color indicates individual bird. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

These data display a wide range of movement behaviors, most of which do not clearly belong to a single movement strategy. This may be due the young age of many of the birds and much of the data collection taking place in a desert ecosystem. Thus, we selected datasets from three individuals which were visually identified as clearly displaying one of each of the three movement strategies: migration, residence, and dispersal. We focused this analysis on comparing models fit for these three records.

The latent-state models require regular time intervals, but the data are irregular in time. To resolve this issue, we thinned the data to only one observation per day and linearly interpolated missing intervals that were all shorter than 30 days. We chose to use daily observations because we are interested in movement behaviors that happen over the course of five days at the least. See Appendix A for details.

3. A stochastic differential equation model framework for animal movement

We considered a flexible SDE model framework following Russell et al. (2018) and Hanks et al. (2017). We adopt the notation of Eisenhauer and Hanks (2020). The continuous time model for an animal's position \mathbf{r}_t at time t is

$$d\mathbf{r}_t = \mathbf{v}_t dt$$



E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 3. Plot replicated with permission from Spitz et al. (2017).

$$d\mathbf{v}_t = -\beta(\mathbf{v}_t - \boldsymbol{\mu}(\mathbf{r}_t))dt + \sigma \mathbf{I} d\mathbf{w}_t$$
⁽²⁾

where \mathbf{v}_t is the velocity of the animal at time t, β is the coefficient of friction (Nelson, 1967) which controls autocorrelation in movement, $\mu(\mathbf{r}_t)$ is the mean drift in the direction of movement, σ is a scalar that controls the variance in the stochastic term, \mathbf{I} is a 2 × 2 identity matrix, and \mathbf{w}_t is independent Brownian motion in \mathbb{R}^2 .

We adopt the additional simplification of a constant motility surface and regular time intervals of 1 day. Thus an Euler–Maruyama scheme (Kloeden and Platen, 1992) approximates (1)–(2) by

$$\mathbf{r}_{t+1} = \mathbf{r}_t + \mathbf{v}_t \tag{3}$$

$$\mathbf{v}_{t+1} = \mathbf{v}_t - \beta(\mathbf{v}_t - \boldsymbol{\mu}(\mathbf{r}_t)) + \sigma \mathbf{I} d\mathbf{w}_t$$
(4)

and substituting (3) into (4) results in

$$\mathbf{r}_{t+2} - 2\mathbf{r}_{t+1} + \mathbf{r}_t = \beta \left(-\nabla p(\mathbf{r}_t) - \mathbf{r}_{t+1} + \mathbf{r}_t \right) + \sigma \epsilon_t$$
(5)

where we have modeled the mean drift $\mu(\mathbf{r}_t)$ as the negative gradient of a potential function $p(\mathbf{r}_t)$ (Brillinger et al., 2012; Preisler et al., 2013; Eisenhauer and Hanks, 2020; Russell et al., 2018). We define this potential function $p(\mathbf{r}_t)$ as a weighted sum of distances to *m* fixed attractors with *x*-coordinates $a_{x1}, a_{x2}, \ldots, a_{xm}$ and *y*-coordinates $a_{y1}, a_{y2}, \ldots, a_{ym}$, i.e.,

$$p(\mathbf{r}_t) = \sum_{i=1}^m k_{it} \sqrt{(x_t - a_{xi})^2 + (y_t - a_{yi})^2}$$
(6)

where x_t and y_t are the *x*- and *y*-coordinates of \mathbf{r}_t and k_{it} is the coefficient of attraction to the *i*th attractor. The models considered in this paper all follow this framework but have varying number of attractors *m* and coefficients of attraction k_{it} that may or may not change over time *t*.

Spatial Statistics xxx (xxxx) xxx

4. Flexible model for partially migratory species

The varying coefficient model we considered fixes the number of attractors m = 8 and allows k_{it} for i = 1, 2, ..., m to change smoothly over time. We chose to use m = 8 as an overestimate of the number of attractors, and we used shrinkage methods to effectively select a subset of the attractors (Marra and Wood, 2011). It is clear that at least two attractors are needed for a migrant or disperser model, and the additional attractors might capture stopover sites or other irregular behavior. We chose the eight attractors with k means clustering of the daily locations with eight clusters, but they could be chosen with any method of the researcher's choice or considered latent variables to be estimated jointly with other parameters. Each coefficient of attraction k_{it} is a weighted sum of cyclic cubic regression spline basis functions. Thus, the potential function becomes

SPASTA: 100637

$$p(\mathbf{r}_t) = \sum_{i=1}^m \sum_{j=1}^J \alpha_{ij} B_j(t) \sqrt{(x_t - a_{xi})^2 + (y_t - a_{yi})^2}$$
(7)

where α_{ij} is the coefficient of each cyclic cubic basis function $B_j(t)$ for attractor *i*, and the number of basis functions *J* is bounded above by 30. We fit the model expressed by plugging the gradient of (7) into (5) using the gam function in the **mgcv** R package. We also penalized the null space of the basis functions to implement selection of the attractor coefficients via regularization (Marra and Wood, 2011).

5. Resident example

We define a resident as an individual attracted to the same location throughout the year. In this section, we compare our varying coefficient model from Section 4 to a single-state model formulated specifically for the residential movement strategy. For this comparison, we chose a path consisting of a single year of data for one individual. We visually determined that this path, shown in Fig. 4A, displayed a residential movement strategy.

5.1. Single-state residence model for comparison

We formulate a model in the SDE framework which is specific to residential behavior. Since we assume only one type of attraction, e.g., a nest, occurs throughout the year for a resident individual, m = 1 and we estimated the single attractor as the median of all data points. The potential function becomes

$$p(\mathbf{r}_t) = k \sqrt{(x_t - a_{x1})^2 + (y_t - a_{y1})^2}$$
(8)

where a_{x1} and a_{y1} are the *x*- and *y*-coordinates of the single attractor and *k* is the single coefficient of attraction. We restricted *k* to be positive so that the attractor is forced to have a positive attraction as opposed to a negative attraction or repulsion. The unknown parameters of this simple model are β , *k*, and σ .

We fit this model in a Bayesian framework with the no-U-turn sampler (NUTS) implemented in Stan for consistency with the latent-state models for the migrant and dispersal, but it was not necessary to fit this model in a Bayesian framework. We also fit the same model using the **Im** function in the **stats** R package and observed similar results. Two Markov chains of length 10,000 were assessed visually to assess convergence. The first 5000 samples were discarded as burn-in, and the last 5000 iterations from the first chain were used for inference.

5.2. Resident results

The attractors for the single-state residence model from Section 5.1 and our varying coefficient model from Section 4 are labeled on top of daily locations in blue in Fig. 4D and Fig. 4E, respectively. For the single-state residence model, the posterior mean of the coefficient of attraction, *k*, is 6773.0.



Spatial Statistics xxx (xxxx) xxx



SPASTA: 100637

Fig. 4. (A) Movement path for the resident in 2015. (B) Potential surface for the varying coefficient model. Attractors are numbered. (C) Potential surface for the single-state residence model. The label "1" is located at the single attractor. Bounds are the same as in B. The original path is plotted in blue with the simulation in red from (D) the single-state model and (E) the varying coefficient model. (F) Density plots of average distance from the simulations to the true path for varying coefficient and single-state models. Vertical lines indicate the means for each model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We used this estimate to construct the potential surface for the single-state model (8) in Fig. 4B. More information about the posterior samples can be found in Appendix B.

For our varying coefficient model described in Section 4, model diagnostics tests – e.g., **mgcv:: gam.check()** in R – did not reveal evidence of a substantive departure from the model assumptions. This includes verification that the imposed maximum of 30 basis functions had been sufficient (i.e., not restrictively low). We estimated $\beta k_{it} = \beta \sum_{j=1}^{J} \alpha_{ij} B_j(t)$ for $i \in \{1, 2, ..., 8\}$ in the potential function (7). The smooth estimates of βk_{it} over time *t* are shown in Appendix E with standard errors. Values of βk_{it} above 0 indicate attractor *i* has a positive attraction at time *t*, and values of βk_{it} below 0 indicate attractor *i* has a negative attraction or a repulsion at time *t*. We used these estimates to construct a potential surface which changes smoothly over time. A series of snapshots of this potential surface throughout the year is shown in Fig. 4C. The degree of attraction to different attractors 2, 4, and 5.

One way to assess how well a model fits a movement path is to simulate from the model and measure how close the simulation is to the original path. To compare the single-state and varying coefficient model, we simulated from both models fit with the same path 100 times. We simulated 365 days of data, the same length as the original path. One representative simulation from each



model is shown in Fig. 4D–E. For each model, we calculated the distance between each simulated location and the original path on the same day of the year. Mean distances for each simulation from both models are shown in Fig. 4F. On average, the varying coefficient model simulations more closely resemble the original path, indicating that our varying coefficient model better describes the movement behavior (Fig. 4F).

All computing was performed on a laptop computer (2.9 GHz Dual-Core Intel Core i5). The computing time for the varying coefficient model is less than 1 s while the computing time for the single-state model is 27 s to run 2 chains for 10,000 iterations.

6. Dispersal example

We define dispersal as an individual that switches from being attracted to one location to being attracted to a second location at some point in the year and remains attracted to the second location for the rest of the year. The path we analyzed as a path displaying dispersal was collected in the year 2018. We compare our varying coefficient model described in Section 4 to a latent-state model formulated specifically for dispersal.

6.1. Latent-state dispersal model for comparison

We formulated a model in the SDE framework which is specific to dispersal. In this model, we estimated two attractors using k means. To capture the shift from attractor 1 to attractor 2, we used a framework similar to a Hidden Markov Model (HMM) framework consisting of 2 states. We chose to use the term latent-state model instead of HMM following Zucchini et al. (2008) due to the dependence between observations after accounting for the latent states. Thus, the Markov assumption is violated. However, the distinction is relatively unimportant since the methods are functionally almost identical to those used for a HMM, including the forward algorithm and Viterbi algorithm (Zucchini et al., 2017).

In state 1, the individual is attracted to attractor 1, and in state 2, the individual is attracted to attractor 2. Thus we estimated β and σ as in Section 5.1 but now we also estimated 2 different values for k, called k_1 and k_2 , corresponding to the strength of attraction to attractor 1 while in state 1 and the strength of attraction to attractor 2 while in state 2, respectively. We also estimated the probability of transitioning from state 1 to state 2 on any particular day. We characterized this probability of transitioning with a vector of length 2 called **g** where the first element is the probability of staying in state 1 and the second is the probability of transitioning. Once in state 2, the probability of transitioning back to state 1 was set equal to 0 to ensure only one transition.

As for the single-state resident model, we fit this model in a Bayesian framework with NUTS implemented in Stan. Two Markov chains of length 10,000 were assessed visually to assess convergence. The first 5000 samples were discarded as burn-in, and the last 5000 iterations from the first chain were used for inference.

6.2. Dispersal results

Attractors chosen with k means for each of the latent-state and varying coefficient model are shown in Fig. 5D–E along with the true path in blue. The posterior mean estimates from the latent-state dispersal model for k_1 and k_2 are 9228.7 and 7067.3, respectively. We used these values to construct the potential surfaces for the two states in the latent-state model as shown in Fig. 5B. More information about the posterior samples can be found in Appendix C.

We estimated the posterior probability of being in each state on each day using the Viterbi algorithm (see Zucchini et al. (2017) for a description of the algorithm). In Appendix C the most probable state sequence is depicted using different colors for each state. The state switches when the individual begins dispersal.

For our varying coefficient model described in Section 4, model assumptions were satisfied including the maximum of 30 basis functions not being restrictively low. We estimated $\beta k_{it} = \beta \sum_{i=1}^{J} \alpha_{ii} B_i(t)$ for $i \in \{1, 2, ..., 8\}$ in the potential function (7). The smooth estimates of βk_{it} over

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 5. (A) Movement path for the individual showing dispersal in 2018. (B) Potential surfaces for the two states in the latent-state model for dispersal. The attractors are identified with the numbers 1 and 2. (C) Potential surface for the varying coefficient model with same bounds as B. Attractors are numbered. The original path is plotted in blue with the simulation in red from (D) the latent-state model and (E) the varying coefficient model. (F) Density plots of average distance from the simulations to the true path for varying coefficient and latent-state models. Vertical lines indicate the means for each model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



time t are shown in Appendix E with standard errors. The smooth estimates were used to construct a potential surface which changes smoothly over time (see Fig. 5C). Around the end of April, a shift can be seen which corresponds to the time of the dispersal from one location to another.

We compared 100 365-day-long simulations from the varying coefficient and the latent-state dispersal model. One representative simulation from each model is shown in Fig. 5D–E. For both models, we calculated the distance between each simulated location and the original path on the same day of the year. Mean distances for each simulation from both models are shown in Fig. 5F. As was the case for the resident strategy, the varying coefficient model simulations more closely resemble the original path on average (Fig. 5F).

All computing was performed on a laptop computer (2.9 GHz Dual-Core Intel Core i5). The computing time for the varying coefficient model is less than 1 s while the computing time for the latent-state dispersal model is 110 s to run 2 chains for 10,000 iterations each. Convergence of the chains was assessed visually (see Appendix C).

7. Migrant example

We defined a migrant as an individual who switches seasonally from being attracted to a southern location to a northern location and back to the original southern location throughout the year. The path we analyzed was collected from an eagle in the year 2012. We compare our varying coefficient model described in Section 4 to a latent-state model formulated specifically for migration.

7.1. Latent-state migrant model for comparison

We formulated a model in the SDE framework which is specific to migration. To fit this model, we began by estimating two attractors using k means. We used a latent-state model framework, this time with five states which are analogous to northern migration, northern residence, southern migration, and two states for southern residence. States 1 (northern migration) and 2 (northern residence) both have an attraction to attractor 2, which is the northern-most attractor, but the coefficient of attraction k in (8) for state 1, which we call k_1 , was restricted to be larger than the k for state 2, which we call k_2 . This means the attraction is stronger in state 1 than in state 2. We defined states 3 (southern migration) and 4 (southern residence) similarly with the same southern attractor and restricting $k_3 > k_4$. State 5 is identical to state 4 in all ways except for the transition probabilities. Thus, the coefficient of attraction for state 5 is $k_5 = k_4$.

Since the year starts in January, we assumed the path to begin in a state of attraction to the southern residence, which we call state 4. We then allowed a positive transition out of state 4 so the individual may transition into northern migration. We also assumed the individual ends the year, in December, attracted to the southern residence again. We restricted the probabilities to only allow one cycle through the states according to a prescribed sequence. This restriction prevents the use of these states to characterize brief excursions by quickly cycling through the states repeatedly. In order to specify this rule, we created the fifth state, state 5, which is identical to state 4 but which cannot be transitioned out of. Thus the required order of transitions is 4 to 1 to 2 to 3 to 5 (i.e., southern residence starting on Jan 1, migration, northern residence, migration, southern residence ending on Dec 31).

Instead of directly estimating the transition probabilities, we allowed the transition probabilities to change over time by letting them be functions of a covariate, daily change in daylight length. To illustrate why we chose to use this covariate, we have plotted the latitude of the daily observations and the covariate, which both change in time, in Appendix D. It appears that the covariate should be high during transitions from states 4 to 1 and from states 1 to 2. Similarly, the covariate is low during transitions from states 2 to 3 and from states 3 to 5. Thus, we intended to model this relationship.

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 6. (A) Movement path for the migrant in 2012. (B) Potential surfaces for the five states in the latent-state model for the migratory strategy. The attractors are identified with the numbers. (C) Potential surface for the varying coefficient model with same bounds as B. The original path is plotted in blue with the simulation in red from (D) the latent-state model and (E) the varying coefficient model. (F) Density plots of average distance from the simulations to the true path for varying coefficient and latent-state models. Vertical lines indicate the means for each model. (F) for interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

7.2. Migrant results

We compared the latent-state migrant model to our varying coefficient model described in Section 4. Attractors chosen with k means for each of models are shown in Fig. 6D–E on top of true paths in blue. For the latent-state migrant model, the posterior means of k_1, k_2, \ldots, k_5 are as

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 7. In (A), (C), and (E), the original paths are plotted in blue with the simulations from the varying coefficient model plotted in red. The attractors are labeled. In (B), (D), and (F), the varying coefficients corresponding to the attractors in (A), (C), and (E), respectively, are shown over time. For example, k1 in (B) is the varying coefficient corresponding to attractor 1 in (A). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

follows: $\hat{k}_1 = 33826.8$, $\hat{k}_2 = 10022.8$, $\hat{k}_3 = 39865.8$, and $\hat{k}_4 = \hat{k}_5 = 12950.6$. We used these values to construct the potential surfaces for the five states in the latent-state model as shown in Fig. 6B. More information about the posterior samples can be found in Appendix D.



Spatial Statistics xxx (xxxx) xxx



SPASTA: 100637

Fig. 8. Potential surfaces for the varying coefficient model for the resident with (A) daily observations and (B) observations every 2 days. Black dots indicate attractors.

We estimated the posterior probability of being in each state on each day using the Viterbi algorithm. In Appendix D the most likely state sequence is depicted using different colors for each state. The latent-state migrant model is reasonably subdividing the path (see Fig. 19 in Appendix D).

For our varying coefficient model described in Section 4, model assumptions were satisfied including the maximum of 30 basis functions not being restrictively low. We estimated $\beta k_{it} = \beta \sum_{j=1}^{J} \alpha_{ij} B_j(t)$ for $i \in \{1, 2, 3, 4\}$ in the potential function (7). The smooth estimates of βk_{it} over time *t* are shown in Appendix E with standard errors. The smooth estimates were used to construct a potential surface which changes smoothly over time (see Fig. 6C). The snapshots of the potential surface throughout the year indicate that southern migration occurred around April and May, and northern Migration occurred in August and September.

We compared 100 337-day-long simulations from the varying coefficient and the latent-state migrant model. One representative simulation from each model is shown in Fig. 6D–E. For both

SPASTA: 100637

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 9. Potential surfaces for the varying coefficient model for the migratory path with (A) daily observations and (B) observations every 2 days. Black dots indicate attractors.

SPASTA: 100637

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 10. Potential surfaces for the varying coefficient model for the path displaying dispersal with (A) daily observations and (B) observations every 2 days. Black dots indicate attractors.

models, we calculated the distance between each simulated location and the original path on the same day of the year. Mean distances for simulations from both models are shown in Fig. 6F. As was the case for the resident and dispersal strategies, the varying coefficient model simulations more closely resemble the original path on average (Fig. 6F).

SPASTA: 100637

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx





C.Cahone_2015_stan_output



Fig. 12. Traceplots of MCMC samples from the marginal posterior distributions.



Fig. 13. Histograms of MCMC samples from the marginal posterior distributions for the latent-state model for dispersal.



Fig. 14. Traceplots of MCMC samples from the marginal posterior distributions for the latent-state model for dispersal.



Fig. 15. States for the dispersal model. In the top plot, the line width is proportional to the posterior probabilities of being in each state across time. In the bottom plot, the *y* axis is the *y*-coordinate of location (in meters). The observations are colored by most likely state. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

All computing was performed on a laptop computer (2.9 GHz Dual-Core Intel Core i5). The computing time for the varying coefficient model was less than 1 s, while the computing time for the latent-state migrant model was >10 h to run 2 chains for 5000 iterations each. Convergence of the chains was assessed visually (see Appendix D).

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 16. Histograms of MCMC samples from the marginal posterior distributions for the latent-state model for the migratory strategy.



Fig. 17. Traceplots of MCMC samples from the marginal posterior distributions for the latent-state model for the migratory strategy.

8. Fitting boundary individuals

An important benefit of the flexible movement model described in Section 4 is its ability to fit a wide range of movement behaviors, including those that do not clearly fit the migrant, resident, or dispersal stereotypes. There are many such individuals in the golden eagle dataset since they are a partially migrating species. Three examples of less-stereotyped paths are shown in Fig. 7 along with simulations from the varying coefficient model fit to each path and the varying coefficients. Simulations from the varying coefficient model are reasonable even for more irregular movement behavior such as these examples.

9. Discussion and future work

We have described a flexible model using varying coefficients for fitting individual movement paths for a partially migrating species. We compared our varying coefficient model described in Section 4 to latent-state models within the same SDE model framework for three individual golden eagles. For these three individuals displaying migration, residence, and dispersal, simulations from our varying coefficient model more closely resembled the true paths. We also illustrated the ability of our varying coefficient model to fit boundary individuals which do not clearly exhibit migration, residence, or dispersal. The latent-state model is also restricted by needing discrete time steps. In our examples, we fit the varying coefficient model to paths with observations at discrete daily time



Spatial Statistics xxx (xxxx) xxx



Fig. 18. On the top, we see the latitude of the individual changing over time, and on the bottom, we see the covariate, daily change in day length, changing over time. All data is for NM Tredwell's location in 2012.



Fig. 19. States for the migrant model. In the top plot, the line width is proportional to the posterior probabilities of being in each state across time. In the bottom plot, the y axis is the y-coordinate of location. The observations are colored by most likely state. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

SPASTA: 100637 ARTICLE IN PRESS

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 20. The smooth estimates of βk_{it} where k_{it} is the coefficient of attraction to attractor *i* at time *t* are plotted as solid black lines. Dotted lines are drawn two standard errors above and below the estimate of the smooth. The *x*-axis is the day of the year.

steps for the sake of model comparison, but our varying coefficient model could be fit to irregular time steps as well.

Ecologists could use our flexible modeling framework described in this paper to better understand movement behavior in less-stereotyped individuals. Ultimately, however, it is important to make inference at the population level to inform conservation efforts. We have taken the first step toward population level inference by defining a model that is flexible enough to fit all individuals in the population including individuals that do not easily conform to pre-specified movement strategies. The next step is a structured population level model, which would require a model for the locations of attractors and could incorporate covariates. One could also build a model using individual random effects with population level means to make population level inferences.

Attractive points were fixed in our model, but the attractors could be modeled at the population level by implementing clustering approaches or spatial point process models. Clustering approaches such as k means clustering of all attractors across all individuals could identify popular attractive regions. Incorporating covariates in the model is important to ecologists who want to understand



n

Fig. 21. As in Fig. 20, smooth estimates are plotted with dotted lines two standard errors above and below.

why movement decisions are made by individuals. Covariates could be included in a model for the attractors by utilizing a spatial point process model for the attractors (Warton and Shepherd, 2010). To model speed as a function of spatial covariates, a motility function could be included as in Russell et al. (2018) but allowing the motility function to depend on covariates.

While latent-state models are more popular in the animal movement literature, our work shows that our varying coefficient model could help researchers better understand the nuances inherent in animal telemetry data. While the latent-state models must be tailored to each movement strategy, our varying coefficient model does not need such adjustments. The implementation of a general SDE framework using potential functions with varying coefficients controlling the degree of attraction or repulsion to given attractors allows for a reasonable model fit for many movement strategies.

Appendix A. Justification for daily observations

We chose to use daily observations in our analysis since the patterns we were interested in modeling (e.g., migration, dispersal) took place over several days at the least. We did not use a



E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx



Fig. 22. As in Fig. 20, smooth estimates are plotted with dotted lines two standard errors above and below.

finer resolution because of the inconsistent resolution of the original data and since we were not interested in more fine scale behaviors. The choice of sampling resolution is an important one and can be explored by fitting the same model to the data at different resolutions.

We fit the varying coefficient model for the three paths described in Sections 5–7 at two different resolutions: daily and every 2 days. Resulting potential surfaces are shown in Figs. 8–10. In each case, attractors chosen with k means were located in very similar locations at both resolutions. The potential surfaces are centered at zero to facilitate a fairer comparison. For the resident and migrant, potential surfaces are very similar when we use daily observations or observations every 2 days (Figs. 8–9). For dispersal, the move from one attractor to the other occurs over about 5 days, so we would lose that event by using one observation every 2 days (Fig. 10).

Appendix B. Additional plots for single-state resident model

See Figs. 11 and 12.

E. Eisenhauer, E. Hanks, M. Beckman et al.

Spatial Statistics xxx (xxxx) xxx

Appendix C. Additional plots for latent-state dispersal model

See Figs. 13-15.

Appendix D. Additional plots for latent-state migrant model

See Figs. 16–19.

Appendix E. Varying coefficients

See Figs. 20–22.

References

- Ball, J.P., Nordengren, C., Wallin, K., 2001. Partial migration by large ungulates: characteristics of seasonal moose alces alces ranges in northern Sweden. Wildlife Biology 7 (3), 39–47.
- Brillinger, D.R., Preisler, H.K., Ager, A.A., Kie, J., 2012. The use of potential functions in modelling animal movement. In: Selected Works of David Brillinger. Springer, pp. 385–409.
- Brown, J.L., Bedrosian, B., Bell, D.A., Braham, M.A., Cooper, J., Crandall, R.H., DiDonato, J., Domenech, R., Duerr, A.E., Katzner, T.E., et al., 2017. Patterns of spatial distribution of golden eagles across North America: how do they fit into existing landscape-scale mapping systems? J. Raptor Res. 51 (3), 197–215.
- Cagnacci, F., Focardi, S., Ghisla, A., Van Moorter, B., Merrill, E.H., Gurarie, E., Heurich, M., Mysterud, A., Linnell, J., Panzacchi, M., et al., 2016. How many routes lead to migration? Comparison of methods to assess and characterize migratory movements. J. Anim. Ecol. 85 (1), 54–68.
- Chapman, B.B., Brönmark, C., Nilsson, J.-Å.k., Hansson, L.-A., 2011. The ecology and evolution of partial migration. Oikos 120 (12), 1764–1775.
- Eisenhauer, E., Hanks, E., 2020. A lattice and random intermediate point sampling design for animal movement. Environmetrics e2618.
- Fullman, T.J., Person, B.T., Prichard, A.K., Parrett, L.S., 2021. Variation in winter site fidelity within and among individuals influences movement behavior in a partially migratory ungulate. PLoS One 16 (9), e0258128.
- Hanks, E.M., Johnson, D.S., Hooten, M.B., 2017. Reflected stochastic differential equation models for constrained animal movement. J. Agric. Biol. Environ. Stat. 22 (3).
- Hooten, M.B., Buderman, F.E., Brost, B.M., Hanks, E.M., Ivan, J.S., 2016. Hierarchical animal movement models for population-level inference. Environmetrics 27 (6), 322–333.
- Kloeden, P.E., Platen, E., 1992. Stochastic differential equations. In: Numerical Solution of Stochastic Differential Equations. Springer, pp. 103–160.
- Marra, G., Wood, S.N., 2011. Practical variable selection for generalized additive models. Comput. Statist. Data Anal. 55 (7), 2372–2387.
- McClintock, B.T., Michelot, T., 2018. momentuHMM: R package for generalized hidden Markov models of animal movement. Methods Ecol. Evol. 9 (6), 1518–1530.
- Michelot, T., Glennie, R., Harris, C., Thomas, L., 2020. Varying-coefficient stochastic differential equations with applications in ecology. arXiv preprint arXiv:2008.09111.
- Michelot, T., Langrock, R., Patterson, T.A., 2016. Movehmm: an r package for the statistical modelling of animal movement data using hidden Markov models. Methods Ecol. Evol. 7 (11), 1308–1315.
- Morales, J.M., Moorcroft, P.R., Matthiopoulos, J., Frair, J.L., Kie, J.G., Powell, R.A., Merrill, E.H., Haydon, D.T., 2010. Building the bridge between animal movement and population dynamics. Philos. Trans. R. Soc. B 365 (1550), 2289–2301.
- Mueller, T., Fagan, W.F., 2008. Search and navigation in dynamic environments-from individual behaviors to population distributions. Oikos 117 (5), 654–664.
- Nelson, E., 1967. The ornstein-uhlenbeck theory of Brownian motion. In: Dynamical Theories of Brownian Motion. Princeton University Press, pp. 45–52.
- Patterson, T.A., Parton, A., Langrock, R., Blackwell, P.G., Thomas, L., King, R., 2017. Statistical modelling of individual animal movement: an overview of key methods and a discussion of practical challenges. AStA Adv. Stat. Anal. 101 (4), 399–438.
- Pirotta, E., Katzner, T., Miller, T.A., Duerr, A.E., Braham, M.A., New, L., 2018. State-space modelling of the flight behaviour of a soaring bird provides new insights to migratory strategies. Funct Ecol 32 (9), 2205–2215.
- Poessel, S.A., Bloom, P.H., Braham, M.A., Katzner, T.E., 2016. Age-and season-specific variation in local and long-distance movement behavior of golden eagles. Eur J Wildl. Res. 62 (4), 377–393.
- Preisler, H.K., Ager, A.A., Wisdom, M.J., 2013. Analyzing animal movement patterns using potential functions. Ecosphere 4 (3), 1–13.
- Russell, J.C., Hanks, E.M., Haran, M., Hughes, D., 2018. A spatially varying stochastic differential equation model for animal movement. Ann. Appl. Stat. 12 (2), 1312–1331. http://dx.doi.org/10.1214/17-AOAS1113.
- Russell, J.C., Hanks, E.M., Modlmeier, A.P., Hughes, D.P., 2017. Modeling collective animal movement through interactions in behavioral states. J. Agric. Biol. Environ. Stat. 22 (3), 313–334.



Spatial Statistics xxx (xxxx) xxx

- Spitz, D.B., Hebblewhite, M., Stephenson, T.R., 2017. 'MigrateR': extending model-driven methods for classifying and quantifying animal movement behavior. Ecography 40 (6), 788–799.
- Warton, D.I., Shepherd, L.C., 2010. Poisson point process models solve the" pseudo-absence problem" for presence-only data in ecology. Ann. Appl. Stat. 1383-1402.
- Zucchini, W., MacDonald, I.L., Langrock, R., 2017. Hidden Markov Models for Time Series: An Introduction using R. CRC Press.
- Zucchini, W., Raubenheimer, D., MacDonald, I.L., 2008. Modeling time series of animal behavior by means of a latent-state model with feedback. Biometrics 64 (3), 807–815.