Agent-Based Models for Collective Animal Movement: Proximity-Induced State Switching

Authors: Andrew Hoegh, Frank T. van Manen, and Mark Haroldson

The final publication is available at Springer via https://doi.org/10.1007/s13253-021-00456-0

Hoegh, A., van Manen, F.T. & Haroldson, M. Agent-Based Models for Collective Animal Movement: Proximity-Induced State Switching. Journal of Agricultural, Biological and Environmental Statistics (2021). https://doi.org/10.1007/s13253-021-00456-0

Agent-Based Models for Collective Animal Movement: Proximity-Induced State Switching

May 26, 2021

Abstract

Animal movement is a complex phenomenon where individual movement patterns can be influenced by a variety of factors including the animal's current activity, available terrain and habitat, and locations of other animals. Motivated by modeling grizzly bear movement in the Greater Yellowstone Ecosystem, this article presents an agent-based model represented in a state-space framework for collective animal movement. The novel contribution of this work is a collective animal movement model that captures interactions between animals that can trigger changes in movement patterns, such as when a dominant grizzly bear may cause another subordinate bear to temporarily leave an area. The modeling framework enables learning different movement patterns through a state-space representation with Particle-MCMC methods for fully Bayesian model fitting and the prediction of future animal movement behaviors.

Keywords: Bayesian statistics, MCMC, Particle-MCMC

1 Introduction

Animal movement is a complex phenomenon driven by a variety of factors including interactions with other animals. However, until recently, most statistical models for animal movement did not consider interactions among individual animals. Collective animal movement models, where individual animals interact and influence each other, present a challenge. Part of the challenge is the model specification, which can require complicated non-linear interactions between individuals, whereas the other part is computational, in that jointly considering the simultaneous position of all animals prevents parallel computation and can require evaluating complicated non-linear interactions among individuals.

A variety of approaches are used for animal movement models: one foundation is based on velocity between consecutive measurement fixes (Jonsen et al., 2005), which is often constructed using step length and turning angles. The simplest of these animal movement models use correlated or biased random walks, where the animal's directional heading depends on previous angle or is biased toward a location or direction. Habitat or terrain can influence animal movement, hence, these can also be incorporated into animal movement models (Christ et al., 2008; Peck et al., 2017). For a comprehensive overview of animal movement modeling, readers are referred to Hooten et al. (2017).

More complex animal movement patterns can also be included, such as those that depend on the animal's behavioral state (e.g., exploring or foraging). Langrock et al. (2012) provides an overview of hidden Markov models for discrete time, which enable different turning angles and step lengths for a set of animal states. State switching models can also be useful as an animal can have different movement patterns corresponding to different behavioral states. Morales et al. (2004), Haydon et al. (2008), and McClintock et al. (2012) present frameworks for multistate models that account for different behavioral states. In a related setting, Scharf et al. (2019) models polar bear movement and incorporates a dynamic covariate response to correspond to changing resources, allowing behavior to change as resources change.

Collective animal movement models require jointly evaluating positions of all animals in a

study area; although, models have been implemented in different ways. In the ecology literature, rule-based methods are common. Couzin et al. (2002) presents a rule-based framework based on animals minimizing the distance between each individual coupled with a zone of repulsion. These models can capture many different collective dynamics including swarming behavior often exhibited in insects or parallel groups often exhibited by flocks of birds or schools of fish. Couzin et al. (2005) also presents a framework that allows groups of animals to make collective movement decisions based on social interaction within the group. Ballerini et al. (2008) summarizes rule-based ecological collective movement models by stating, "...the models agree on three general behavioral rules: move in the same direction as your neighbours; remain close to them; avoid collisions." While ecological models for collective movement tend to be rule-based, statistical frameworks for animal movement are generally model-based which enable learning model parameters directly from observed collective movement datasets and can be used to generate k-step ahead predictions given current animal locations. Langrock et al. (2014) uses correlated and biased random walks to develop a collective animal movement model, where animals are attracted to the moving centroid of the group. In addition to attraction to the groups centroid, Langrock et al. (2014)'s framework also incorporates a state that permits the animal to explore independently of the group by temporarily removing or reducing the attraction to the group centroid. Agent-Based Models (ABMs) are another approach for collective animal movement that have gained popularity in the last decade (Hooten and Wikle, 2010; Bonnell et al., 2016; McDermott et al., 2017).

A major benefit of ABMs for animal movement is that complicated interactions between individuals can be modeled in a relatively simple manner. ABMs provide a simulation approach built on behaviors of a group of *agents*. The agents interact with each other based on a set of rules to provide a mechanism for modeling individual-level behavior, but the collective response of the agents can be used to understand population-level characteristics. Individual agents can be calibrated with fairly simple behaviors; but the collective response of the agents can model complex population dynamics. While ABMs have been used in

many situations (for instance, Gilbert and Terna (2000); Brown et al. (2005); Gilbert (2008); Farmer and Foley (2009)), statistical parameter estimation with uncertainty is fairly new, particularly in an animal movement framework.

One challenge with ABMs from a statistical perspective has been capturing the uncertainty in parameter estimation. The individual dynamics and nonlinear functions that result from a set of behavioral rules for agents can make evaluating likelihoods, and thus many likelihood-based or Bayesian approaches, difficult or impossible. One recent approach (McDermott et al., 2017) for ABMs with animal movement models used Approximate Bayesian Computation (ABC) (Beaumont et al., 2002; Beaumont, 2010) for model fitting. In addition to ABC approaches, Particle-MCMC (P-MCMC) methods (Andrieu et al., 2010) are also recommended for model fitting with nonlinear state space models (Fasiolo et al., 2016); however, we are not aware of their use in ABMs for collective animal movement.

The motivating research goal for this work is to calibrate an ABM to capture the collective dynamics of grizzly bear (*Ursus arctos*) movement with a focus on long term land use patterns. With collective grizzly bear movements and the associated ABMs, there are several features that inform model choice, which is specified through the agent rules. The collective behavior of grizzly bears is fundamentally different than schooling behavior modeled with a self-propelled particle (Vicsek et al., 1995) as formulated in McDermott et al. (2017) or the herd-like behavior in Langrock et al. (2014), and thus, requires a new model formulation. Dominance hierarchies in grizzly bears can induce repulsive forces; however, there may also be times when bears are attracted to each other: during breeding season, for family dynamics, or when there is a local abundance of food resources. Hence, models of collective behavior should allow for both attractive and repulsive forces. Furthermore, the models can include hierarchical treatment of movement patterns that allow each individual bear to have its own movement parameters. In addition to specifying an ABM for collective grizzly bear movement, another novel contribution of this modeling framework is the inclusion of a Markov switching component where changes in behavioral states can be influenced

by other bears.

This article extends existing ABMs for animal movement models through the lens of state-space models and introduces P-MCMC methods for computing those models. Whereas the application focuses on grizzly bears in the Greater Yellowstone Ecosystem (GYE), the methods are applicable to a range of species and contexts. In Section 2 we describe the structure of the data, Section 3 provides an overview of our ABM modeling framework for grizzly bears, Section 4 details computation for fitting these models, Section 5 provides a set of simulation studies and an analysis of grizzly bear movement data in the GYE, and Section 6 concludes with a discussion.

2 Data Overview

The dataset used for this analysis contains telemetric GPS locations of 91 male grizzly bears in the GYE from 2005 – 2015. Captures were conducted under U.S. Fish and Wildlife Service Endangered Species Permit [Section (i) C and D of the grizzly bear 4(d) rule, 50 CFR17.40(b)], with additional permits from the National Park Service, Wyoming, Montana, and Idaho, and conformed to the Animal Welfare Act and to U.S. Government principles for the use and care of vertebrate animals used in testing, research, and training (USGS ACUC no. 201201). The bears are instrumented with telemetric tracking devices that record consecutive location fixes. Given the crepuscular behavioral patterns that bears exhibit, the location fixes are filtered to include one data point from the morning twilight hours and one from evening twilight hours. In particular, the location fixes nearest to 7:30 AM and PM on each day are retained for the analysis. This corresponds to time periods when bears are most active and hence, most likely to interact with other animals. Models are fit using data from May through August and predictions are made for September of each year. Data are recorded until the tracking devices failed, were shed as scheduled using pre-set drop-off mechanisms, or were removed during recaptures. Some of bears were recaptured and collared at a later

date, so it is possible to have location fixes that are separated by many years. Bears without a minimum of 20, twice-daily consecutive location fixes are removed from the dataset.

Grizzly bear locations, and animal movement in general, are best visualized with animation, such as using the anipaths package in R (Scharf, 2019); however, Figure 1 contains static summaries of grizzly bear locations created in R with packages from Pedersen (2019), Wickham (2016), and Kahle and Wickham (2013). Figure 1 contains telemetric locations of bears observed in 2014 and 2015. Each bear in a given year is represented in a different color and locations are bounded with a convex hull to capture an approximate home range.

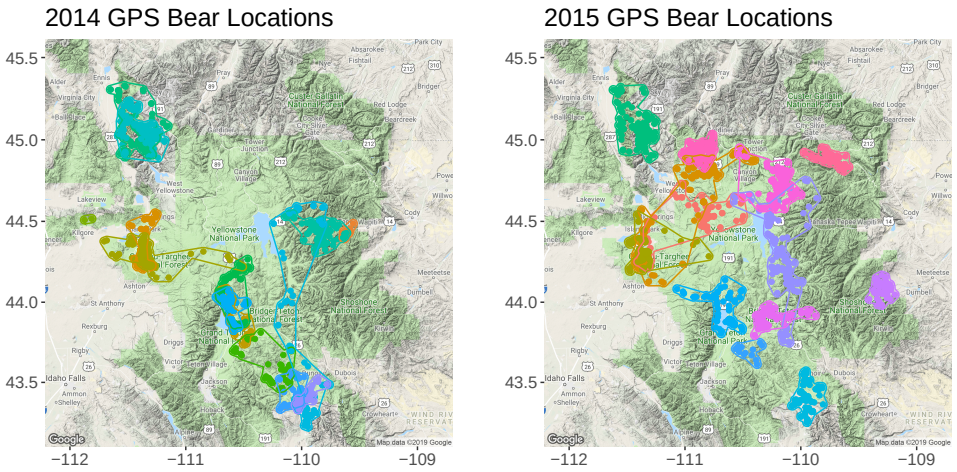


Figure 1: Locations of male grizzly bears during 2014 - 2015. While there are similar shades of colors, the same color across both years indicates a bear with GPS location data for both years. The outlines represent approximate range of each bear, constructed with a convex hull.

Past research has shown that different behavioral states can be inferred from telemetric movement patterns (Ebinger et al., 2016). For instance, when bears are exploring and searching for new territory the step size will tend to be larger and the turning angles, between consecutive fixes, will be smaller. In other words, the bears tend to move in the same heading for an extended time. This can be seen in Figure 2, which displays the turning angle as a function of step size to illustrate differences associated with smaller versus larger movement steps. With the smaller steps, bears tend to have a more uniform turning angle, with addi-

tional mass near 180 degree turns that would correspond to returning to a previous location. With larger step size, bears tend to maintain an existing heading, as a large proportion of the turning angles reside within 60 degrees of 0. These longer, directed movements are often associated with attracting or repulsive forces (e.g., mating, avoidance of dominant animals or areas of high density, dispersal) or shifts in resource use.

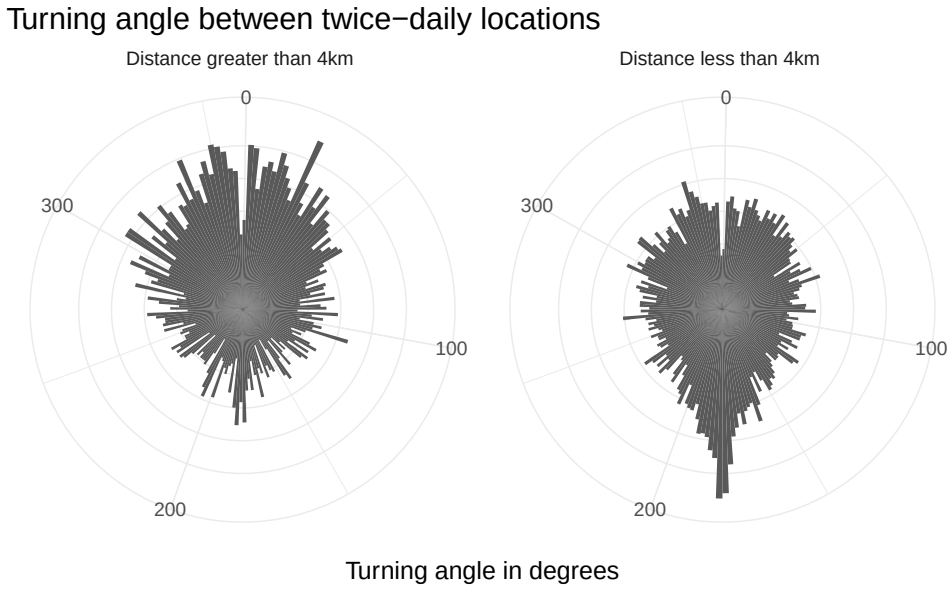


Figure 2: Turning angles associated with step sizes less than or greater than 4 kilometers. The longer step sizes tend to maintain the previous heading, while the shorter step sizes tend to have a more uniform distribution, with a large number of movement patterns around 180 degrees, which would correspond to returning to a previous location. A distance of 4 kilometers represents about the 80th percentile of step sizes, but the figure is fairly robust to changes in the percentile value.

3 Modeling Framework

The main contribution of this article is to formalize a state-space framework for ABMs with collective animal movement focusing on species that do not exhibit schooling or herd-like behavior. The article also details particle methods for fully Bayesian computation. The

modeling framework enables collective animal movement models that incorporate attractive-repulsive forces in a behavioral switching model along with a hierarchical model structure.

3.1 Model Specification

Building on the ABM model specification in McDermott et al. (2017), we describe our model through a state-space framework and add additional features necessary to capture the complex dynamics that drive grizzly bear movement. Given the applicability for general animal movement models, the model specification is explained in general terms using agents in place of bears.

Observation Equation

The observation equation is specified using model notation from McDermott et al. (2017). Assume that spatial locations of the $i = 1, .., N$ agents are obtained at discrete times, $t = 1, ..., T$. Then the observed locations at time t are defined as $\underline{\mathbf{z}}_\mathbf{t} \equiv \{\mathbf{z}'_{\mathbf{i,t}}\}_{i \in 1,...,m_t}$, where $\mathbf{z}'_{\mathbf{i,t}} \equiv (x_{i,t}, y_{i,t})$ and $x_{i,t}$ and $y_{i,t}$ are the coordinates of the i^{th} agent at time t. The locations may be observed for an m_t subset of the N agents, meaning all of the agents may not have location history for all T time points. In practice, this could result from a collar being enabled or disabled during the study period or the failure to record a location at that time. Similarly, the latent locations of the agents are denoted as $\underline{\mathbf{s}}_\mathbf{t} \equiv \{\mathbf{s}'_{\mathbf{1,t}}, ..., \mathbf{s}'_{\mathbf{N,t}}\}$, where $\mathbf{s}_{\mathbf{i,t}} \equiv (\tilde{x}_{i,t}, \tilde{y}_{i,t})$ and $\tilde{x}_{i,t}$ and $\tilde{y}_{i,t}$ are the latent coordinates of the i^{th} agent at time t. Then the observation equation is expressed as:

$$\underline{\mathbf{z}}_\mathbf{t} = \mathbf{H}_\mathbf{t}\underline{\mathbf{s}}_\mathbf{t} + \underline{\epsilon}_\mathbf{t}, \qquad \underline{\epsilon}_\mathbf{t} \sim N(\mathbf{0}, \sigma^2_\epsilon I), \tag{1}$$

where $\mathbf{H}_t$ is a $2m_t \times 2N$ incidence matrix of zeros are ones used to account for agents that are are not observed at time t. The measurement error in the model is represented by σ^2_ϵ, which in this case incorporates uncertainty from both the GPS signal as well as impacts from

aligning the bear location fixes as measurements do not occur at precisely the same time.

Evolution (State) Equation

In the state-space literature, the terms state equation and evolution equation are used interchangeably. Here we use evolution equation to avoid confusion between the state equation and the behavior states of the switching model. The evolution equation for an individual agent, agent i in this case, is specified as

$$\underline{\mathbf{s}}_{\mathbf{i},\mathbf{t}} \sim N\left(\underline{\mathbf{s}}_{\mathbf{i-1},\mathbf{t}} + u_{i,j,t}\underline{\delta}_{\mathbf{i},\mathbf{j},\mathbf{t}}, \sigma_\eta^2 I\right), \tag{2}$$

where the term $u_{i,j,t}$ is a scalar variable representing agent i's speed when in behavioral state j at time t, $\underline{\delta}_{\mathbf{i},\mathbf{j},\mathbf{t}} = (\delta_{x,i,j,t}, \delta_{y,i,j,t})$ is a unit vector describing the directional component of agent i's velocity when in state j at time t, and σ_η^2 is a random perturbation. The speed and angular direction can be jointly estimated, or as is often the case, estimated separately. The speed and angle parameters largely contribute to animal movement, but the direction of home range and the presence of other individuals can also inform an agent's speed and angular heading. In our particular setting speed is defined as the average distance between the location fixes that are 12 hours apart, which can be different than the actual distance traveled by an agent in that time window. Velocity combines the speed with an angular direction. We chose 12-hr location fixes because the research question motivating this work is focused on longer-term dynamics and range. Furthermore, centering the location fixes near crepuscular hours reflects times when grizzly bears are most active, and thus most likely to interact. Among grizzly bears, primary drivers of interactions are not necessarily direct, physical encounters but involve indirect attracting and repulsive forces that manifest themselves at coarse temporal and spatial scales. Such interactions are mediated, for example, through scent communication or other cues that indicate the presence or density of conspecific bears.

The directional component of velocity is determined by an angle parameter $\theta_{i,j,t}$ such

that $(\delta_{x,i,j,t}, \delta_{y,i,j,t}) = (\cos(\theta_{i,j,t} + \nu_{i,j,t}), \sin(\theta_{i,j,t} + \nu_{i,j,t}))$ is modeled as

$$\theta_{i,j,t} \sim PN(\underline{\mu_{\theta,j}}, \Sigma_j), \tag{3}$$

where $PN()$ is a projected normal distribution, $\underline{\mu_{\theta,j}}$ is a mean vector, and Σ_j is a covariance matrix. The projected normal distribution transforms Cartesian coordinates to polar coordinates to acquire the angular values. The parameters in the projected normal distribution can be modeled as a function of other parameters, such as home range direction, location and turning angle of other agents, and/or behavioral state. The $\theta_{i,j,t}$ variable can be combined with another angle, $\nu_{i,j,t}$, that centers the direction and creates a biased random walk where agents tend to stay in the same home range or on a existing heading.

The speed model is specified as:

$$u_{i,j,t} \sim LN(\mu_{u,j}, \sigma^2_{u,j}) \tag{4}$$

where $LN()$ is a log normal distribution. Different frameworks can be used to model speed; however, a log normal distribution fits well with the data and also provides computational advantages of Gibbs samples inside the P-MCMC framework.

With a switching model, transition probabilities between the behavioral states also need to be specified. In this work, a two-state model is introduced, but the model framework is flexible and could permit multiple states. We use a predefined threshold where agents are aware of surrounding agents to induce different movement patterns. Based on biological understanding of grizzly bear's sphere of influence and uncertainty induced by using 12-hour location fixes, the threshold for our data analysis is set to 10 km. In situations where scientific understanding is not available to inform this distance, the threshold could be calibrated using predictive results or estimated by placing a prior distribution on this parameter. The two-state model requires estimating four probabilities $\underline{p} = (p_{12}[c], p_{12}[f], p_{22}[c], p_{12}[f])$, where $p_{i,j}$ is the switching probability from state i from state j. The proximity of other agents are

specified in a binary framework using $[c]$ for close and $[f]$ for far away. Using conditionally conjugate beta priors allows efficient samples from the posterior distributions.

Hierarchical structure can allow movement patterns to differ by agents within a population. For instance, building off the model for speed in Equation 4, rather than the same speed term for all agents, $\mu_{u,j}$, the speed term could be formulated in a hierarchical model so that each agent has a different speed term, $\mu_{u,i,j}$. This would result in speed parameters that can vary by agent:

$$\mu_{u,i,j} \sim N(\mu_{u,j}, \tau_j^2), \tag{5}$$

where $\mu_{u,j}$ would be the population level step length parameters, whereas $\mu_{u,i,j}$ could vary by agent. Similarly, angle and transition probabilities can also be modeled in a hierarchical framework.

To complete the Bayesian specification, prior distributions are needed for the following parameters: σ_ϵ^2, σ_η^2, Σ_j, $\mu_{\theta,j}$, $\mu_{u,j}$, σ_u^2, and $\underline{p}$. In general, conjugate and semi-conjugate priors are used. Subject matter knowledge can be used to help inform hyperparameters for the prior distributions. For instance, biological knowledge can be used for determining reasonable distances a bear could travel in 12 hours. The appendix contains the complete model formulation, including prior specification, for the data analysis of grizzly bears in the GYE.

4 Computation

Fasiolo et al. (2016) outlines two common approaches for highly complex, nonlinear state-space models: ABC and P-MCMC. The extended Kalman filter has also been used in this setting (Orderud, 2005). The challenge in state-space models, such as this ABM formulation, is to jointly sample from the model parameters, denoted Θ, along with the state parameters, denoted $\mathcal{X}$. With the model specified in Section 3.1, there are a large number of state variables that need to be estimated. In particular, for each agent and time point, a step size

variable, turning angle variable, and a behavioral state parameter are required, which are combined with noise to result in two latent location parameters, a latitude and longitude. Thus, estimating the state parameters, which depend on Θ, requires jointly evaluating many state variables. Hence, exploring the space of the joint posterior distribution

$$p(\Theta, \mathcal{X}|\mathcal{Z}), \tag{6}$$

where $\mathcal{Z}$ contains the observed locations of the agents, is difficult.

Approximate Bayesian Computation (ABC) is used for scenarios where evaluating the likelihood is difficult or impossible. Rather than directly evaluating the likelihood, a dataset is generated with the mechanistic model using a proposed value of the parameter set Θ. Then, a summary statistic, or set of summary statistics, are computed for the generated dataset. This summary statistic on the generated data is compared with a summary statistic on the observed data. If the distance between the summary statistics is sufficiently small, then the proposed value of Θ is accepted; otherwise, the proposed value is rejected. McDermott et al. (2017) presents a hybrid MCMC algorithm for jointly estimating additional model parameters along with the ABC approach for state parameters.

An alternative to ABC methods involves using P-MCMC where particle-based methods are used for MCMC proposals. One benefit of P-MCMC, versus ABC, is that the estimation is fully Bayesian, rather than approximate. As stated, the challenge with fitting this model is jointly estimating the model parameters and state variables in the joint posterior distribution, $p(\Theta, \mathcal{X}|\mathcal{Z})$. Using a standard Metropolis-Hasting approach for this model can be difficult, given the high dimensionality and correlation in the state variables. Particle MCMC uses particle filtering techniques, Sequential Monte Carlo (SMC), to propose $p(\mathcal{X}|\mathcal{Z}, \Theta)$. In other words, given the parameter values, Θ, a particle filter is used to propose a set of state values.

The common criticism of SMC techniques is that the methods suffer from particle degeneracy as the number of distinct samples becomes small enough that the estimates of the

joint likelihood are unreliable. This issue is exacerbated as the total number of time points becomes larger and model fitting requires a, perhaps unfeasible, large number of particles. The problem is lessened in P-MCMC as Andrieu et al. (2010) state "PMCMC sampling is much more robust and less likely to suffer from this depletion problem. This stems from the fact that PMCMC algorithms do not require SMC algorithms to provide a reliable approximation of [the sampling distribution of the state variables], but only to return a single sample approximately distributed according to [sampling distribution of the state variables]." With that said, with the P-MCMC algorithm we opted to use the Particle Marginal Metropolis-Hastings approach rather than the Particle Gibbs approach as we found better mixing of the state parameters. Additionally, a much longer time series, such as using hourly data rather twice-daily location fixes and/or a large increase in the number of agents, would certainly tax the algorithm.

P-MCMC Algorithm

Let Θ be a set of model parameters, which in this case have closed form full conditional distributions. An overview of a basic P-MCMC algorithm can be seen in Algorithm 1.

---

for *iter = 1 to M* **do**

1. Sample $\Theta|-$ from full conditional distributions ;
2. Propose $\mathcal{X}^*$ using particle filter conditional on $\Theta^{(iter)}$;
3. Accept $\mathcal{X}^*$ according to $p(\mathcal{X}^*|\mathcal{Z}, \Theta^{(iter)})$ with the standard Metropolis-Hastings acceptance ratio

end

Algorithm 1: Basic P-MCMC Algorithm (Particle Marginal Metropolis-Hastings)

A detailed algorithm for the models shown in Section 3.2 can be found in the appendix and computer code is provided in the supplemental materials section. A rather obvious drawback of P-MCMC is computing time, as each iteration of an MCMC algorithm requires a particle

filter with a large number of particles.

In all scenarios with weakly informative prior specification, the P-MCMC approach is considerably easier to tune than ABC, in that it does not require as many researcher choices. In particular, ABC requires the researcher to choose: 1.) a summary statistic, 2.) a distance metric, and 3.) a kernel associated with that distance between the observed and simulated data. While P-MCMC model fit is based on the statistical likelihood, ABC can be calibrated on metrics that are not part of the likelihood, such as some summary of collective movement. Complicating factors for ABC include that the summary statistics are typically application dependent, and even in the simulation studies conducted in this article, where the true behavior parameters are known, it is challenging to calibrate the algorithm to return accurate results. In contrast, the only parameter to tune in the P-MCMC framework is the number of particles, which controls the acceptance rate of the Metropolis proposal for the state parameters. The trade off is that P-MCMC approach requires fitting a SMC algorithm with a large number of particles for every iteration of the MCMC sampler, which presents a computational burden.

Given the computational requirements of the P-MCMC approach, the algorithms are implemented in the programming language Julia (Bezanson et al., 2017). Julia has been shown to execute similar programs several orders of magnitude quicker than R and results in a substantial reduction in the analysis time. In our experience and with our coding, the particle filter component of the model is roughly an order of 50 - 100 times faster in Julia than in R.

5 Model Demonstration

5.1 Synthetic Data Analysis

With this section we simulate data similar to the grizzly bear data to determine the efficacy of estimating the model switching parameters and making predictions of future locations.

In particular, we simulate synthetic bear data from the models in Equations 1 - 6 and then evaluate the ability of P-MCMC to recover model parameters and predict future movement. Finally the predictive capacity of models with and without collective behavior are compared.

Typically a manuscript makes comparisons between a proposed method and the best existing alternatives. In this case, existing agent-based collective movement models have been primarily designed for animals with herding or schooling behaviors, neither which would be applicable to grizzly bears. McDermott et al. (2017) present a general approach for collective animal movement, but, as previously discussed, the ABC approach presents difficulties in choosing summary statistics, kernels, and tuning. Hence, the logical comparison is with a model that does not implement any collective movement characteristics, that is one that uses the principles of correlated random walks, but does not consider locations of other animals.

Four different simulation scenarios are conducted: the collective movement signal varies across the first three simulations and the fourth has no collective movement. Weakly informative, conjugate priors are used in both simulations. Complete specification and Julia code for re-creating the synthetic data analyses are available in the supplemental materials. For each scenario, 90 time points for 30 synthetic bears are simulated. Another 120 time points are reserved for predictive comparisons. Each simulation setting is replicated twenty different times. For simulation 1, the switching behavior is largely influenced by proximity of other agents and the movement patterns are well separated between the two states. For simulation 2 and simulation 3, the switching behavior as a function of agent proximity becomes less pronounced. Similarly, for simulation 2 and simulation 3, the movement patterns become progressively less separated between the two states. For simulation 4, there is no collective movement in the simulated data. The parameters used to generate the movement data can be seen in Figure 4. Across these four simulated scenarios, we compare the ability of the two model specifications to estimate the model parameters, with special attention given to the model switching behavior, and also compute the accuracy of k-step ahead predictions.

5.1.1 Parameter Estimation

Simulated paths for simulation 2, which has a moderate collective movement signal, are shown in Figure 3. The shape icon denotes whether or not the agent is in the exploratory

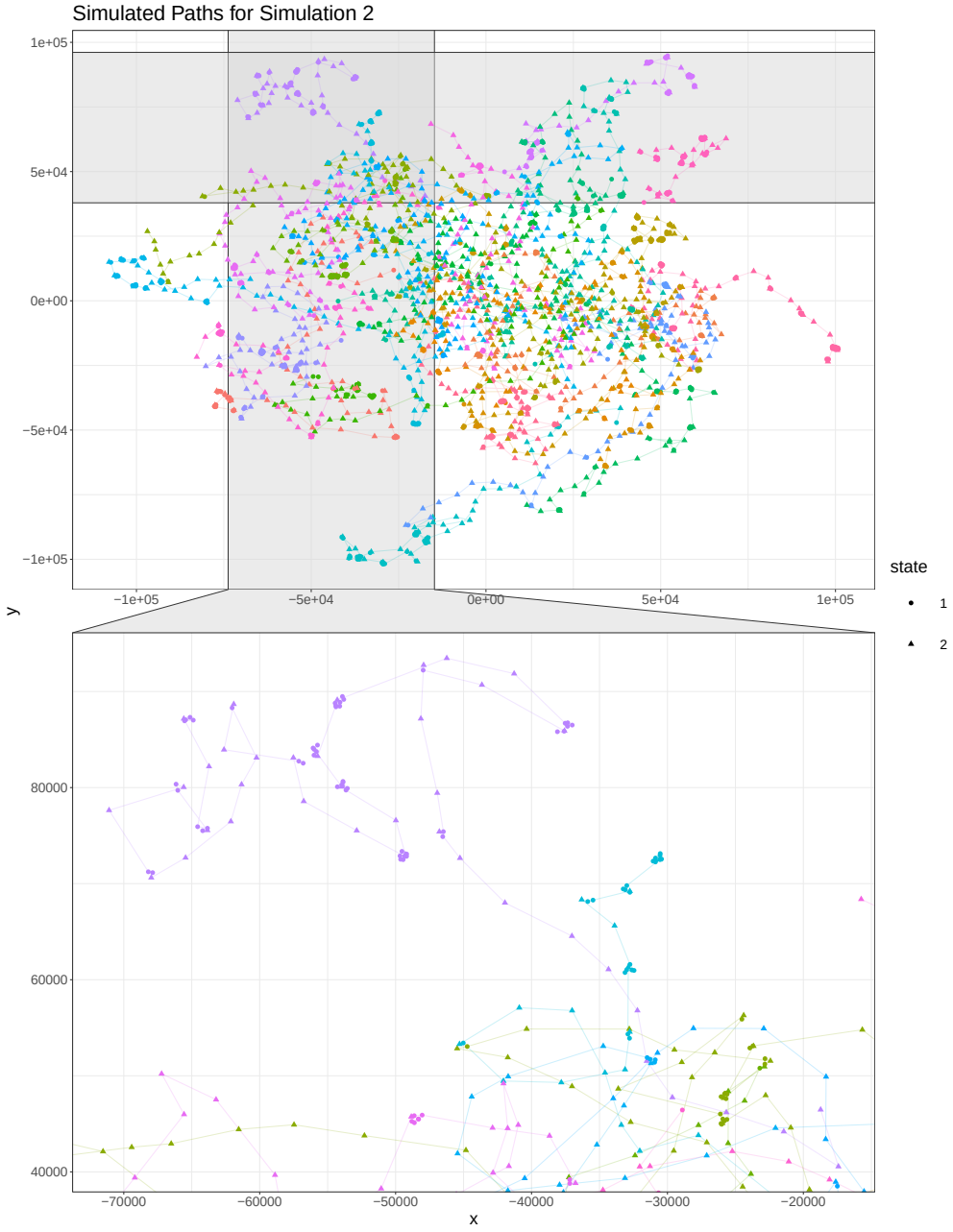


Figure 3: A sample of simulated bear paths (top). Close up look at simulated paths of bear id 3 (bottom), which is shown in purple. State two, the triangles, corresponds to exploring.

behavioral pattern and the lines are estimated paths. These synthetic data are from one

realization of simulation 2. From this figure, the inertia of the synthetic bears is apparent, as the bears tend to stay either in state one or state two. What isn't as obvious, without the benefit of animation, is that when bears are in proximity it tends to result in transitions to state 2, with larger steps and similar heading to previous steps.

The true parameters as well as estimates from the simulation realizations can be seen in Figure 4. In general, the parameters are similar for the two models and match the values from the simulated data. Both modeling frameworks also tend to identify the states fairly accurately. However, the notable exception is the switching probabilities. In the three simulation scenarios with collective movement, the independent movement model does a poor job of estimating the switching probabilities. This is not surprising as there is not a mechanism in those models to switch states based on other agents in the area.

Without a mechanism for accounting for collective movement, or the proximity of other agents, the independent model does not capture the switching probability when agents are inside a threshold that causes behavioral changes. The proportion of time points where the agents are close together is fairly small, but these events are important drivers of land use. Furthermore, the predictive densities, when agents are close together, can be greatly influenced by failing to account for agent proximity and differing behaviors. Figure 5 contains an illustrative example, using the estimated parameters from simulation 3, where the independent model fails to account for the proximity of the nearest agent.

5.1.2 Predictive Densities

To evaluate the k-step ahead prediction we use the mean absolute deviation (MAD) between the observed locations and the mean of the posterior predictions. The expected difference in MAD between a collective movement model and the independent model can be seen in Figure 6. The expected difference in MAD is calculated using the twenty replicates for each simulation setting. In all three collective movement scenarios, simulations 1, 2, and 3, the collective movement model contains a predictive advantage. For simulation 1, where

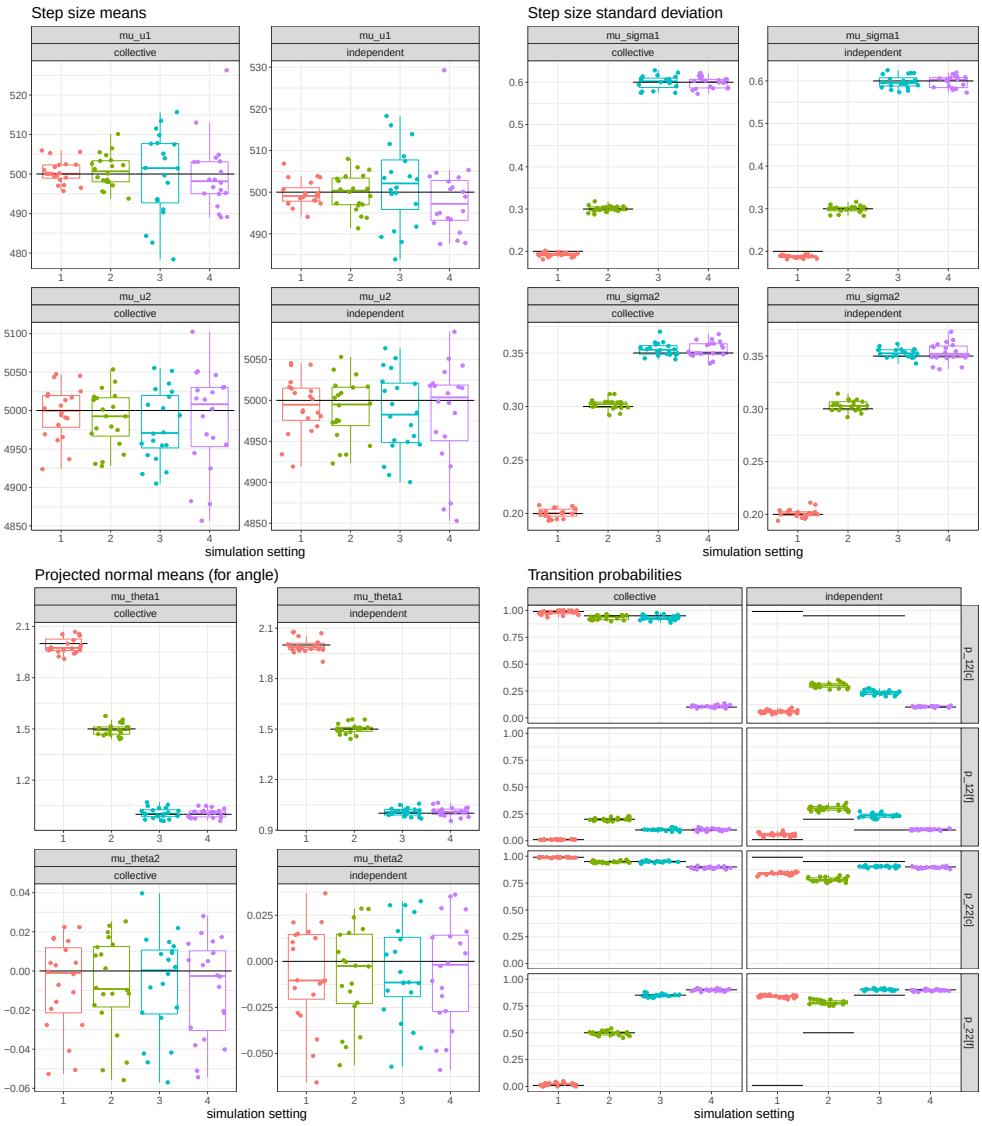


Figure 4: Known parameters from the simulation (black bars) and simulated results for a set of twenty simulated realizations are shown. The step size means, step size standard deviations, and angle parameters are estimated fairly precisely for both the collective movement and independent movement models. The key difference is the independent model's inability to estimate different behaviors when other agents are in close proximity. The top panel of the bottom right figure corresponds to the probability of switching to an exploring state when other agents are close by.

the collective movement signal is strong and the synthetic bears tend to stay in the same state for longer periods, the prediction improvement continues to grow as k increases. For simulations 2 and 3, which have more moderate, and likely realistic, collective movement, the

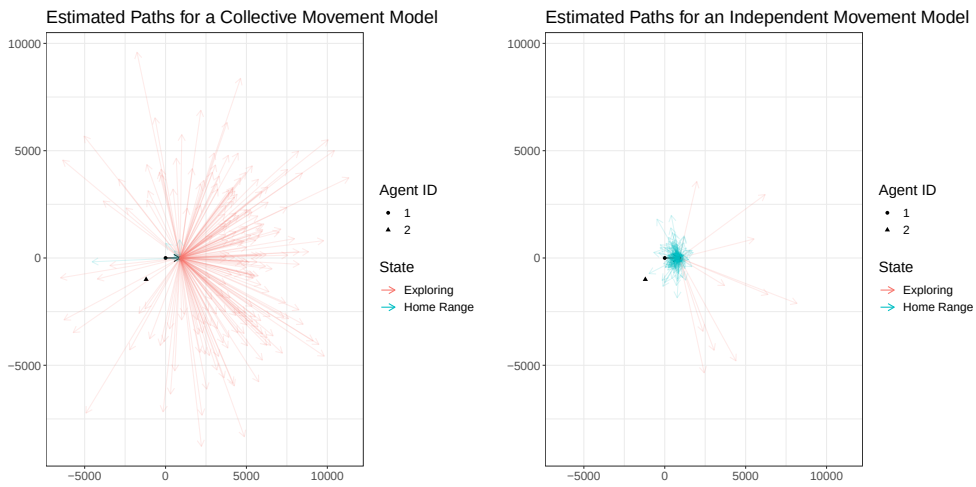


Figure 5: This figure demonstrates the model switching behavior in the collective movement model. The proximity of an other agent, denoted with a triangle, switches the exploratory behavior for the collective movement model, whereas, the independent movement model stays largely in the home range state.

prediction increases level out, or increase minimally, after the first few time periods. Finally, in the final scenario with no collective movement, there are no discernible differences in the predictive results between the collective movement and independent models.

For the first three simulation settings, the predictive difference is magnified when synthetic bears are in close proximity to other individuals. The difference in MAD between collective movement models and and independent models is larger, for the first step or two, than the average across all sythetic bears that is displayed in Figure 6. A demonstration of the differences between the two models, when other bears are in close proximity, can be seen in Figure 5, which shows a set of predicted paths for an agent that has another agent in close proximity. Using the estimated parameters from the model, the collective movement model would largely result in a state switch to the exploratory state, whereas, with the independent model bears would tend to stay in the home-range state.

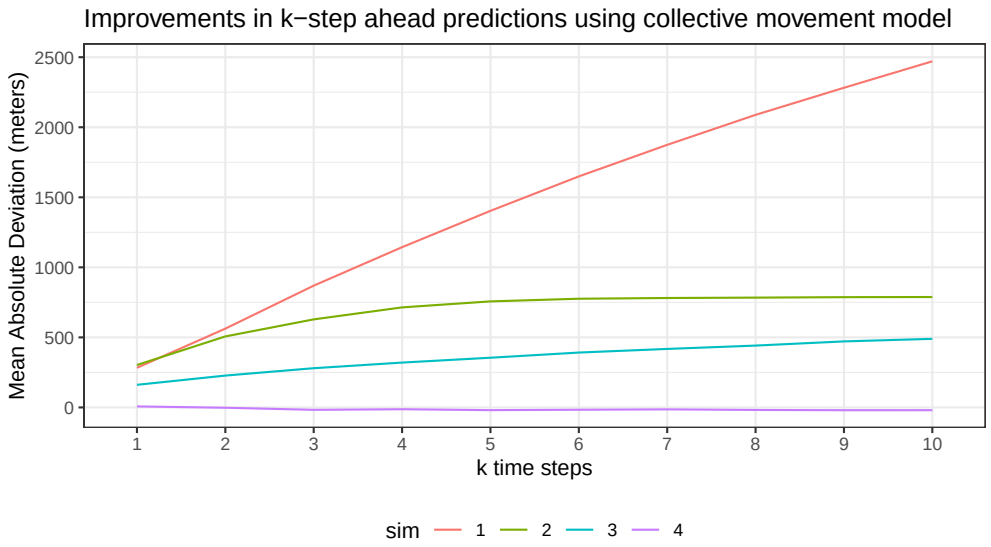


Figure 6: The expected improvement in Mean Absolute Deviation (MAD) for a collective movement model relative to the independent model is shown. With the three collective movement settings (simulations 1 - 3), there is a predictive improvement, but the improvement is lessened as the collective movement model becomes weaker. With no collective movement, in simulation 4, there is not a discernible difference between the two approaches.

5.1.3 Summary

For the simulation settings when collective movement exists in the data, the collective movement model improved parameter estimation and predictive performance. In simulation 4, when no collective movement exists, there are little differences between the two models. When collective movement does occur, the collective movement model is superior. When collective movement is not present, the model does not do appreciably worse.

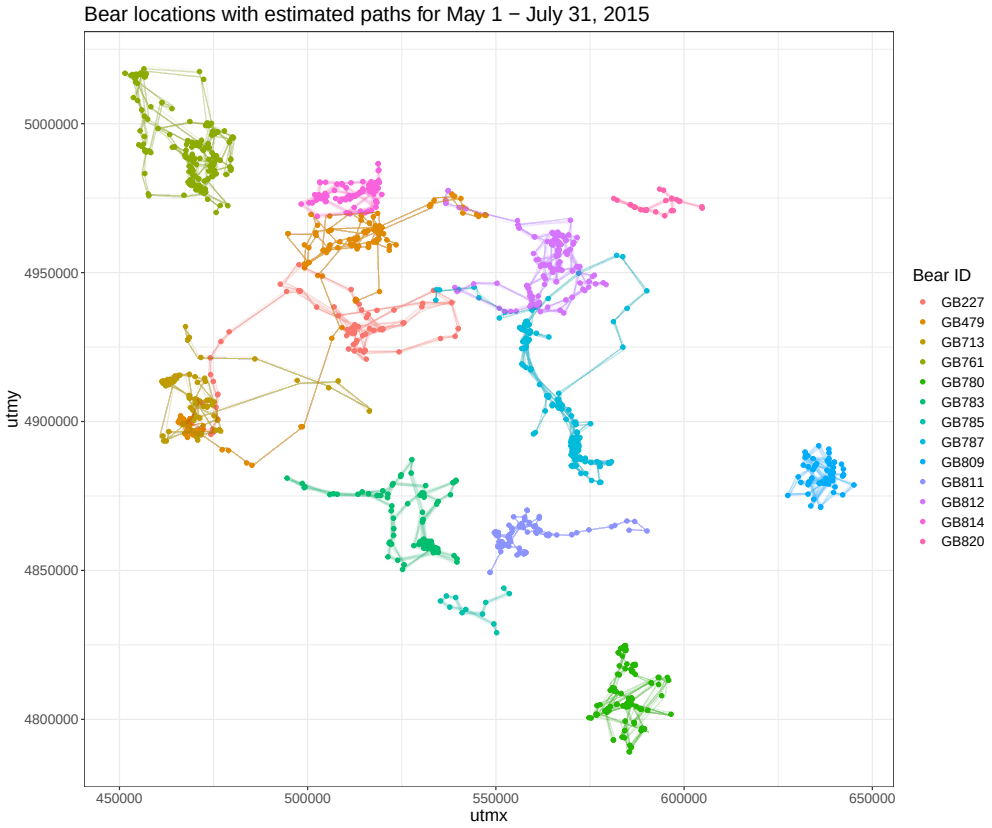
5.2 Application: Grizzly Bear Movement

The collective movement model is applied to the grizzly bear data described in Section 2. With this model, we find bears spend 47.1 (46.5, 47.5) percent of their time in the state with shorter steps oriented toward a moving average of previous positions and the remaining 52.9 (52.5, 53.5) percent of the time in the state with longer steps oriented on the same heading. While the movement types may suggest associated behavior states, such as ”home range” and

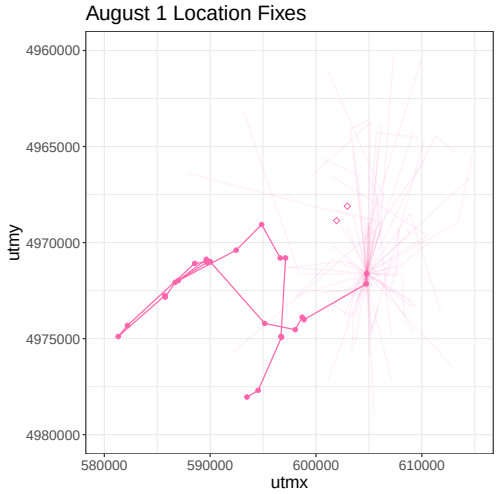
343 "exploring," that correspond to the states of the model, but without labeled observational
344 data the suggested behavioral states do not necessarily have a one-to-one correspondence
345 with the movement patterns. The posterior means for the step size parameters for the
346 lognormal distribution are $\mu_u = 5.78$ (5.70, 5.86) and $\sigma_u = 1.06$ (1.02, 1.10) for the "home
347 range" state and $\mu_u = 8.29$ (8.27, 8.31) and $\sigma_u = 0.72$ (0.70, 0.74) for the "exploring" state,
348 which results in median step size of around 300 and 4000 meters, respectively. We note we
349 use the term "exploring" in a broad biological sense of movements influenced by the presence
350 of other bears while remaining within their respective home ranges. The parameters for
351 the projected normal distribution, which correspond to the angular heading are estimated
352 as 0.56 (0.52, 0.59) and -0.05 (-0.09, 0.02). This suggests, unsurprisingly, that the angular
353 heading of the bears is more uncertain than in the simulation setting.

354 The bears tend to switch states at a higher rate than the simulated datasets. Furthermore,
355 there is evidence that switching behavior does change when in the proximity of other bears.
356 The switching probabilities are $p_{12[f]} = 0.32$ (0.31, 0.34), $p_{12[c]} = 0.39$ (0.35, 0.43), $p_{22[f]} =$
357 0.72 (0.70, 0.73), and $p_{22[c]} = 0.68$ (0.60,0.69). In particular, a bear in "home range" state
358 is more likely to switch to "exploring" when in the proximity of other bears. The credible
359 intervals associated with the switching probabilities provide evidence in favor of using the
360 collective movement model in this situation.

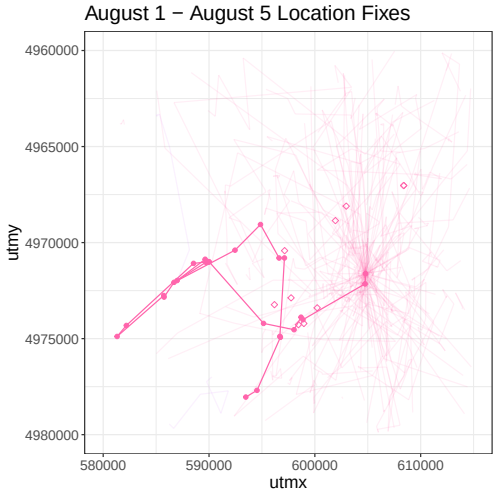
361 Given the motivation of this work is to understand longer-term grizzly bear land use
362 patterns, we use the holdout datat from August to evaluate prediction of future locations.
363 With this framework, the question of interest is not focused on where will the bear be located
364 on step $t+k$, but rather the focus is on all the locations that a bear could visit between time t
365 and $t+k$. Figure 7 shows recorded locations and estimated paths for the period of May - July
366 in 2015. The figure also contains an up-close view of predicted paths for 1 day (2-steps) and
367 5 days (10-steps) for bear GB820. From a practical purpose, these k-step ahead predictions
368 from this model can also be used to calibrate changes in space use. Predictions for the first
369 week have a mean absolute deviation of 9800 meters (and a median absolute deviation of



(a) The points represent recorded locations and lines correspond to the estimated paths.



(b) Predicted paths for August 1.



(c) Predicted paths for August 1 - August 5.

Figure 7: The top figure contains observed location fixes and estimated paths for all bears from May 1 - July 31 in 2015. The bottom two figures contain observed location and predicted paths for GB820 for 2-step ahead predictions on August 1 (left) and 10-step ahead predictions for August 1 - August 5 (right). The diamonds represent the actual 2 or 10-step ahead predictions.

6600 meters). Extending the predictions to the entire month, the mean absolute deviation is about 14,600 meters, while the median absolute deviation is just more than 10,000 meters.

6 Discussion

For many species the movement patterns of an individual depend upon the collective locations of other individuals. Existing research in collective movement models has largely focused on species that exhibit herding or schooling behavior. With this article, we present an ABM, along with fully Bayesian parameter estimation, designed for species in which both attractive or repulsive effects of conspecific animals can influence movement dynamics. In particular, our model is developed with a focus on grizzly bears, but it could be widely applicable to other species with common behaviors.

Our synthetic data analysis suggests that using a model designed to capture collective movement does not result in adverse predictions, when collective movement does not exist, however, it results in improved performance relative to a model that does not consider locations of other individuals. Using this model with twice-daily grizzly bear locations, there is evidence of different behavioral patterns in the presence of other bears.

The model also provides a mechanism for predicting future movement patterns. Using holdout data, median absolute prediction error for the next week is roughly 6600 meters and the median absolute deviation prediction error for the next month is just more than 10,000 meters. The model provides a framework that enables predictive paths for k-steps ahead, which can be used, for example, to estimate future occupied range, a critical information need for wildlife managers. To predict expansion of future occupied range, future research will focus on collective movement patterns and dispersal dynamics in combination with habitat and terrain attributes.

Acknowledgements

We thank the member agencies of the Interagency Grizzly Bear Study Team for data contributions: U.S. Geological Survey, National Park Service; U.S. Fish and Wildlife Service; U.S. Forest Service; Wyoming Game and Fish Department; Montana Fish, Wildlife and Parks; Idaho Department of Fish and Game; and the Eastern Shoshone and Northern Arapaho Tribal Fish and Game Department. We thank Joseph D. Clark for his review as part of the U.S. Geological Survey's Fundamental Science Practices. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

We also thank Stephen Walsh for the motivation to try Julia and the patience for answering our questions.

References

Andrieu, C., Doucet, A., and Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342.

Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A., Cisbani, E., Giardina, I., Orlandi, A., Parisi, G., Procaccini, A., Viale, M., et al. (2008). Empirical investigation of starling flocks: a benchmark study in collective animal behaviour. *Animal behaviour*, 76(1):201–215.

Beaumont, M. A. (2010). Approximate Bayesian computation in evolution and ecology. *Annual review of ecology, evolution, and systematics*, 41:379–406.

Beaumont, M. A., Zhang, W., and Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162(4):2025–2035.

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM review*, 59(1):65–98.

Bonnell, T. R., Henzi, S. P., and Barrett, L. (2016). Direction matching for sparse movement data sets: determining interaction rules in social groups. *Behavioral Ecology*.

Brown, D. G., Riolo, R., Robinson, D. T., North, M., and Rand, W. (2005). Spatial process and data models: Toward integration of agent-based models and GIS. *Journal of Geographical Systems*, 7(1):25–47.

Christ, A., Ver Hoef, J., and Zimmerman, D. L. (2008). An animal movement model incorporating home range and habitat selection. *Environmental and Ecological Statistics*, 15(1):27–38.

Couzin, I. D., Krause, J., Franks, N. R., and Levin, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025):513–516.

Couzin, I. D., Krause, J., James, R., Ruxton, G. D., and Franks, N. R. (2002). Collective memory and spatial sorting in animal groups. *Journal of theoretical biology*, 218(1):1–12.

Ebinger, M. R., Haroldson, M. A., van Manen, F. T., Costello, C. M., Bjornlie, D. D., Thompson, D. J., Gunther, K. A., Fortin, J. K., Teisberg, J. E., Pils, S. R., et al. (2016). Detecting grizzly bear use of ungulate carcasses using global positioning system telemetry and activity data. *Oecologia*, 181(3):695–708.

Farmer, J. D. and Foley, D. (2009). The economy needs agent-based modelling. *Nature*, 460(7256):685.

Fasiolo, M., Pya, N., Wood, S. N., et al. (2016). A comparison of inferential methods for highly nonlinear state space models in ecology and epidemiology. *Statistical Science*, 31(1):96–118.

Gilbert, N. (2008). *Agent-based models*. Sage.

Gilbert, N. and Terna, P. (2000). How to build and use agent-based models in social science. *Mind & Society*, 1(1):57–72.

441 Haydon, D. T., Morales, J. M., Yott, A., Jenkins, D. A., Rosatte, R., and Fryxell, J. M.
442 (2008). Socially informed random walks: incorporating group dynamics into models of
443 population spread and growth. *Proceedings of the Royal Society of London B: Biological*
444 *Sciences*, 275(1638):1101–1109.

445 Hooten, M. B., Johnson, D. S., McClintock, B. T., and Morales, J. M. (2017). *Animal*
446 *movement: statistical models for telemetry data*. CRC press.

447 Hooten, M. B. and Wikle, C. K. (2010). Statistical agent-based models for discrete spatio-
448 temporal systems. *Journal of the American Statistical Association*, 105(489):236–248.

449 Jonsen, I. D., Flemming, J. M., and Myers, R. A. (2005). Robust state–space modeling of
450 animal movement data. *Ecology*, 86(11):2874–2880.

451 Kahle, D. and Wickham, H. (2013). ggmap: Spatial visualization with ggplot2. *The R*
452 *Journal*, 5(1):144–161.

453 Langrock, R., Hopcraft, J. G. C., Blackwell, P. G., Goodall, V., King, R., Niu, M., Patterson,
454 T. A., Pedersen, M. W., Skarin, A., and Schick, R. S. (2014). Modelling group dynamic
455 animal movement. *Methods in Ecology and Evolution*, 5(2):190–199.

456 Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., and Morales, J. M.
457 (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models
458 and extensions. *Ecology*, 93(11):2336–2342.

459 McClintock, B. T., King, R., Thomas, L., Matthiopoulos, J., McConnell, B. J., and Morales,
460 J. M. (2012). A general discrete-time modeling framework for animal movement using
461 multistate random walks. *Ecological Monographs*, 82(3):335–349.

462 McDermott, P. L., Wikle, C. K., and Millspaugh, J. (2017). Hierarchical nonlinear spatio-
463 temporal agent-based models for collective animal movement. *Journal of Agricultural,*
464 *Biological and Environmental Statistics*, 22(3):294–312.

Morales, J. M., Haydon, D. T., Frair, J., Holsinger, K. E., and Fryxell, J. M. (2004). Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology*, 85(9):2436–2445.

Nuñez-Antonio, G., Ausín, M. C., and Wiper, M. P. (2015). Bayesian nonparametric models of circular variables based on Dirichlet process mixtures of normal distributions. *Journal of Agricultural, Biological, and Environmental Statistics*, 20(1):47–64.

Orderud, F. (2005). Comparison of kalman filter estimation approaches for state space models with nonlinear measurements. In *Proc. of Scandinavian Conference on Simulation and Modeling*, pages 1–8. Citeseer.

Peck, C. P., van Manen, F. T., Costello, C. M., Haroldson, M. A., Landenburger, L. A., Roberts, L. L., Bjornlie, D. D., and Mace, R. D. (2017). Potential paths for male-mediated gene flow to and from an isolated grizzly bear population. *Ecosphere*, 8(10):e01969.

Pedersen, T. L. (2019). *ggforce: Accelerating 'ggplot2'*. R package version 0.3.1.

Ritter, C. and Tanner, M. A. (1992). Facilitating the Gibbs sampler: the Gibbs stopper and the griddy-Gibbs sampler. *Journal of the American Statistical Association*, 87(419):861–868.

Scharf, H. (2019). *anipaths: Animation of Observed Trajectories Using Spline-Based Interpolation*. R package version 0.9.7.

Scharf, H. R., Hooten, M. B., Wilson, R. R., Durner, G. M., and Atwood, T. C. (2019). Accounting for phenology in the analysis of animal movement. *Biometrics*.

Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I., and Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical review letters*, 75(6):1226.

Wang, F. and Gelfand, A. E. (2013). Directional data analysis under the general projected normal distribution. *Statistical methodology*, 10(1):113–127.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Appendix

P-MCMC Algorithm

Using conjugate priors, the sampling largely consists of Gibbs steps along with a particle approach for the state parameters.

1. Propose new set of state parameters using marginal particle proposal. Specifically, a particle filter is used to propose $\mathcal{S}$, $\mathcal{X}$,δ, and *state* conditional on the remaining parameters in the model. The script notation refers to all of the variables of that type, across agents, years, and time points. Proposals are accepted with the typical Metropolis-Hastings ratio as detailed in Andrieu et al. (2010). This procedure updates one agent at a time, within a given year, but the procedure can be parallelized across years.

2. The probability parameters associated with the state transitions are fit using conjugate priors from a beta distribution.

3. The step size ($\mu_{u,j}$) and variance parameters ($\sigma^2_{u,j}$) in the lognormal distribution for each state can be sampled using a Gibbs sampler.

4. The variance associated with the measurement error (σ_ε) can be sampled from an inverse gamma distribution.

5. The final piece is taking samples from the projected normal, Nuñez-Antonio et al. (2015) outlines a Griddy-Gibbs approach (Ritter and Tanner, 1992) for this procedure. With this approach, the angular data (θ) is converted to Cartesian coordinates, where $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Integrating out r and using the x and y data enables Gibbs samples for the mean of the projected normal. Wang and Gelfand (2013) presents a more general approach for sampling from a projected normal distribution.

6.1 Model Specification and Priors for Data Analysis

Observation Equation

$$\underline{\mathbf{z_t}} \sim N\left(\mathbf{H_t}\underline{\mathbf{s_t}} + \underline{\epsilon_\mathbf{t}}, \sigma_\epsilon^2 I\right)$$

Evolution Equation

$$\begin{aligned}
\underline{\mathbf{s_{i,t}}} &= \underline{\mathbf{s_{i-1,t}}} + u_{i,j,t}\underline{\delta}_{\mathbf{i,j,t}} \\
\underline{\delta}_{\mathbf{i,j,t}} &= (\cos(\theta_{i,j,t} + \nu_{i,j,t}), \sin(\theta_{i,j,t} + \nu_{i,j,t})) \\
\theta_{i,j,t} &\sim PN(\underline{\mu_{\theta,j}}, I) \\
u_{i,j,t} &\sim LN(\mu_{u,j}, \sigma_{u,j}^2) \\
p_{12[c]} &= Pr[j_t = 1 \rightarrow j_{t+1} = 2 | d_t < thr] \\
p_{12[f]} &= Pr[j_t = 1 \rightarrow j_{t+1} = 2 | d_t > thr] \\
p_{22[c]} &= Pr[j_t = 2 \rightarrow j_{t+1} = 2 | d_t < thr] \\
p_{22[f]} &= Pr[j_t = 2 \rightarrow j_{t+1} = 2 | d_t > thr],
\end{aligned}$$

where $j = 1$ is the model state with shorter steps that tend to stay in the same area, $j = 2$ is the model state with larger steps that tend to maintain the existing heading, $\theta_{i,j,t}$ is the stochastic component of the heading, where $\nu_{i,j,t}$ aligns the heading to a collection of previous locations for state i or the most recent heading for state j. The switching component of the model is determined by whether the distance to the nearest animal, d_t, is less than a threshold, thr.

Prior Distributions

In general, the prior distributions are weakly informative and aligned with biological understanding.

$$\sigma_\epsilon^2 \sim IG(5, 50000)$$

The prior for σ_ϵ^2 has a mean that roughly corresponds to a standard deviation of 100 meters.

$$\mu_\theta \sim N(\underline{(1,0)}^T, .1I)$$

The prior for μ_θ is centered at $(\underline{(1,0)}^T)$, with relatively weak precision, has little impact on the posterior.

$$\mu_{u1} \sim N(\log(500), .5)$$
$$\mu_{u2} \sim N(\log(5000), .5)$$
$$\sigma_{u1}^2 \sim IG(5, 1.8)$$
$$\sigma_{u2}^2 \sim IG(5, 0.6)$$

The mean step size parameters are centered at 500 meters and 5000 meters (after accounting for lognormal parameterization), but have small enough precision that they can be largely informed by the data. Similarly, the standard deviation values are not particularly influential.

$$p_{12[c]} \sim Beta(1.9, .1)$$

$$p_{12[f]} \sim Beta(.1, 1.9)$$

$$p_{22[c]} \sim Beta(1.9, .1)$$

$$p_{22[f]} \sim Beta(1.9, .1)$$

The priors for the transition probabilities suggest inertia (95 percent probability of staying in same state), except for when another bear is in close proximity and a bear is in state 1, then the prior probability of switching is 95 percent. However, the prior probabilities correspond to 2 data points, so there is little impact on the posterior probability.