

**ALTERNATIVE LEARNING STRATEGIES FOR
SPATIO-TEMPORAL PROCESSES OF COMPLEX ANIMAL
BEHAVIOR**

A Dissertation presented to
the Faculty of the Graduate School
at the University of Missouri

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

by
TORYN L. J. SCHAFER
Christopher K. Wikle, Dissertation Supervisor

JULY 2020

The undersigned, appointed by the Dean of the Graduate School, have examined the dissertation entitled:

Alternative Learning Strategies for Spatio-temporal Processes of Complex Animal
Behavior

presented by Toryn L. J. Schafer,

a candidate for the degree of Doctor of Philosophy and hereby certify that, in their opinion,
it is worthy of acceptance.

Dr. Christopher K. Wikle

Dr. Mitch D. Weegman

Dr. Erin M. Schliep

Dr. Mevin B. Hooten

Dr. Athanasios C. Micheas

ACKNOWLEDGMENTS

First, I would like to thank my advisor (Chris Wikle), and my doctoral committee (Erin Schliep, Mitch Weegman, Mevin Hooten, Sakis Micheas) for their guidance, support, and patience. I am especially grateful to Chris for his constant encouragement and thoughtful insights.

Specifically, I would like to thank Stephanie Cunningham, Bart Ballard, Patrick McDermott, Henry Scharf, and Jay VonBank for providing datasets, feedback, and code used to complete this dissertation.

Lastly, I am forever grateful for the support and love from my family, friends, and pets.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES	vii
LIST OF FIGURES	ix
ABSTRACT	xiv
CHAPTER	
1 Statistical Learning Strategies for Spatio-temporal Ecological Data	1
1.1 Markov Models (MM) for Individual Animal Behavior	3
1.1.1 Ecological MMs	6
1.2 Echo State Networks (ESNs) for Ecological Forecasts	7
1.2.1 Leaky Integrator ESN	10
1.2.2 ESNs for Long-lead Ecological Forecasting	11
1.3 Reinforcement Learning (RL) for Collective Animals	12
1.3.1 Markov Decision Process (MDP)	13
1.3.2 Solving the MDP	14
1.3.3 Guppy MDP Example	16
1.3.4 Inverse Reinforcement Learning	22
1.3.5 Linearly-Solvable MDP (LMDP)	24
1.4 Overview of the Chapters	26
2 A Bayesian Markov Model with Pólya-Gamma Sampling for Estimating Individual Behavior Transition Probabilities from Accelerometer Classifications	28

2.1	Introduction	28
2.2	Materials and Methods	35
2.2.1	Dataset	35
2.2.2	Model	38
2.2.3	Model Fitting	41
2.3	Results	44
2.4	Discussion	51
3	Hybrid Statistical and Deep Neural Models for Ecological Prediction	55
3.1	Introduction	55
3.2	Methodology	58
3.2.1	D-EESN with GLM	59
3.2.2	BD-EESN with GLM	62
3.3	Mallard Settling Patterns	65
3.3.1	Background and Data	65
3.3.2	Forecasting Application Specifics	66
3.3.3	Results	68
3.4	Discussion	73
4	Bayesian Inverse Reinforcement Learning (IRL) for Collective Animal Movement	77
4.1	Introduction	77
4.2	IRL Methodology	80
4.2.1	LMDP	80
4.2.2	Inverse Reinforcement Learning (IRL)	82

4.3	Vicsek LMDP	84
4.4	Guppy Application	87
4.5	Discussion	93
4.5.1	LMDP and Animal Movement Modeling	94
5	General Conclusions	96
APPENDIX		
A	Supplement for Chapter 2	98
A.1	CropScape Habitat Categories	98
A.2	Acceleration Classification	100
A.3	Conditional Distributions	103
A.4	Model Validation	104
A.5	Habitat Multiple Comparison Matrices	106
A.6	Sensitivity of Estimates	109
B	Supplement for Chapter 3	112
B.1	Hyperparameters	112
B.2	MCMC Distributions	113
B.2.1	Basis Coefficients	114
B.2.2	Intercept Vector	115
B.2.3	Regression Coefficients and HS parameters	116
C	Supplement for Chapter 4	117
C.1	STAN Algorithms and Code	117
C.1.1	STAN Model Code	119

BIBLIOGRAPHY	121
VITA	138

LIST OF TABLES

Table	Page
1.1 Reward system for actions related to collective behavior by zones defined by neighbor distance. The immediate reward received for taking one of the collective behavior actions is based completely on the distance to perceived neighbor in the current state, d	19
1.2 Reward system for actions related to tank boundary by distance to nearest tank wall. The immediate reward received for taking one of the tank boundary actions is based completely on the distance to nearest boundary in the current state, b	19
2.1 Means and 95% credible intervals for odds ratios for habitat coefficients for the feeding to feeding transition arranged by increasing mean estimated in a Bayesian framework for six greater white-fronted geese in March 2018. A significant difference between habitats is indicated by non-overlapping credible intervals which can be visualized in Figure 2.4. The values correspond to exponentiating the estimates of the coefficients depicted in the from feeding to feeding panel of Figure 2.2	48

2.2	A selection of means, 95% credible intervals and proportion of samples with an estimate greater than 1 for the odds ratios for quantitative covariate coefficients estimated in a Bayesian framework for six greater white-fronted geese in March 2018. The quantitative variable has a significant effect on the transition probability if the credible interval does not overlap with 1 which corresponds to a credible interval in Figure 2.3 not overlapping 0.	50
3.1	The mean squared prediction error (MSPE), Pearson correlation, and spatial median of the continuous rank probability score (CRPS) for one-step-ahead forecast of mallard settling patterns in 2009 and 2014 with various architectures of the D-EESN and BD-EESN. Smaller indicates better prediction for MSPE and CRPS while larger correlations are better.	69
A.1	Habitat categories used in the Bayesian Markov model with the associated CropScape categories from the USDA. CropScape categories were matched to GPS data for six greater white-fronted geese during March 2018.	99
B.1	A table of hyperparameters for the D-EESN. For all layer specific hyperparameters, values were tuned for each layer unless otherwise noted.	113

LIST OF FIGURES

Figure	Page
1.1 A directed acyclic graph (DAG) of the deep ESN. The vertical direction of the graphic depicts feedforward layers while the horizontal direction depicts recurrent layers. The reservoirs are the large circles between the input layer, $\tilde{\mathbf{x}}_t$, and output layer, α_t . All connections to reservoirs are random while all connections to the output layer are estimated.	9
1.2 This graphic depicts the environment. The blue area is the neutral zone. The red area is the negative zone. The green area is the positive zone. All agents are initiated to start at a random point within the yellow square. . . .	17
1.3 The episode length is defined as the minimum of the number of steps before an agent passes the boundary or 200. The agents quickly learn to stay within the boundary given the frequency of episodes with 200 steps after about episode 70.	21
1.4 The behavior trajectories of the 10 agents in episode 70, episode 250, and under the final policy. The colored boxes delineate the different habitat zones as shown in Figure 1.2	22

2.1	Spring migration paths in North America of 6 individual greater white-fronted geese (<i>Anser albifrons frontalis</i>) from GPS equipped tracking devices. The highlighted section (white) are the locations for March 2018 used in the analysis of behavior transition probabilities.	36
2.2	95% credible intervals for habitat effects on the log-odds scale estimated in a Bayesian framework by behavior transition for six greater white-fronted geese in March 2018. There are significant pairwise differences between intervals that do not overlap.	45
2.3	95% credible intervals for diurnal (a) and weather (b) covariate effects on the log-odds scale estimated in a Bayesian framework by behavior transition for six greater white-fronted geese in March 2018. Significance refers to whether or not the credible interval overlaps zero.	46
2.4	Matrix of pairwise comparisons between habitat coefficients for the transition feeding to feeding for six greater white-fronted geese in March 2018. The values indicate the proportion of posterior MCMC iterations in which the habitat coefficient down the row was greater than the coefficients along the column; the upper triangular values and lower triangular values sum to 1. For example, in the first row, corn, the proportion of times the estimate for the corn habitat coefficient was greater than the open water habitat coefficient was 1. The rest of the transitions can be found in Appendix A.5 .	49

3.1	Spatial maps of the true mallard settling pattern for 2009 (a) and side-by-side comparisons of forecast medians, standard deviations, and the 97.5th percentiles of the forecast distribution for the best D-EESN (b) and BD-EESN (c) model based on the CRPS found in Table 3.1. Grayed circles indicate values that exceed the legend threshold.	70
3.2	Spatial maps of the true mallard settling pattern for 2014 (a) and side-by-side comparisons of forecast medians, standard deviations, and 97.5th percentiles of the forecast distribution for the best D-EESN (b) and BD-EESN (c) model based on the CRPS found in Table 3.1. Grayed circles indicate values that exceed the legend threshold.	71
3.3	Hidden units with a pseudo-inclusion probability greater than 0.1 for the top two BD-EESN models based on CRPS found in Table 3.1	73
4.1	Estimated costs-to-go for a Vicsek SPP model using LMDP IRL for Bayesian MCMC sampling and variational approximation under known passive dynamics. The models either used Gaussian basis functions (dashed lines) or independent state parameters (solid lines). The shaded regions correspond to the 95% C.I. The black curve is the true costs-to-go.	86
4.2	Variational mean costs-to-go for the guppy experiments for a gridded state space of target and local misalignment across three sets of features: full (identity matrix), radial basis functions, and bisquare basis functions. The passive dynamics are assumed to be discrete uniform and the mean estimated costs have been shifted to have a minimum of 0. The yellow indicates states with lower costs-to-go and therefore states to which the guppies choose to transition.	88

4.3	Variational mean costs-to-go (left panel) and standard deviations (right panel) for the guppy experiments for a gridded state space of target and local misalignment with passive dynamics assumed to be a normal random walk and bisquare basis functions. The mean estimated costs-to-go have been shifted to have a minimum of 0. The yellow indicates states with lower costs-to-go and therefore states to which the guppies choose to transition.	90
4.4	Trajectories of all 26 experiments of groups of 10 guppies in a tank. The target is located at the point marked "T".	91
4.5	Marginal costs-to-go of target and local misalignment for the guppy experiments for a gridded state space of target and local misalignment with passive dynamics assumed to be a normal random walk and bisquare basis functions. The mean estimated costs have been shifted to have a minimum of 0. The lower costs-to-go increase the probability of transitioning into that state.	92
A.1	Raw acceleration time series for (a) the 481 training observations in the X, Y, and Z axis for the four behavior categories flight (150), feeding (106), stationary (150), and walking (75) and (b) a sample of 1% of the classified observations. In (b), the behaviors correspond to the most likely category based on the random forest predicted probabilities.	101
A.2	Box plots for the top five variables used in the classification random forest for the ground-truthed behaviors (a) and the classified observations (b). In (b), the behaviors correspond to the most likely category based on the random forest predicted probabilities.	102

A.3	Comparisons of the imputed distribution of transition frequencies based on 200 imputation data sets and the posterior distribution of transition frequencies for all behavior transitions for six greater white-fronted geese in March 2018. The posterior covers the imputed distribution in all cases.	105
A.4	Matrix of pairwise comparisons between habitat coefficients for the all behavior transitions with estimated coefficients for six greater white-fronted geese in March 2018. The values indicate the proportion of posterior MCMC iterations in which the habitat coefficient down the row was greater than the coefficients along the column; the upper triangular values and lower triangular values sum to 1.	108
A.5	95% credible intervals for habitat effects on the log-odds estimated in a Bayesian framework with either the behavior class with the largest classification probability (Max Prob.), 100 imputation data sets, or 200 imputation data sets by behavior transition for six greater white-fronted geese in March 2018.	110
A.6	95% credible intervals for diurnal (a) and weather (b) covariate effects on the log-odds estimated in a Bayesian framework with either the most likely behavior class (Max Prob.), 100 imputation data sets, or 200 imputation data sets by behavior transition for six greater white-fronted geese in March 2018.	111

Alternative Learning Strategies for Spatio-temporal Processes of Complex Animal Behavior

Toryn L. J. Schafer

Christopher K. Wikle, Dissertation Supervisor

ABSTRACT

The estimation of spatio-temporal dynamics of animal behavior processes is complicated by nonlinear interactions. Alternative learning methods such as machine learning, deep learning, and reinforcement learning have proven successful for approximating nonlinear system mechanisms for prediction and classification. These alternative learning frameworks can be linked to statistical models in a hierarchical framework to improve ecological inference and prediction in the presence of uncertainty. This dissertation provides three methodological extensions of alternative learning with statistical uncertainty quantification for modeling animal behavior dynamics at different scales. First, an efficient Bayesian Markov model is developed to provide inference on white-fronted geese behavior from individual accelerometer and location data while accounting for classification uncertainty. Second, nonlinear basis function expansions produced by a spatio-temporal echo state network are used as features in a hierarchical generalized linear model for predicting spatial patterns of mallard duck settling pattern counts. Lastly, Bayesian inverse reinforcement learning is developed to estimate the behavioral state costs for collective animal groups.

Chapter 1

Statistical Learning Strategies for Spatio-temporal Ecological Data

With the prevalence of high volume and individual-specific ecological data sources, the quantitative analysis of ecological systems is challenging due to multiscale nonlinearities in space and time including inter- and intra-system interactions. In animal behavior modeling, inference for individual decision-making from new sources of individual tracking data is complicated due to habitat and social interactions. Remote sensing has also increased the quantity of data for monitoring large scale ecological systems, but efficient methods are needed for understanding the underlying mechanisms of spatial and temporal dependence. One key to overcoming these challenges is interdisciplinary modeling solutions that hybridize alternative learning strategies and statistical uncertainty quantification.

Machine learning with novel, large data sources can provide more interpretable derived data products and may be implemented by black box software not developed by the ecological modeler. Although machine learning methods typically provide information

about accuracy, there is not much guidance in the literature for incorporating the uncertainty associated with classification into subsequent analyses. Yet, the machine learning method provides probabilistic predictions, and therefore the classification uncertainty can be incorporated in further analysis using a statistical model. Specifically, the classification predictions can be used as a distribution for the true, unobserved values within a multiple imputation framework. Incorporating the uncertainty in predictions instead of using the most likely classification is necessary to avoid biases in further estimation (Scharf et al., 2017). The probabilistic predictions can naturally be incorporated into a prior on the data in a Bayesian framework, but the literature is still lacking in examples. We present an efficient Bayesian multiple imputation framework applied to accelerometer data classifications in Chapter 2.

In dynamic ecological system modeling, the scales of temporal dependence are typically constrained to the current or most recent past, but this assumption is not generally appropriate. While the stochastic antecedent or ecological memory models of Ogle et al. (2015) can learn the length of temporal dependence in a system, the dependence is still primarily linear. Deep learning with echo state networks (ESNs), which are sparse and random recurrent neural networks (RNNs), can learn the multiscale, nonlinear features of importance in a dynamic system (Jaeger, 2001). In particular, an ensemble of randomly initialized ESNs acts as a large nonlinear, multiscale basis expansion of the inputs. The only parameters in the ESN are the output weights, which can be estimated with uncertainty in a frequentist or Bayesian regression framework with shrinkage. As shown in Chapter 3, this idea can be extended if the output corresponds to a generalized linear model, which broadens the applicability for ecological data sources.

In animal behavior ecology, collective animal movement modeling in heterogeneous

environments has been challenging. A holistic framework for inference of habitat use in social animal groups is lacking. Many collective movement models are "forward models" in that they assume the local rules governing behavior *a priori* and the learned behavior is not transferable to various situations. Inverse modeling allows the local rules and objectives to be learned from observations of sequential decision-making providing insight into the decision-making process from tracked individuals (Yamaguchi et al., 2018). Inverse reinforcement learning (IRL) is a method for inverse learning in a Markov decision process framework. Chapter 4 utilizes IRL in a Bayesian framework to infer collective animal behavior.

The chapters of this dissertation present methodology that links ecology, alternative learning approaches, and statistical uncertainty. In the remaining sections of this Chapter, the key features of the analyses of animal behavior at various scales found in the subsequent chapters are previewed. Additionally, Section 1.3 includes an illustrative example of reinforcement learning for collective animal behavior, given the novelty of this approach in statistics.

1.1 Markov Models (MM) for Individual Animal Behavior

A Markov model (MM) describes a dynamic process with the Markov property (i.e., the future state of the system is independent of past states given the current state). A homogeneous, discrete time MM with J discrete states is determined by its transition probability

matrix, \mathbf{P} :

$$\mathbf{P} \equiv \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1J} \\ p_{21} & p_{22} & \cdots & p_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ p_{J1} & p_{J2} & \cdots & p_{JJ} \end{bmatrix}, \quad (1.1)$$

where $p_{ij} \equiv P(s_t = j | s_{t-1} = i)$, and the row probabilities sum to one, $\sum_{j=1}^J p_{ij} = 1$. The non-zero entries of \mathbf{P} determine the one step reachable states (e.g., if $p_{12} > 0$ then state 2 is reachable from state 1). When the transition probabilities are not stationary, the MM is non-homogeneous with a transition probability matrix indexed by time, \mathbf{P}_t .

Generally, the transition probability matrix is estimated by multinomial logistic regression. Consider observations $\mathcal{S} = \{s_0, \dots, s_T : s_t \in \{1, \dots, J\}\}$ and let $\mathbf{y}_{it} = (y_{i1t}, \dots, y_{iJt})'$ be the vector of length J with elements $y_{ijt} = 1(s_t = j | s_{t-1} = i)$. Then, \mathbf{y}_{it} is an observation of a multinomial trial with probabilities $\mathbf{p}_{it} = (p_{i1t}, \dots, p_{iJt})'$. The likelihood of \mathcal{S} is:

$$P(\mathcal{S}) = p_0(s_0) \prod_{t=1}^T \prod_i^J \prod_j^J p_{ijt}^{y_{ijt}}, \quad (1.2)$$

for $i = 1, \dots, J; j = 1, \dots, J$, where a n_b -dimensional covariate vector, \mathbf{x}_t , is introduced on the transitions probabilities with the multinomial logistic link function:

$$\log(p_{ijt}/p_{iC_it}) = \mathbf{x}_t' \boldsymbol{\beta}_{ij}, \quad (1.3)$$

where C_i is a reference category, $\boldsymbol{\beta}_{ij}$ is a n_b -dimensional vector of coefficients, and $\boldsymbol{\beta}_{iC_i} = 0$ for all $i = 1, 2, \dots, J$ to ensure identifiability due to the sum to one constraint on the probabilities. In Chapter 2, the reference category is set to the last category, J , for all $i = 1, 2, \dots, J$. However, in transition probability matrix estimation it is also common to

consider effects of covariates on transitioning out of the current state, and in that framework, $C_i = i$, for all for all $i = 1, 2, \dots, J$. Lastly, MMs generally do not estimate the initial state distribution of the Markov chain and model the likelihood of the transitions only (Rugg and Buech, 1990; Sung et al., 2007).

In Bayesian multinomial logistic models, the posterior of the logistic regression coefficients can be sampled using Metropolis-Hastings. However, tuning the Metropolis-Hastings sampler can be difficult for moderate numbers of covariates or low observed frequencies of categories (Albert and Chib, 1993; Holmes and Held, 2006; Frühwirth-Schnatter and Frühwirth, 2010; Polson et al., 2013). Albert and Chib (1993) introduced a data augmentation approach that used truncated normal random variables for category membership using probit link regression. Hooten et al. (2010) used the Albert and Chib sampling for a multinomial movement model. Holmes and Held (2006) expressed a multinomial logistic model as a product of binary models and used truncated scale mixture normals to define category membership. Alternatively, Polson et al. (2013) developed a class of Pólya-Gamma distributions for automatic sampling of Bayesian logistic regression models that is faster and provides exact inference when compared to the scale mixture latent variables. The introduction of Pólya-Gamma latent variables induces simple Gibbs updating steps for the logistic regression coefficients when using a normal (Gaussian) prior. Additionally, Polson et al. (2013) showed the Pólya-Gamma scheme is significantly faster than other data augmentation methods and more efficient than Metropolis Hastings for mixed effects logistic models. More recently, Bradley et al. (2019) introduced an approach that gives conjugate updates for multinomial models with a new class of conditional multivariate logit-beta priors.

1.1.1 Ecological MMs

MMs are frequently used in population ecology with transitions defined by a Leslie matrix. Leslie matrices are mathematical models for discrete age classes with transition probabilities defined by age specific birth and survival rates (Leslie, 1945) and the elements can be estimated in a statistical MM framework (Buckland et al., 2007); similar MMs for transitions between stage classes (e.g., juveniles and adults) are used when age specific demographic rates are not available (Cochran and Ellner, 1992).

In animal behavior ecology, inference about behavioral time allocation is important for inference about energy expenditure, which can scale up to affect larger scale processes such as reproduction and survival (Harrison et al., 2011). Similarly, Hooten et al. (2019) showed energetic recharge dynamics of individuals is related to differences in movement modes. Until the advent of individual remote tracking technology in wildlife ecology, data about proportion of time spent in behavioral states had to be collected by direct observation of the individuals often limiting the habitats and times of day data of monitoring. Often, the proportions were then modeled in a logistic regression framework ignoring the temporal state transitions. Now, data from accelerometers on individual tracking devices provide novel insight into individual decision making by continually tracking individuals. One method for obtaining behavioral inference from acceleration data is by using a hidden Markov model (HMM) where the acceleration data are treated as observations from a distribution whose parameters are determined by the hidden true behavioral state and transitions between the hidden states are modeled with a MM (Zucchini et al., 2016; Leos-Barajas et al., 2017; McClintock et al., 2017).

In systems where the visual observations can be temporally matched to acceleration data, a labeled training data set can be made. In these situations, a strong relationship between

acceleration profiles and behavior has been established by the use of machine learning classification models and tools (Resheff et al., 2014; Fehlmann et al., 2017; Chakravarty et al., 2019). However, there is not an established next step for obtaining behavioral inference from the classified behavior data that considers the temporal dependence in the behavior categories and the uncertainty in classification. In Chapter 2, we present a framework to address both of these points.

The methodology of Chapter 2 incorporates the uncertainty in classification by using the probabilistic classifications as a model for the missing true categories in a multiple imputation framework. A MM captures the temporal dependence between the behavior categories and links the acceleration data to habitat data as measured by a Global Positioning System (GPS). The estimation is done in an efficient Bayesian framework with Pólya-Gamma latent variables and offline imputation of plausible data sets. It is the first comprehensive example of Pólya-Gamma latent variables and multiple imputation within a Markov model.

1.2 Echo State Networks (ESNs) for Ecological Forecasts

MM analysis of ecological time series may not be adequate when the scales of temporal dependence are not known. Recurrent neural networks (RNNs) provide a model architecture to learn multiscale temporal features (Hopfield, 1982). Big ecological data generally consists of many individuals and/or spatial locations and considerably fewer time points limiting the ability to train an RNN. However, the echo state network (ESN) is a sparse RNN that can learn multiscale temporal dependence with significantly fewer parameters (Lukoševičius and Jaeger, 2009) and therefore are more appropriate for smaller datasets.

In neural modeling, the RNN is a neural network with temporal evolution of hidden layers

and therefore, is a dynamical system (Lukoševičius and Jaeger, 2009). Indeed, an RNN is a universal approximator of dynamical systems under certain, mild assumptions (Funahashi and Nakamura, 1993). However, the optimization of a traditional RNN is challenging primarily due to the so called vanishing and exploding gradient problem (Bengio et al., 1994). The vanishing gradient problem is caused by the multiplication of many small gradients through time during backpropagation and the exploding gradient is caused by propagation of large gradients.

The echo state network (ESN) is the reservoir computing counterpart to the RNN. Reservoir computing is characterized by the random generation and fixing of internal network weights prior to estimation of output weights (Fig. 1.1; Lukoševičius and Jaeger, 2009). Thus, the ESN is a competitive forecast model with only a fraction of the number of weights (parameters) requiring estimation compared to the RNN and therefore, offers a competitive advantage in computation time.

Due to the separation of the internal structure from the output structure, the ESN reservoir is interpretable as a random, nonlinear, temporal basis expansion of the inputs (Lukoševičius and Jaeger, 2009). Therefore, the reservoir features (and weights) can be used as predictors for other outputs. For example, the reservoirs in Chapter 3 are nonlinear transformations of Pacific Ocean Sea Surface Temperature (SST) for forecasting mallard duck settling patterns. However, SST has an indirect relationship with many other ecological processes and therefore the SST features can be used as predictors in another ESN model.

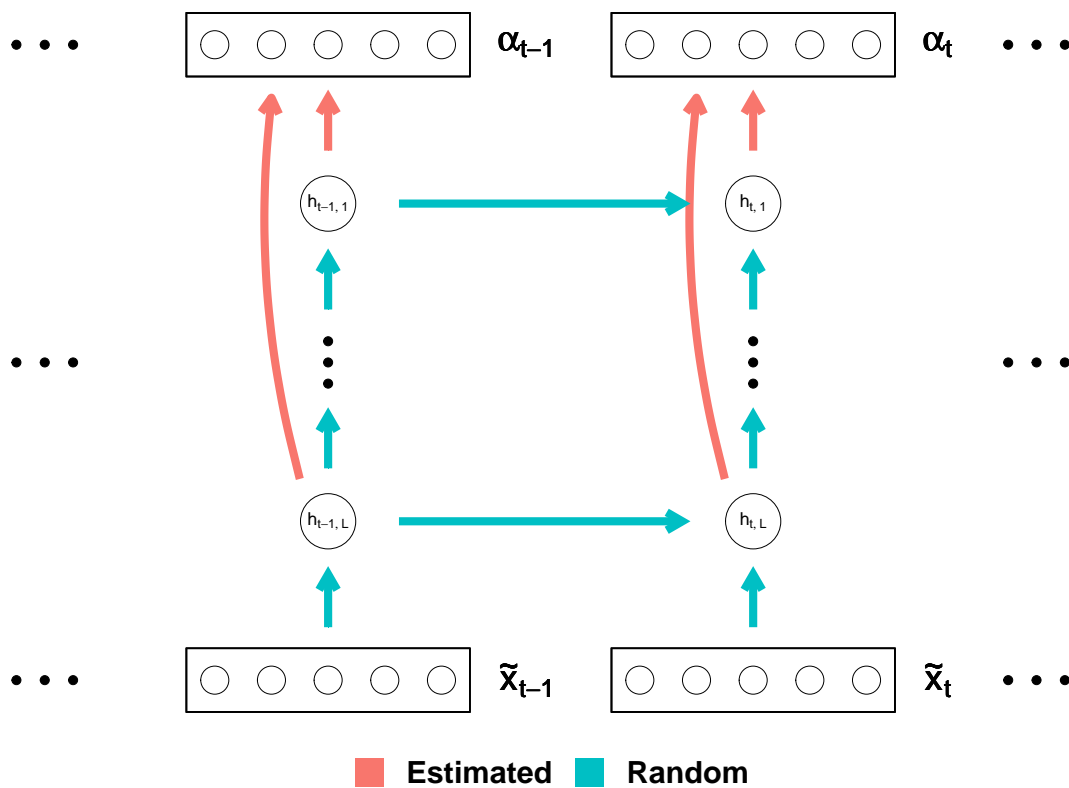


Fig. 1.1. A directed acyclic graph (DAG) of the deep ESN. The vertical direction of the graphic depicts feedforward layers while the horizontal direction depicts recurrent layers. The reservoirs are the large circles between the input layer, $\tilde{\mathbf{x}}_t$, and output layer, α_t . All connections to reservoirs are random while all connections to the output layer are estimated.

1.2.1 Leaky Integrator ESN

A leaky integrator ESN is defined by the following:

$$\mathbf{Output\ Stage:} \quad \alpha_t = \mathbf{V}\mathbf{h}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathbf{N}(0, \mathbf{I}) \quad (1.4)$$

$$\mathbf{Shrinkage:} \quad \min_{\mathbf{V}} \sum_t (\alpha_t - \mathbf{V}\mathbf{h}_t)^2 + \lambda \|\mathbf{V}\|_2^2, \quad (1.5)$$

$$\mathbf{Leaky\ Units:} \quad \mathbf{h}_t = (1 - \zeta)\mathbf{h}_{t-1} + \zeta\mathbf{h}_t \quad (1.6)$$

$$\mathbf{Input\ Stage:} \quad \mathbf{h}_t = f\left(\frac{\nu}{|\lambda_W|} \mathbf{W}\mathbf{h}_{t-1} + \mathbf{U}\tilde{\mathbf{x}}_t\right), \quad (1.7)$$

where the output weights, \mathbf{V} , linking the outputs, α_t , and the hidden units, \mathbf{h}_t , are estimated by ridge regression with penalty, λ , while the recurrent weights, \mathbf{W} , and input weights, \mathbf{U} , are random and sparse as determined by a hierarchical discrete mixture distribution:

$$\begin{aligned} w_{k,q} &= \gamma_{k,q}^w \text{Unif}(-a_w, a_w) + (1 - \gamma_{k,q}^w) \delta_0, & \gamma_{k,q}^w &\sim \text{Bern}(\pi_w), \\ u_{k,r} &= \gamma_{k,r}^u \text{Unif}(-a_u, a_u) + (1 - \gamma_{k,r}^u) \delta_0, & \gamma_{k,r}^u &\sim \text{Bern}(\pi_u), \end{aligned} \quad (1.8)$$

where δ_0 is a Dirac measure at 0 (Jaeger, 2001).

The scaling of the recurrent weight matrices, \mathbf{W} , by the largest eigenvalue, λ_W , ensures that the network is stable and gradually "forgets" over time when the activation function, $f()$, is tanh (i.e., the echo state property; Jaeger, 2001). Furthermore, the hyperparameter $\nu \in (0, 1]$ controls the rate of forgetting where values closer to 1 forget more slowly. The leaky integrator feature controls reactivity of the output to the input with smaller values of $\zeta \in [0, 1]$ reacting more slowly to the inputs providing the opportunity for long memory (Lukoševičius and Jaeger, 2009).

The single reservoir ESN (1.4) cannot learn instantaneous relationships well (i.e., the

relationship between now and the current inputs.) However, deepening and widening the ESN improves the learning of important multiscale features, which lends to its success in forecasting problems (Ma et al., 2017; Carmichael et al., 2019; McDermott and Wikle, 2019b). The widening of the ESN provides an ensemble of feedforward reservoirs.

1.2.2 ESNs for Long-lead Ecological Forecasting

Although there is an increase in ecological data sources, for annual ecological processes such as annual migration patterns considered in Chapter 3, the number of temporal observations remains small. Traditional deep RNN models require more data than is typically available in these settings. However, ESNs can still provide prediction and estimation of nonlinear linkages between output annual processes and input features in these small data situations due to the random weights and stochastic basis functions.

Despite its utility in small data situations, the generic ESN is lacking the key uncertainty component required for ecological forecasting (Clark et al., 2001; Dietze, 2017a). To account for uncertainty in the initialization of the random weights matrices, \mathbf{W} and \mathbf{U} , McDermott and Wikle (2017) introduced the ensemble ESN, which randomly initializes a large number of reservoirs and estimates the reservoir specific output weights to create an ensemble of forecasts. The estimation of output weights can be done in a frequentist or Bayesian framework with shrinkage.

An extension to the ESN for increased applicability to ecological systems, explored here, is the use of a generalized linear model for the output layer. Generalized linear models (GLM) are important tools in ecology as many data types are not continuous, real-values such as population counts, occupancy, and mark recapture histories. To the best of our knowledge, a generalized linear model has not been used in the ESN model for machine

learning application regardless of the architecture. Therefore, the ESN models with GLM output in Chapter 3 are the first examples.

1.3 Reinforcement Learning (RL) for Collective Animals

This Section includes more relevant background than the previous two sections because reinforcement learning is a relatively new topic for statisticians.

Coordinated collective movement processes are typically based on individual interactions and implemented with agent-based models (Vicsek et al., 1995; Couzin et al., 2002; McDermott et al., 2017). The simple individual-based rules employed in agent-based models lead to complex group dynamics. Such rules attempt to reflect the internal processes of an agent, but generally imply automatic behavior and do not include memory (Ried et al., 2019). Specifically, the lack of autonomous learning by agents makes it difficult to incorporate interactions with habitat as every situation requires *a priori* assumptions about the agent's behavior in those specific environments.

Reinforcement learning (RL) is goal-oriented learning from the continuous interaction between an agent and its environment. The RL framework extends the definition of an agent from typical agent-based models to include features of perception and memory (Ried et al., 2019). In RL, perception by an agent is encoded into the agent's state. Memory is incorporated by allowing the parameters controlling behavior to be learned from an agent's experience.

Lastly, examples of agent-based movement models in heterogeneous habitats exist, but typically assume independent individual movement (Hooten et al., 2010; Hooten and Wikle, 2010). The incorporation of the interaction between habitat and agents in a traditional

collective movement model is not straightforward, but important for realistic inference (Strandburg-Peshkin et al., 2017). However, complex group dynamics in heterogeneous habitats can easily be incorporated in the RL framework.

1.3.1 Markov Decision Process (MDP)

The mathematical definition of reinforcement learning is based on the Markov decision process (MDP; Sutton and Barto, 1998), which is an extension of the Markov model in Section 1.1 to include a set of actions in the transition probability function. An MDP is defined as the tuple (S, A, P, γ, R) where:

- S : set of **states**
- $A = \{a_1, \dots, a_k\}$: set of **actions**
- $P: S \times A \times S \rightarrow [0, 1]$ is a **transition probability function**
- $\gamma \in [0, 1)$ is a **discount factor**
- $R: S \rightarrow R$ is a **reward function**

For an MDP, a policy is a map for choosing an action in a given state, $\pi : S \rightarrow A$. The value function of a policy, π , for a state s , $V^\pi(s)$, is defined as

$$V^\pi(s) = E_\pi\left[\sum_{i=0}^{\infty} \gamma^i R(s_i) | s_0 = s\right], \quad (1.9)$$

where $E_\pi[\cdot]$ is the expectation function with respect to the policy π . The values can be interpreted as the long term, discounted reward of following policy π from state s . Similarly,

we can define a function based on the state-action pairs, $Q^\pi(s, a)$, as

$$Q^\pi(s, a) = E_\pi\left[\sum_{i=0}^{\infty} \gamma^i R(s_i) | s_0 = s, a_0 = a\right], \quad (1.10)$$

which can be interpreted as the long term, discounted reward of following policy π after taking action a in state s .

The goal of reinforcement learning is to learn the policy, π , for which the value function is maximized. A policy $\pi = \pi^*$ is said to be optimal if the value function, and hence the Q function, is maximized $\forall s \in S$, i.e.

$$\pi^*(s) = \mathit{arg} \max_{a \in A} Q^\pi(s, a). \quad (1.11)$$

The above expression is referred to as Bellman optimality. The resulting expression for the Q function under Bellman optimality is

$$Q^{\pi^*}(s, a) = R(s) + \gamma \max_{a' \in A} Q^{\pi^*}(s', a'), \quad (1.12)$$

where s' is the next state. The above is referred to as the Bellman optimality equation (Sutton and Barto, 1998).

1.3.2 Solving the MDP

In model-free methods, the reinforcement learning problem is solved by iteratively estimating the Q function from an agent's interactions with the environment. Let s_t and a_t denote the state-action pair for an agent at iteration t . In problems with relatively small state-action spaces, the most common algorithm is tabular Q-learning. Q-learning updates the current

estimate of the Q function, $Q_t(\cdot, \cdot)$, at iteration t , of a state-action pair by the following

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \alpha [R(s_{t+1}) + \gamma \max_{a' \in A} Q_t(s_{t+1}, a') - Q_t(s_t, a_t)], \quad (1.13)$$

where α is a learning rate (Sutton and Barto, 1998).

In problems with larger state-action spaces, the updating of the Q function in 1.13 is typically replaced with a parameterized model, which learns the relationship between state-action pairs and the Q function. The prediction of the model is then used to choose the next action. To illustrate using the notation of Van Hasselt et al. (2016), let the Q function be a function of the state-action pair and a set of unknown parameters, θ , then

$$Y_t^Q \equiv R(s_{t+1}) + \gamma \max_{a' \in A} Q(s_{t+1}, a'; \theta_t) \quad (1.14)$$

is the target value learned by the model where θ_t is the estimate of the parameters at iteration t (Van Hasselt et al., 2016).

A deep Q-network (DQN) algorithm uses a deep neural network model to estimate the relationship between state-action pairs (input) and the Q function (Equation 1.14) with two novel features: Experience replay and the use of a target network (Mnih et al., 2015). Experience replay refers to storing observed transitions, (s_t, a_t, r_t, s_{t+1}) , and randomly sampling the observed transitions for updating the online network. In 1.14, the target, Y_t^Q , is calculated using the online values of parameters, θ_t . The target of the DQN is calculated from a copy of the online network called the target network with parameters θ^- where the parameters of the target network are copied from the online network only every τ steps:

$$Y_t^{DQN} \equiv R(s_{t+1}) + \gamma \max_{a' \in A} Q(s_{t+1}, a'; \theta^-) \quad (1.15)$$

The DQN uses the same estimated model for action choice and policy evaluation, which makes it more prone to overfitting and overoptimistically estimating the value. Van Hasselt et al. (2016) proposed the double DQN algorithm to improve the accuracy of the value estimation. The target of the double DQN,

$$Y_t^{doubleDQN} \equiv R(s_{t+1}) + \gamma Q(s_{t+1}, \operatorname{argmax}_{a' \in A} Q(s_{t+1}, a'; \theta_t); \theta'_t) \quad (1.16)$$

uses online network with parameters θ_t to select the action and a second network with parameters θ'_t to evaluate the value.

Note, the above mentioned methods, Q-learning and deep Q-learning are considered to be model-free, because they do not consider models of the environment (Sutton and Barto, 1998). The algorithms do not explicitly reference the environment model, P , it only learns from samples of trial and error.

1.3.3 Guppy MDP Example

The following is a simulation of guppies in a tank to illustrate the MDP and RL concepts. The simulation environment was comprised of three main components: a boundary, positive and negative habitat zones, and N agents (Figure 1.2). The agents perceived other agents, the boundary, the current habitat zone, and its coordinate position. Therefore the state, $\mathbf{s} = (s_1, s_2, s_3, s_4, s_5)'$, for each agent was a five-dimensional vector where $s_1 = d \in \mathbb{R}^+$ is the distance to a neighbor perceived with probability $d^{-\beta}$ (Morihiro et al., 2006), $s_2 = b \in \mathbb{R}^+$ is the distance to the nearest boundary, $s_3 = c \in \{0, 1, 2, 3\}$ where the indices correspond

to outside the boundary, neutral zone, negative zone, and positive zone, respectively, and $(s_4, s_5) = (x, y) \in \mathbb{R}^2$ is the position of the agent.

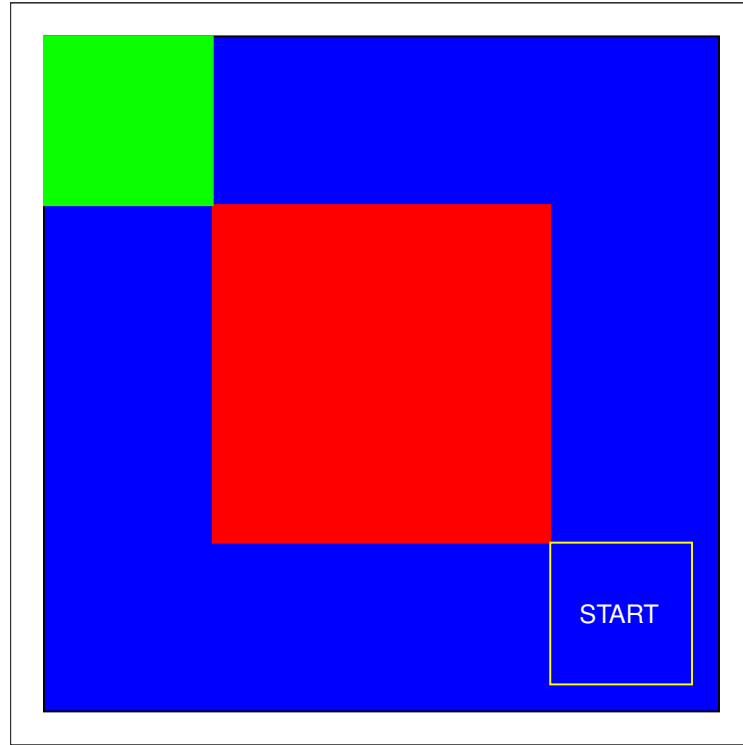


Fig. 1.2. This graphic depicts the environment. The blue area is the neutral zone. The red area is the negative zone. The green area is the positive zone. All agents are initiated to start at a random point within the yellow square.

The actions space consisted of eight unit directional vectors where the first four correspond to collective motion and the last four correspond to staying within the boundary. The assumption is that the agent is moving one body length at each step represented by the unit directional vector (Morihiro et al., 2006). Specifically, the eight actions are attraction to neighbor (a_1), positive parallel alignment with neighbor (a_2), negative parallel alignment with neighbor (a_3), repulsion to neighbor (a_4), attraction to nearest boundary (a_5), parallel alignment to nearest boundary toward positive habitat zone (a_6), parallel alignment to near-

est boundary away from positive habitat zone (a_7), and repulsion to nearest boundary (a_8). As a reminder, the neighbor is determined with probability proportional to $d_{ij}^{-\beta}$ where d_{ij} is the euclidean distance between agent i and j for all $i \neq j$.

The boundary alignment actions were selected such that the agent could choose the same action, a_6 , to go toward the positive zone regardless of which side of the environment it is located. This makes learning easier than using just clockwise and counterclockwise directions for boundary alignment. Finally, direction between the current and next position of the agent was then calculated as a weighted average of the agent's current direction and the chosen action unit directional vector given inertia:

$$\mathbf{m}_{it} = \frac{(1 - \kappa)\mathbf{m}_{i,(t-1)} + \kappa\mathbf{m}_a}{|(1 - \kappa)\mathbf{m}_{i,(t-1)} + \kappa\mathbf{m}_a|} \quad (1.17)$$

where \mathbf{m}_{it} is the direction vector for agent i at time t , \mathbf{m}_a is the directional vector for the action a , and κ is a parameter that controls inertia (Morihiro et al., 2006).

1.3.3.1 Reward System

The immediate reward received by the agent after taking an action is the mean of three values. The three values are associated with the collective behavior, tank boundary, and habitat zone. The magnitude and direction of the rewards are the biggest influence on the final learned policy. The reward system for the collective behavior is a zonal type system similar to that used by Couzin et al. (2002) with a zone of repulsion, alignment, and attraction based (Table 1.1). Outside of these three zones, an agent cannot perceive a neighbor. Again, since the actions are normalized, all the distances used in the reward system are interpreted as distances in body lengths. The values used to delineate the different zones are the same

Table 1.1. Reward system for actions related to collective behavior by zones defined by neighbor distance. The immediate reward received for taking one of the collective behavior actions is based completely on the distance to perceived neighbor in the current state, d .

Zone	Neighbor Distance, d	Action			
		a_1	a_2	a_3	a_4
Repulsion	$d < 4$	-1	-1	-1	+1
Alignment	$4 < d < 20$	-1	+1	-1	-1
Attraction	$20 < d < 50$	+1	-1	-1	-1
No Perception	$d > 50$	0	0	0	0

Table 1.2. Reward system for actions related to tank boundary by distance to nearest tank wall. The immediate reward received for taking one of the tank boundary actions is based completely on the distance to nearest boundary in the current state, b .

Boundary Distance, b	Action			
	a_5	a_6	a_7	a_8
$b < 4$	-1	+1	+1	+1
$4 < b$	0	0	0	0

as used by Morihiro et al. (2006).

The agent is rewarded for staying within the tank boundary. The agent can only perceive or not perceive the boundary based on the distance to the nearest edge (Table 1.2). If the agent perceives the tank boundary, it is penalized for moving toward the boundary (action a_5) and rewarded for taking actions a_6 , a_7 , or a_8 . This reward system allows agents to learn to travel near the boundary, but not cross over it.

Finally, the rewards for the different habitat zones are based on the zone observed after the agent executes the chosen action. The rewards for the different zones are -10 , 0 , -1 , and $+5$ for outside the boundary, neutral zone, negative zone, and positive zone, respectively.

The group size was $N = 10$. An episode was defined to be the minimum of the number

of steps observed before any agent moved outside the boundary and 200 steps. The MDP discount parameter, γ , was set to 0.8.

The state and action sequences were evaluated asynchronously for each agent. For calculating the nearest neighbor distance in the state vector, the perception parameter β was set to 0.5. To encourage exploration of the action space early in learning, we used an ϵ -greedy action selection with decaying exploration parameter ϵ . Under an ϵ -greedy action policy, the agent chooses a random action with probability ϵ or the best action based on the current Q function with probability $1 - \epsilon$. Because the agents were defined identically, each state-action sequence for the 10 agents was treated as an independent replicate from the environment and stored in the replay memory.

The deep neural network model used for the double DQN was a feed forward model with 3 hidden layers of size 64, 32, and 16 nodes, respectively. The parameters of the online neural network were updated each episode with a minibatch of 128 samples randomly selected from the replay memory. The parameters of the evaluation model in the double DQN were copied from the action selection model every 20 episodes and the simulation was run for 500 episodes. It is worth noting that the differences in using a DQN and double DQN were negligible for the proposed agent environment. This suggests learning the optimal policy in the scenario was not being severely overfit by a DQN.

The agents quickly learn to stay within the boundary for the set maximum of 200 steps per episode after about 70 episodes and consistently have episode lengths of 200 after about 280 episodes 1.3.

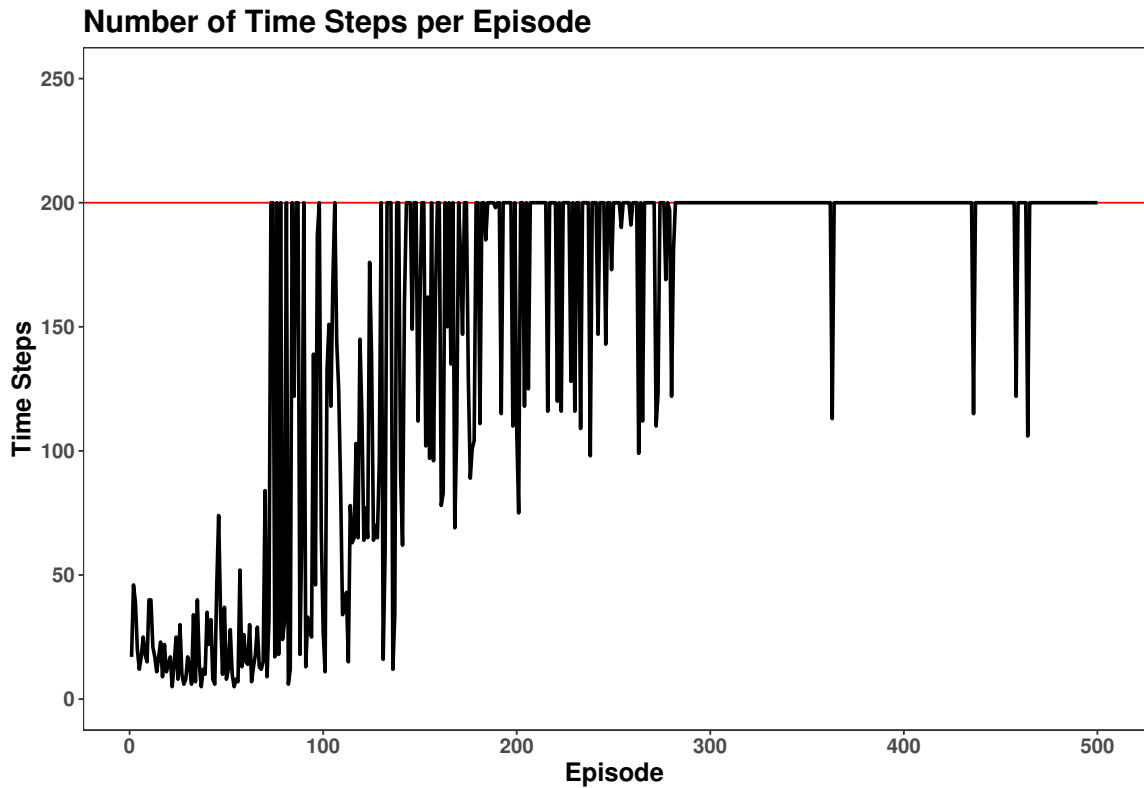


Fig. 1.3. The episode length is defined as the minimum of the number of steps before an agent passes the boundary or 200. The agents quickly learn to stay within the boundary given the frequency of episodes with 200 steps after about episode 70.

The optimal policy learned by the agents encouraged collective movement as one group toward the positive habitat zone. Once the agents were in the positive habitat zone, the collective movement was secondary to staying in the positive habitat zone and therefore smaller groups were able to form.

Figure 1.4 shows the progression of learning by the agents. At episode 70, the movements appear to mostly be random with some small groups forming. At episode 250, the agents have learned to stay in the boundary for the period of 200 steps, but do not end up in the positive habitat zone. Under the final policy, the agents travel as a group to the positive habitat zone and smaller groups form as they trade off the reward for

the positive zone and collective behavior. We provide animations of the sampled behavior for episode 70, episode 250, and the final policy at the following github page: <https://github.com/schafert/GuppyReinforcementLearning>.



Fig. 1.4. The behavior trajectories of the 10 agents in episode 70, episode 250, and under the final policy. The colored boxes delineate the different habitat zones as shown in Figure 1.2

1.3.4 Inverse Reinforcement Learning

Although the guppies learned complex, collective behavior, defining the reward system subjectively to mimic natural systems is challenging. Alternatively, the reward system that governs local behavior can be learned from observations of behavior trajectories with inverse methods.

In systems where observations of behavior trajectories can be observed, inverse rein-

forcement learning (IRL) methods aim to learn the state rewards or values that govern the agents' decisions. Ng and Russell (2000) introduced the first IRL algorithms including dynamic programming, which solves a system of equations based on the transition probabilities and a grid search method from trajectory samples. Many more methods have since been used and developed to tackle problems of meaningful size. A popular class of IRL methods is entropy optimization with a potential function based on feature matching along the observed trajectory (Ziebart et al., 2008; Arora and Doshi, 2018). Bayesian IRL methods assume a trajectory likelihood with a state-action value, Q , potential function resulting in a multinomial likelihood. Various Bayesian IRL implementations include using *maximum a posteriori* estimation with gradient optimization (Choi and Kim, 2011, 2014), Gaussian Process IRL (Jin et al., 2017), and full MCMC sampling (Šošić et al., 2017b). Bayesian nonparametric models have been used for reward feature construction (Choi and Kim, 2013) and learning the representation of state information used by the agent (Šošić et al., 2017b).

The likelihood for the Bayesian IRL methods assuming a Q potential function based on observed state-action pairs, $\mathcal{D} = \{(s_0, a_0), (s_1, a_1), \dots, (s_n, a_n)\}$, is a MM with transition probabilities proportional to the Q :

$$P(\mathcal{D}) \propto \prod_n \prod_t \frac{\exp(\beta Q^*(s_t^n, a_t^n | R))}{\sum_{a \in A} \exp(\beta Q^*(s_t^n, a | R))}, \quad (1.18)$$

where the parameter β reflects the degree to which agents are behaving optimally (i.e., $\beta = 1$ corresponds to the agents following the optimal policy and $\beta = 0$ corresponds to choosing actions uniformly at random; Ramachandran and Amir, 2007). The main computational issue with computing the likelihood is for every proposed value of the rewards, R , the Q function (1.12) must be estimated by a forward reinforcement learning algorithm in both Bayesian and optimization methods.

1.3.5 Linearly-Solvable MDP (LMDP)

An alternative class of MDP, the linearly-solvable MDP, introduced by Todorov (2009) is linear in its solution for the optimal policy and thus less computationally costly. The LMDP is defined by the tuple $(S, \bar{\mathbf{P}}, \gamma, R)$ where S , γ , and R are defined as the state space, discount factor, and state rewards, and $\bar{\mathbf{P}}$ are a set of passive transition dynamics between the states (i.e., $\bar{\mathbf{P}}: S \times S \rightarrow [0, 1]$). For a discrete state space, $S = \{1, \dots, J\}$, $\bar{\mathbf{P}}$ is a $J \times J$ transition probability matrix with elements \bar{p}_{ij} for $i = 1, \dots, J$ and $j = 1, \dots, J$ corresponding to the transition from state i to state j .

The LMDP replaces the discrete action set, A , of an MDP with a set of continuous controls, $\mathbf{u} = (u_{ij})$ for $i = 1, \dots, J$ and $j = 1, \dots, J$, such that the controlled dynamics are expressed as:

$$p(s_t = j | s_{t-1} = i) = p_{ij}(u_{ij}) = \bar{p}_{ij} \exp(u_{ij}), \quad (1.19)$$

where $\{u_{ij} \in \mathbb{R}: \text{if } \bar{p}_{ij} = 0, \text{ then } u_{ij} = 0\}$. The controlled dynamics define the policy (e.g., how to choose the next state) of the LMDP. The controls, u_{ij} , are interpretable as the cost the controller is willing to pay to go against the passive dynamics (Todorov, 2009). For a given policy, the state control costs, $l(i, \mathbf{u})$, for a discrete state space LMDP are:

$$l(i, \mathbf{u}) = q(i) + KL(\mathbf{p}_i(\mathbf{u}) || \bar{\mathbf{p}}_i), \quad (1.20)$$

where $q(i) = -r(i)$ is the immediate state cost for states $i = 1, \dots, J$ where $r(i)$ is an element of the state reward vector, R , and $KL(\cdot)$ is the Kullback-Leibler (KL) divergence between the controlled transition probability, $\mathbf{p}_i(\mathbf{u}) = (p_{i1}(u_{i1}), \dots, p_{iJ}(u_{iJ}))'$, and passive transition probabilities, $\bar{\mathbf{p}}_i = (\bar{p}_{i1}, \dots, \bar{p}_{iJ})'$. The KL divergence penalty requires the agent to "pay" a larger price for behavior that strays further from the passive dynamics (Todorov, 2007).

The state costs-to-go, v_i for $i = 1, \dots, J$, are the discounted sum of future expected costs incurred from beginning in state i :

$$v_i = E[l(i, \mathbf{u}) + \gamma \sum_{t=1}^T l(j, \mathbf{u})] \quad (1.21)$$

where the expectation is with respect to the controlled transitions (1.19). The value of T determines whether the problem has finite- or infinite-horizon (e.g., $T < \infty$ or $T = \infty$). A finite-horizon LMDP can be modeled as an infinite-horizon LMDP by assuming the agent remains in the final observed state and incurs no future costs (Todorov, 2007). Costs-to-go can also be interpreted as relative time to goal completion where a smaller cost-to-go indicates that the agent can reach a desirable state more quickly by transitioning to that state than transitioning to a state with a higher cost-to-go. The Bellman optimality for the state costs-to-go, v_i for $i = 1, \dots, J$, of an LMDP is:

$$v_i = \min_{\mathbf{u}} \left(l(i, \mathbf{u}) + \gamma \sum_j p_{ij}(u_{ij}) v_j \right). \quad (1.22)$$

The computational advantage of the LMDP is the Bellman optimality can be solved analytically using the method of Lagrange multipliers for the optimal transition probabilities:

$$p^*(s_t = j | s_{t-1} = i) = \frac{\bar{p}_{ij} \exp(-\gamma v_j)}{\sum_{k=1}^J \bar{p}_{ik} \exp(-\gamma v_k)}. \quad (1.23)$$

In Chapter 4, we use a basis function reduction on the state space to estimate the state value function in a Bayesian framework with MCMC sampling and variational approximation for a simulation and data application. There exist relatively few applications of IRL for real animal movement data and we illustrate the first application of IRL for the LMDP

framework on a collective animal movement dataset.

1.4 Overview of the Chapters

The chapters of this dissertation consider different learning strategies for ecological applications and follows the order of the previous three sections.

In Chapter 2, we demonstrate a flexible and efficient MM with multinomial logistic link for inference of habitat and other covariate effects on behavioral transition rates. The uncertainty from probabilistic machine learning classifications of accelerometer data is accounted for in the MM by the method of multiple imputation. We use Pólya-Gamma auxiliary variables for efficiency in the Bayesian MCMC sampling with a multinomial logistic likelihood. Ultimately, the model provides a unifying framework for including both accelerometer and GPS data.

In Chapter 3, deep ensemble ESN (D-EESN) and Bayesian D-EESN models are used for a long-lead ecological forecast of count data for mallard settling patterns. The annual settling patterns of mallards surveyed during the breeding season provide insight into productivity, which is used to set hunting regulations in the fall. The accuracy of a long-lead forecast of settling patterns could lengthen the management planning period. We show that D-EESN and BD-EESN are competitive forecast models compared to the current literature. The work is also the first use of a generalized linear model (GLM) output with an ESN for ecological forecasting. Furthermore, it is the first example of output weight learning for a BD-EESN with the horseshoe prior in a fully Bayesian inference paradigm.

In Chapter 4, the local rules governing observed trajectories of groups of animals (guppies) are learned by inverse reinforcement learning in a Bayesian framework. We

assume the guppy perceives its local misalignment and target misalignment forming the state space of an LMDP. By using inverse reinforcement learning, we learn the tradeoff the guppies make between traveling toward the target and traveling in a similar direction as its neighbors. It is the first application of IRL for the LMDP framework on a real collective animal movement dataset.

The methodology and datasets of the following chapters cover a wide spectrum, but the chapters are linked by an overarching theme. All of the applications and methodology estimate the transitions of an ecological state through time. In Chapters 2 and 4, the discrete state transitions are modeled explicitly through the transition probability function of a MM. On the other hand, in Chapter 3, the temporal evolution of the output system is approximated by a deep model where the internal network connections (e.g., the random weights in Figure 1.1) provide a fixed transition function for the hidden layers.

Chapter 2

A Bayesian Markov Model with Pólya-Gamma Sampling for Estimating Individual Behavior Transition Probabilities from Accelerometer Classifications

2.1 Introduction

Animals make decisions daily that can result in differential fitness (i.e., survival or reproductive success; Brown et al., 2004; Breed and Moore, 2015). The knowledge of an animal's behavior provides insight into its decision-making process. Historically, behavioral studies of animal populations were conducted using methods such as direct observation of focal individuals or instantaneous scan sampling of the group (Altmann, 1974). However, data collection by direct observation prevents a comprehensive understanding of the

decision-making process because it is limited to specific times and places when the animal is observable. The introduction of wildlife tracking devices has largely mitigated these constraints and provided observations over greater time and space. Recent improvements in tracking technology have allowed researchers to gather high frequency data over extended periods of time (e.g., >1 year), which has led to unprecedented insights into animal decision-making (Nathan et al., 2012; Leos-Barajas et al., 2017).

Behavioral inference from animal movement models has long been a research goal. Early work by Morales et al. (2004) showed the accuracy of movement state-space models based on location data only could be improved by including latent behavioral states allowing the model to switch between movement modes such as exploratory and encampment. Specifically, an exploratory state was assumed to be characterized by longer step lengths and smaller turning angles, whereas an encamped state was characterized by shorter step lengths and larger turning angles. State-space models considering more than two states typically require more information. Michelot et al. (2017) included constraints in the temporal sequence of latent behavioral states based on expert knowledge of the annual cycle of the study species to distinguish among four states. McClintock et al. (2017) used a three state model, and states were distinguished by the addition of an auxiliary data stream. State estimation in this context typically is accomplished via hidden Markov models (McClintock, 2017), which have proven to be an efficient method for estimating transition probabilities of an unobserved sequence of a categorical variable, such as behavior, that is associated with the values of observed quantities (e.g., step length and turning angle; Zucchini et al., 2016). In general, the estimated states are not interpretable as true behaviors, but rather proxies as they arise from a mixture model clustering procedure (Leos-Barajas et al., 2017; Patterson et al., 2017; Michelot and Blackwell, 2019). Therefore, when “ground truth” data are available

and inference on specific behavior states is the primary goal, behavioral inference is better achieved by building a classifier and then applying a behavior model to the classified states as we propose in this paper.

The inclusion of accelerometers in wildlife tracking devices has become increasingly common. An accelerometer is a tool for measuring an object's acceleration (ACC), and when placed on animals, can be used to derive energy expenditure and behavior of the tagged individual independent of location information (Nathan et al., 2012). Data collected from accelerometers are substantially different from Global Positioning System (GPS) observations in quantity, resolution and quality. That is, ACC data can be collected at a high frequency throughout the life of the tracking device, which results in richer and relatively larger data sets. The frequency of ACC collection can range from nearly continuous to more widely spaced intervals and is typically set to the highest frequency possible before affecting battery performance. Importantly, unlike location tracking devices, accelerometers do not typically miss "fixes" because the accelerometer instrument does not require linkages with external equipment such as satellites or radio trackers. In addition, ACC data are collected in two or three axes of movement relative to the device position, which provides the ability to discriminate between behavioral states with similar trajectory profiles.

The high frequency of data collected by accelerometers often allows researchers to "ground-truth" acceleration profiles with visual observations of the animals. When "ground-truth" observations of behavior are available, behavioral classification of ACC data can be conducted using machine learning algorithms (Resheff et al., 2014; Chakravarty et al., 2019). These approaches have greater reliability at identifying more than two states when compared to methods that consider only location data (Resheff et al., 2014). The richness of the ACC data allows for the derivation of many features to build the classification model.

The fitted classification model can be used to assign behavior labels to ACC data. Then the labeled data can be used to make inference on animal behavior in a second stage. Currently, researchers summarize the labeled data within a time scale of interest (e.g., days or hours) and use a traditional activity budget framework, such as linking proportions of observed behaviors to covariates through generalized linear mixed models (Broekhuis et al., 2014; Heurich et al., 2014). However, by analyzing the aggregate summarized data, inherent temporal structure and small scale processes in the data may not be fully utilized. The analysis of the activity budgets generally ignores temporal dependence in the data by modeling the proportions of behaviors separately. Rugg and Buech (1990) showed that a Markov model for behavior resulted in improved estimates of time allocation compared to traditional activity budget analyses. Additionally, the use of the labeled data ignores uncertainty associated with the chosen classification method.

We propose a two-stage framework in which we first build a behavioral state classifier and then explicitly model temporal dependence in the high frequency behavioral observations via a Markov model and multiple imputation. Multiple imputation is often used in missing data scenarios and averages inference across a suite of potential true data sets (Scharf et al., 2017; McClintock, 2017). We build a classifier using features of the accelerometer data then build potential data sets based on the classification probabilities or proportions. For a given behavior data set, our Markov model works directly on the behavioral state transitions and inherently assumes that current behavior depends on recent behavior (i.e., the Markov assumption states that given the most recent past, the current behavior is independent of the long-term past). The transition probabilities are modeled with a logistic function to link covariates to the probability of transition. Therefore, covariate effects retain the useful odds ratio interpretation as in multinomial logistic regression, and provide

insight into how covariates affect the tradeoff between time spent in each behavior. By using additional habitat information from location data, we can infer habitat use differences among behaviors. Resource selection models determine selection of habitats by animals based on differences in frequency of use and availability. As compared to traditional resource selection frameworks, we can make these inferences without having to define an "availability distribution" (Hooten et al., 2017). Although these approaches are important for identifying frequently used habitat, they do not define use and availability with respect to behavior. Our Markov model framework allows one to consider behavior profiles for different used habitats. Our approach provides the ability to answer questions about differential use of selected habitats that may be apparent from the focal animal's behavior. For instance, a resource selection study may identify two different agricultural crops as preferentially used, but these results do not identify the behavior associated with the crops or the differences in behavioral rates between the crops.

The relationship between energy expenditure and fitness is especially critical for migratory animals because the time allocation of behaviors during migration, such as feeding, can impact survival during migration and subsequent reproductive success (Harrison et al., 2011). We consider a long-distance migrant bird, the greater white-fronted goose (*Anser albifrons frontalis*), to demonstrate the inference capabilities of our Markov model. The migration route of white-fronted geese spans a wide range of habitats in mid-continent North America. It is known that geese use different habitat types for different behaviors (e.g., roosting in water and feeding in crops) and thus, we expect differences in behavior transition probabilities or rates depending on habitat (Krapu et al., 1995). Our Markov model has the ability to provide inference on potential differences in behavior transition rates among habitats that serve a similar purpose. For example, feeding may primarily

occur in agricultural fields, but rates of feeding may differ by crop type. Variability in rates of behavior transitions may also be related to environmental factors such as weather. Although activity budget analyses during winter have not found strong effects of weather (Ely, 1992), studies during spring have found important relationships between weather and the timing of migratory movements (Fox et al., 2003), and previous work has demonstrated heterogeneity in the movement of the geese (Hooten et al., 2018). We include weather variables comprising temperature and wind because we anticipate that these features can explain variation in behavior transitions during spring migration (e.g., favorable weather conditions increase rates of flight).

We implemented our Markov model in a Bayesian framework. At the core of our Markov model is multinomial logistic regression. There is a rich literature on Bayesian estimation of logistic models because sampling from the posterior distribution requires Metropolis-Hastings Markov Chain Monte Carlo algorithms that can be difficult to implement (tune) for moderate numbers of covariates or low observed frequencies of categories (Albert and Chib, 1993; Holmes and Held, 2006; Frühwirth-Schnatter and Frühwirth, 2010; Polson et al., 2013). To mitigate this challenge, automatic sampling strategies have been developed that rely on augmentation of data with latent variables. For example, Albert and Chib (1993) introduced the use of truncated normal random variables for category membership using probit link regression. Hooten et al. (2010) used the Albert and Chib sampling for a multinomial movement model. Holmes and Held (2006) expressed a multinomial logistic model as a product of binary models and used truncated scale mixture normals to define category membership. Alternatively, Polson et al. (2013) developed a class of Pólya-Gamma distributions for automatic sampling of Bayesian logistic regression models that is faster and provides exact inference when compared to the scale mixture latent variables.

The introduction of Pólya-Gamma latent variables induces simple Gibbs updating steps for the logistic regression coefficients when using a normal (Gaussian) prior. Additionally, Polson et al. (2013) showed the Pólya-Gamma scheme is significantly faster than other data augmentation methods and more efficient than Metropolis Hastings for mixed effects logistic models. Note that Pólya-Gamma latent variables have been used to estimate coefficients of a transition probability matrix of a hidden Markov model for rainfall data (Holsclaw et al., 2017) and to learn dependence structure in topic models for text mining applications (Chen et al., 2013; Linderman et al., 2015; Glynn et al., 2019). Pólya-Gamma latent variables have yet to be used for analysis of animal behavior data. Moreover, to the best of our knowledge, this is the first comprehensive example of Pólya-Gamma latent variables and multiple imputation within a Markov model.

The presented framework provides rich inference about covariate effects including information related to the location from GPS data on the sequence of observed behaviors derived from ACC data. We expect to estimate behavior transition probabilities for greater white-fronted geese during spring migration that confirm prior knowledge of the diurnal pattern of behavior and habitat use. The coefficients should reflect an increase in stationary behavior overnight in open water and wetland habitats while also suggesting more movement during the day. Additionally, we hypothesize that the estimates would exhibit variability among habitat types, suggesting variability in time allocation. Furthermore, we expect behavioral transitions associated with flight and feeding to be influenced by inclement weather more than walking and stationary behaviors. Our method provides a fine-scale picture of the behavioral decision-making process that is driven by both ACC and GPS data rather than mechanistic drivers of movement.

2.2 Materials and Methods

2.2.1 Dataset

Greater white-fronted geese, hereafter white-fronted geese, migrate from wintering areas in the southern US (i.e., Arkansas, Louisiana, Mississippi and Texas) to breeding areas in Alaska and northern Canada (Figure 2.1; Baldassarre, 2014). Individual female white-fronted geese were captured using rocket nets between December 2017 and January 2018 in Texas. Captured individuals were each fitted with a U.S. Geological Survey metal leg band and solar-powered Ornitela (<http://www.ornitela.com>) neck collar, comprising GPS, ACC and Global System for Mobile communications (i.e., for daily data upload) technology. Age, sex and morphometric measurements (normal wing cord [cm], head, culmen, tarsus, middle toe lengths [mm] and mass [nearest 0.1 kg]) were also collected. We set tracking devices to obtain ACC values at 10 Hz for 3 seconds at 6 minute intervals and GPS locations at 30 minute intervals with data upload every 24 hours. To demonstrate our modeling approach, we used GPS and ACC data from six white-fronted geese subset to 1-31 March 2018, which comprises a portion of the spring migration period (Figure 2.1). For simplicity, we chose a time period when all of the geese were migrating within the U.S. rather than defining the specific dates associated with the geese leaving wintering habitat and arriving at breeding areas. The average number of ACC fixes per individual was 7,257 with a range of 7,054 to 7,319.

Locations from February to May 2018

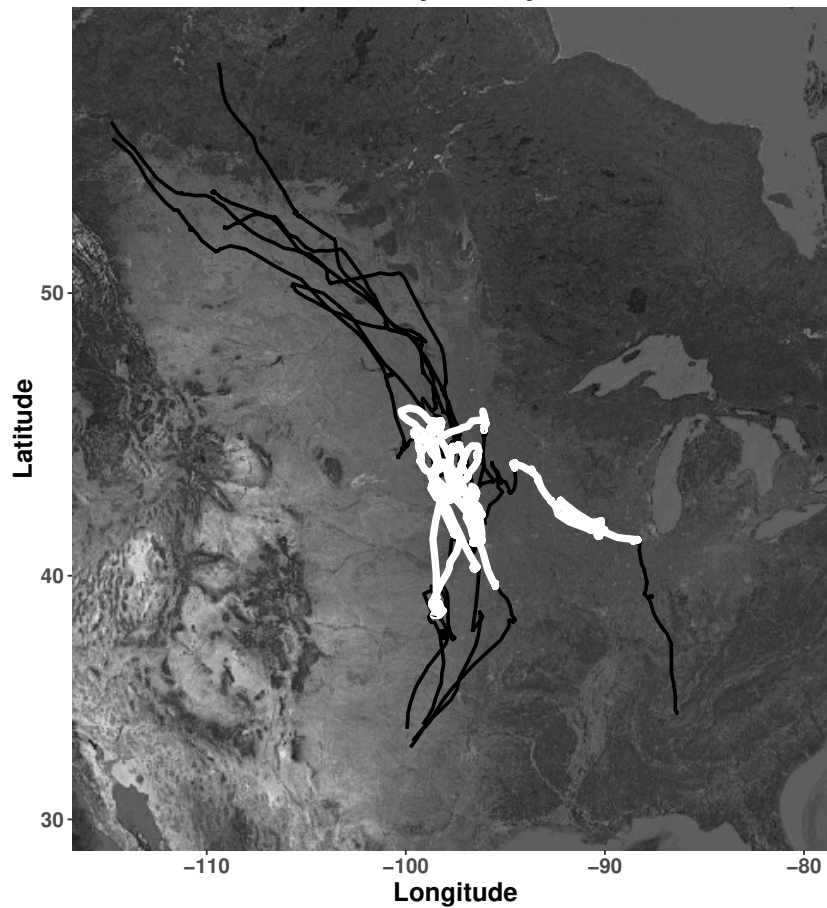


Fig. 2.1. Spring migration paths in North America of 6 individual greater white-fronted geese (*Anser albifrons frontalis*) from GPS equipped tracking devices. The highlighted section (white) are the locations for March 2018 used in the analysis of behavior transition probabilities.

The location data were used to determine habitat and weather factors experienced by geese during spring migration. The U.S. Department of Agriculture maintains a raster of habitat and crop data throughout the contiguous lower 48 states (CropScape; <https://nassgeodata.gmu.edu/CropScape/>), which is updated annually. We assumed the habitat at the GPS fix was best represented by the CropScape grid cell containing the observed point. The original CropScape categories were combined into fewer groups

(Appendix A.1). The National Centers for Environmental Prediction’s North American Regional Reanalysis (NARR; <https://rda.ucar.edu/datasets/ds608.0/>) data provide high resolution historical weather data in space and time. NARR is available eight times daily (i.e., every three hours) on an approximately 32 km grid. We assumed that the weather at the time of GPS fix was best represented by the NARR value corresponding to the nearest time and grid cell containing the observed point. We further assumed the weather and habitat at the time of an ACC fix were best represented by the values assigned to the most recent GPS location. The weather variables obtained were temperature [K], wind direction [°] and wind speed [m/s^2]. Temperature was further summarized to daily minima and maxima. Weather variables were standardized to have mean 0 and standard deviation 1. We assumed diurnal variability in behavior patterns because geese move most predictably near dawn and dusk between roosting and feeding areas; thus, we included such variability by using the local solar time of GPS fixes. We captured the diurnal behavior of the geese through two continuous covariates with a 24 hour period calculated from the local solar time of day (in seconds) by $\cos(2\pi(seconds)/86400)$ and $\sin(2\pi(seconds)/86400)$, referred to as $\cos(\text{time})$ and $\sin(\text{time})$. That is, $\cos(\text{time})$ is a representation of night (high values) and day (low values), while $\sin(\text{time})$ represents the first half of the day (high values) and the second half of the day (low values).

2.2.1.1 Behavior Classification

We summarized the ACC fixes into the 52 features described in the Appendix of Resheff et al. (2014). The features were used to classify the ACC observations into four behavioral categories (i.e., flying, feeding, stationary, and walking) using a random forest model. We chose a random forest model for its efficient handling of the large feature space. The number

of variables used to build the classification trees was varied from 1 to 15 and chosen by repeated 10-fold cross validation. The final model used 4 variables and had 96.5% accuracy on the training data set acquired by video recording. The random forest model was used to predict classification probabilities for the ACC fixes from the time period of interest. From the classification probabilities, we constructed $M = 200$ possible data sets for multiple imputation. We discuss the details of imputation in the Model Fitting section. For more details on fitting the classification model, see Appendix A.4.

2.2.2 Model

2.2.2.1 Behavior Transition Model

Each imputation data set is a regularly spaced time series of categorical behaviors, $\mathcal{S}_{nT_n} = \{s_{n0}, \dots, s_{nT_n} : s_{nt} \in \{1, \dots, J\}\}$, where s_{nt} is the observed behavior category for individual n , $n = 1, \dots, N$, and time point t , $t = 0, 1, \dots, T_n$, from the set of J behavioral categories. A Markov model for categorical time series data is defined by a transition probability matrix, \mathbf{P}_{nt} , describing the time-varying transition probabilities between observed states $s_{n,t-1}$ and s_{nt} as follows:

$$\mathbf{P}_{nt} \equiv \begin{bmatrix} p_{n11t} & p_{n12t} & \cdots & p_{n1Jt} \\ p_{n21t} & p_{n22t} & \cdots & p_{n2Jt} \\ \vdots & \vdots & \ddots & \vdots \\ p_{nJ1t} & p_{nJ2t} & \cdots & p_{nJJt} \end{bmatrix}, \quad (2.1)$$

where $p_{nijt} \equiv P(s_{nt} = j | s_{n,t-1} = i)$, and the row probabilities sum to one, $\sum_{j=1}^J p_{nijt} = 1$. Therefore, given the behavioral state at the previous time step ($t - 1$), the current observed category (at time t) is modeled as a multinomial trial with probabilities from

the corresponding row of the transition matrix (2.1). We let y_{nijt} be an indicator for individual n transitioning from state i at time $t - 1$ to state j at time t , in other words, $y_{nijt} = 1(s_{nt} = j | s_{n,t-1} = i)$ is defined to be 1 for the case when the n th individual is in state i at time $t - 1$ and state j at time t , and 0 otherwise. Then the aggregated transition indices vector $\mathbf{y}_{nit} = [y_{ni1t} \ \dots \ y_{niJt}]'$, along with the corresponding transition probabilities, $\mathbf{p}_{nit} = [p_{ni1t} \ \dots \ p_{niJt}]'$, describe a multinomial trial.

Covariates are introduced by the multinomial logistic link function on the elements of the vector of transition probabilities, \mathbf{p}_{nit} , (e.g., see Sung et al., 2007; Holsclaw et al., 2017):

$$\begin{aligned} \psi_{nijt} &\equiv \log((p_{nijt})/(p_{niJt})) = \mathbf{x}'_{nt} \boldsymbol{\beta}_{ij}, \quad \text{and} \\ p_{nijt} &= \exp(\psi_{nijt}) / \left(\sum_{k=1}^J \exp(\psi_{nikt}) \right), \end{aligned} \tag{2.2}$$

where we assume the vector of parameter coefficients $\boldsymbol{\beta}_{i,j} = 0$ for all $i = 1, 2, \dots, J$ for identifiability. In our application, J refers to the walking behavior. Therefore, coefficients for transitions to walking are 0 and do not appear in subsequent figures 2.2 and 2.3. Note that covariates may vary by individual and/or time, which induces non-homogeneous transition probabilities.

2.2.2.2 Parameter Model

The B -dimensional parameter vectors in the data model multinomial logistic functions, $\boldsymbol{\beta}_{ij}$, for $i = 1, 2, \dots, J$ and $j = 1, 2, \dots, J - 1$ are partitioned into three components:

$$\boldsymbol{\beta}_{ij} = [(\boldsymbol{\alpha}_{ij})' \ (\boldsymbol{\zeta}_{ij})' \ (\boldsymbol{\theta}_{ij})']'. \tag{2.3}$$

where α_{ij} is an $(N - 1)$ -dimensional vector of random individual effects, ζ_{ij} is an H -dimensional vector of habitat intercepts and θ_{ij} are $B - (N - 1 + H)$ -dimensional vectors of fixed quantitative covariate effects. The random individual effects are subject to a “sum to zero” constraint and therefore we only sample $(N - 1)$ coefficients. We assume independent normal distribution priors for each of the sampled coefficients as given below. The habitat coefficients for each transition are assumed to have the same mean, μ_{ij} . This mean is equivalent to the average intercept and, therefore, each habitat coefficient can be interpreted as the mean plus the habitat effect, similar to a cell means model in ANOVA. The common mean is assigned a flat prior. The hierarchical centering of these habitat parameters leads to less correlation between parameters compared to estimating H habitat effects and an intercept in the regression model (Gilks et al., 1995). In addition, centering provides more interpretability compared to including a baseline intercept and $H - 1$ coefficients. In summary, the prior distributions for these coefficients are given by

$$\begin{aligned}\alpha_{ij} &= \left[\alpha_{1ij} \quad \dots \quad \alpha_{(N-1)ij} \right]' \sim \mathbf{N}(\mathbf{0}, \sigma_{\alpha_{ij}}^2 \mathbf{I}), \text{ such that } \sum_{n=1}^N \alpha_{nij} = 0, \\ \zeta_{ij} &= \left[\zeta_{1ij} \quad \dots \quad \zeta_{Hij} \right]' \sim \mathbf{N}(\mu_{ij} \mathbf{1}, \sigma_{\zeta_{ij}}^2 \mathbf{I}), \quad p(\mu_{ij}) \propto 1, \\ \theta_{ij} &= \left[\theta_{1ij} \quad \dots \quad \theta_{B-(N-1+H)ij} \right]' \sim \mathbf{N}(\mathbf{0}, \sigma_{\theta_{ij}}^2 \mathbf{I}).\end{aligned}\tag{2.4}$$

where $\mathbf{0}$ and $\mathbf{1}$ are vectors of zeroes and ones, respectively. Note, all prior variances $(\sigma_{\alpha_{ij}}^2, \sigma_{\zeta_{ij}}^2, \sigma_{\theta_{ij}}^2)$ are fixed at 100 to induce a vague prior and reduce number of parameters needing to be estimated.

2.2.3 Model Fitting

The likelihood of each coefficient vector, β_{ij} , is the product of multinomial logistic functions. As the dimension of β_{ij} increases, the tuning of a Metropolis-Hastings algorithm becomes increasingly difficult. Latent variable schemes provide ways to automatically sample from the posterior without tuning and are generalized from binomial logistic regression to multinomial logistic regression (Albert and Chib, 1993; Holmes and Held, 2006; Frühwirth-Schnatter and Frühwirth, 2007; Polson et al., 2013). Specifically, the latent variable schemes express the likelihood as a product of binary logistic functions by the following transformation:

$$L(\beta_{ij} | \beta_{i,-j}, \mathcal{S}_1, \dots, \mathcal{S}_N) \propto \prod_{n=1}^N \prod_{t=1}^{T_n} \prod_{j=1}^{J-1} \left[e^{\psi_{nijt}} / \left(\sum_{k=1}^J e^{\psi_{nikt}} \right) \right]^{y_{nijt}}, \quad (2.5)$$

$$\propto \prod_{n=1}^N \prod_{t=1}^{T_n} \left[e^{\eta_{nijt}} / (1 + e^{\eta_{nijt}}) \right]^{y_{nijt}} \left[1 / (1 + e^{\eta_{nijt}}) \right]^{1 - y_{nijt}},$$

where $\eta_{nijt} = \psi_{nijt} - C_{nijt}$ and $C_{nijt} = \log \sum_{k \neq j} \exp\{\psi_{nikt}\}$. In this formulation, η_{nijt} is the log odds for a binomial random variable, which indicates whether or not the transition was from state i to j for individual n at time point t (Holmes and Held, 2006; Polson et al., 2013). From Theorem 1 of Polson et al. (2013), the product of logistic functions in the likelihood (2.5) is proportional to the product of an exponential of η_{nijt} and a Pólya-Gamma kernel. Therefore, by introducing the Pólya-Gamma latent variables, the sampling is done by conditional Gibbs updates alternating between the coefficients, β_{ij} , and Pólya-Gamma latent variables. The details for the derivation of the full-conditional distributions are in supplementary material (see Appendix A.3).

2.2.3.1 Multiple Imputation

The uncertainty associated with classifying the ACC data into categorical data by random forest is incorporated in the posterior distribution of the parameters via multiple imputation. Multiple imputation is used in situations involving missing data (e.g., the true behavior category) to provide approximate inference for the parameters based on the observed data (e.g., ACC fixes) (Rubin, 2004; Scharf et al., 2017; McClintock, 2017).

Let β be the collection of regression coefficients for all transitions, \mathcal{S} be the collection of the state transitions for all individuals, and \mathcal{A} be the collection of ACC data for all individuals. Our target distribution is the posterior of the unknown parameters, β , given the behavior labels, \mathcal{S} . However, we do not observe the true behavior labels directly, but we predicted the behavior labels by classifying the observed auxiliary data, \mathcal{A} . Therefore, our target distribution is

$$\begin{aligned} [\beta|\mathcal{A}] &= \int [\beta|\mathcal{S}, \mathcal{A}] [\mathcal{S}|\mathcal{A}] d\mathcal{S}, \\ &= \int [\beta|\mathcal{S}] [\mathcal{S}|\mathcal{A}] d\mathcal{S}, \end{aligned} \tag{2.6}$$

where given the behavior labels, \mathcal{S} , the parameters are conditionally independent of the ACC data. We assumed the distribution of the behavior labels given the ACC data, $[\mathcal{S}|\mathcal{A}]$, is the prediction from the supervised classification random forest.

We follow the multiple imputation MCMC algorithm outlined by Scharf et al. (2017) which numerically marginalizes over $[\mathcal{S}|\mathcal{A}]$ by randomly sampling from a set of M potential data sets where $M = 200$:

1. Draw realizations from the imputation distribution, $\mathcal{S}_{IT_1}^{(m)}, \dots, \mathcal{S}_{NT_N}^{(m)}$, for $m = 1, \dots, M$.
2. For each iteration of the MCMC repeat:

- Select an imputation data set with probability $1/M$.
- Use the Gibbs updating steps in Appendix A.3 with $(\mathcal{S}_{1T_1}, \dots, \mathcal{S}_{NT_N}) = (\mathcal{S}_{1T_1}^{(m)}, \dots, \mathcal{S}_{NT_N}^{(m)})$.

Throughout model fitting, we used “walking” as our reference behavioral category, J , in the multinomial logistic function (2.2). The coefficients for the reference were set to 0 and therefore are not displayed in the results. We chose walking as the reference category because we did not have specific hypotheses for this behavior and there was little variation in acceleration due to walking among individuals. Our method is amenable to choosing a different reference category specific to the study species. If it is of interest to more easily interpret effects on state duration, the reference category would depend on the previous state (i.e., for transitions from flight, the reference category would be flight).

We assessed parameter convergence by monitoring trace plots and initializing different random starting values (Brooks and Gelman, 1998). For inference, we sampled 15000 iterations from the model posterior and the first 5000 were discarded as burn in to ensure summaries were not influenced by the starting values. We generated behavior sequences from the current values of the parameters each iteration to investigate the goodness of fit of the model (Appendix A.4).

The posteriors for odds ratios were obtained by exponentiating the coefficients each iteration. The posterior samples for coefficients and odds ratios were summarized to the posterior mean and 95% credible intervals. We determined significance by whether the 95% credible intervals included zero for the quantitative covariates for weather and time of day. The significance of the quantitative coefficients corresponded to the 95% credible interval for the odds ratio not containing one and a proportion of >0.95 (positive effect on transition probability) or <0.05 (negative effect on transition probability) of iterations in which the odds ratio was >1 . For each transition probability, we investigated the differences in behavioral

transition probabilities among habitats by calculating the pairwise proportion of iterations in which habitat coefficients differed in magnitude, $\zeta_{aij} > \zeta_{bij}$, $a, b \in \{1, \dots, H\}$, and $a \neq b$. Two habitat coefficients for a transition were considered significantly different if the proportion of iterations with a difference in magnitude was >0.95 (the probability of transition from behavior i to j is greater in habitat a than habitat b) or <0.05 (the probability of transition from behavior i to j is lesser in habitat a than habitat b). We did not adjust the proportion cut off or widen credible intervals to account for inherent multiplicity in this case because we modeled the habitat coefficients with a common mean, which pooled the coefficient estimates (2.4). The hierarchical centering shrinks estimates toward the common mean, μ_{ij} , which makes it harder for significant pairwise differences to occur, thus eliminating the need to make additional *post hoc* adjustments for multiple comparisons (Gelman et al., 2012).

The assumption of a discrete time Markov process implicitly accommodates inference on the original sampling scale of the ACC data schedule (e.g., every 6 minutes). We interpreted all coefficient estimates as effects on behavioral transition probabilities from state i to state j at a 6 minute interval, relative to a base behavior state of walking. For brevity, we did not explicitly restate “relative to walking at a 6 minute interval” for each interpretation. Similarly, the odds ratios are interpreted as the multiplicative change in odds of transitions to state j from state i versus transitions to walking from state i .

2.3 Results

Our modeling framework establishes flexibility and efficiency in estimating covariate effects in the behavioral decision-making process by specifying transition-specific coefficients. The

variability in coefficients across from states is indicative of the complexity in the decision-making process in migrating white-fronted geese. The estimated effects of different habitats on behavior transitions rates did not follow the same patterns across the from states (Figure 2.2). By contrast, for weather and time of day covariates, the pattern in the coefficient estimates was similar across from states (Figure 2.3).

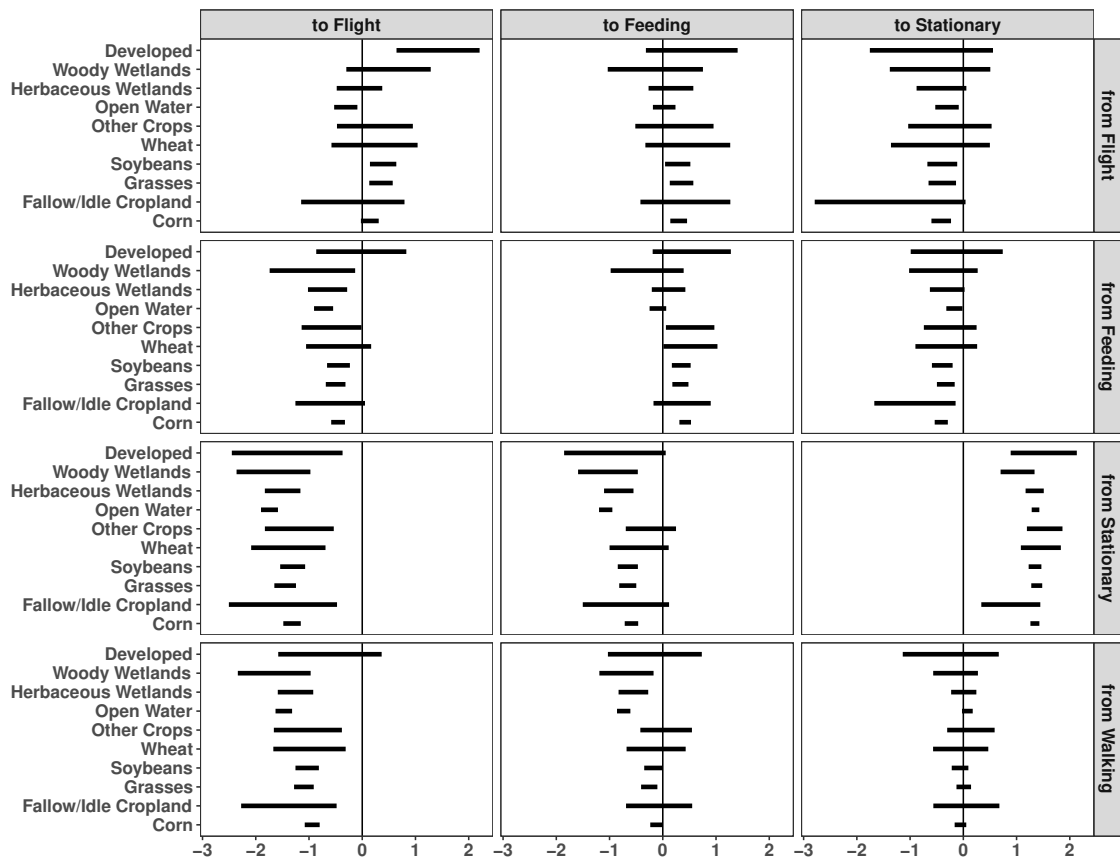


Fig. 2.2. 95% credible intervals for habitat effects on the log-odds scale estimated in a Bayesian framework by behavior transition for six greater white-fronted geese in March 2018. There are significant pairwise differences between intervals that do not overlap.

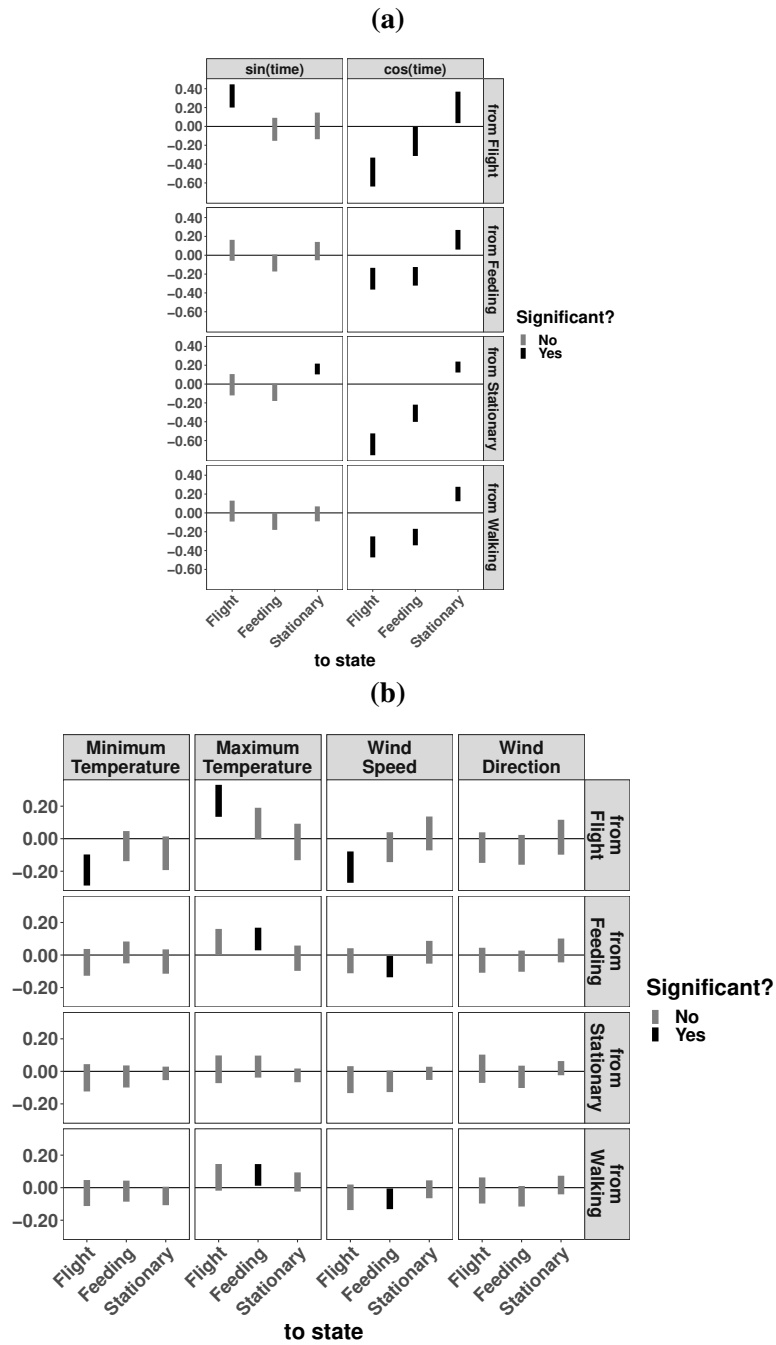


Fig. 2.3. 95% credible intervals for diurnal (a) and weather (b) covariate effects on the log-odds scale estimated in a Bayesian framework by behavior transition for six greater white-fronted geese in March 2018. Significance refers to whether or not the credible interval overlaps zero.

There were significant differences between habitats for every behavior transition probability except flight to stationary (see Appendix A.5). We estimated a significantly higher probability of transitioning to feeding from feeding (i.e., continued to feed), stationary and walking for food crops such as corn and soybeans and a significantly lower probability when the birds were in open water habitat (Figure 2.4). The positive coefficient estimates for corn and soybeans indicated an odds ratio >1 and an increase in the probability for transitioning to feeding from flight and feeding (Figure 2.2 to feeding subplots). For example, the mean effect of corn on the odds of continuing to feed was 0.42 with corresponding mean odds ratio 1.52, indicating the odds of continuing to feed at the next time step was 52% greater than the odds of transitioning to walking from feeding in corn habitat (Table 2.1). Transitions to stationary from feeding and walking were more probable in open water habitats than in corn (see Appendix A.5). In general, the widths of the credible intervals in Figure 2.2 have an inverse relationship with the observed frequency of habitats (see Appendix A.1 for frequencies). For example, the most frequently used habitats were corn and open water, which tended to have the narrowest credible intervals.

Table 2.1. Means and 95% credible intervals for odds ratios for habitat coefficients for the feeding to feeding transition arranged by increasing mean estimated in a Bayesian framework for six greater white-fronted geese in March 2018. A significant difference between habitats is indicated by non-overlapping credible intervals which can be visualized in Figure 2.4. The values correspond to exponentiating the estimates of the coefficients depicted in the from feeding to feeding panel of Figure 2.2

Covariate	Mean	95% CI	
Woody Wetlands	0.80	0.38	1.48
Open Water	0.91	0.78	1.07
Herbaceous Wetlands	1.13	0.81	1.53
Grasses	1.40	1.20	1.62
Soybeans	1.42	1.19	1.69
Fallow/Idle Cropland	1.49	0.84	2.46
Corn	1.52	1.37	1.70
Other Crops	1.72	1.06	2.64
Wheat	1.75	1.01	2.79
Developed	1.82	0.83	3.59

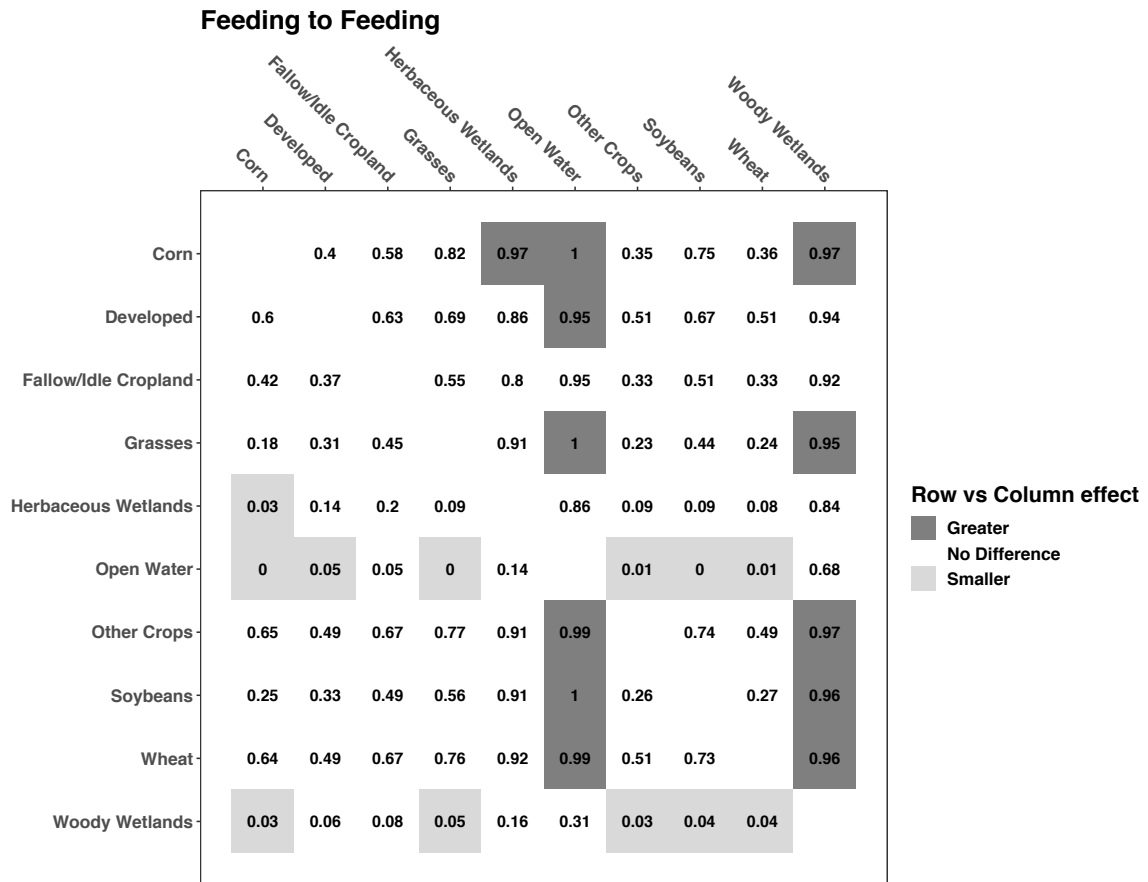


Fig. 2.4. Matrix of pairwise comparisons between habitat coefficients for the transition feeding to feeding for six greater white-fronted geese in March 2018. The values indicate the proportion of posterior MCMC iterations in which the habitat coefficient down the row was greater than the coefficients along the column; the upper triangular values and lower triangular values sum to 1. For example, in the first row, corn, the proportion of times the estimate for the corn habitat coefficient was greater than the open water habitat coefficient was 1. The rest of the transitions can be found in Appendix A.5

The probability of remaining in flight (i.e., transitioning to flight from flight) increased significantly during the first half of the day as indicated by the left most credible interval for $\sin(\text{time})$ in the top left subplot of Figure 2.3a. There were few ACC fixes classified as flight at night, which corresponds to the significant negative effects of $\cos(\text{time})$ on all transition probabilities to flight (Figure 2.3a). ACC fixes were classified as stationary most

frequently and in greater proportions at night. Therefore, it is not surprising that many diurnal coefficients corresponding to transition probabilities to the stationary state were significant. Table 2.2 shows that 6 out of 8 diurnal coefficients for transitions to stationary were significant. The transition probabilities to feeding and to flight from any behavior decreased overnight (negative effect of $\cos(\text{time})$) across all from states to feeding and flight; Figure 2.3a).

Table 2.2. A selection of means, 95% credible intervals and proportion of samples with an estimate greater than 1 for the odds ratios for quantitative covariate coefficients estimated in a Bayesian framework for six greater white-fronted geese in March 2018. The quantitative variable has a significant effect on the transition probability if the credible interval does not overlap with 1 which corresponds to a credible interval in Figure 2.3 not overlapping 0.

Transition	Covariate	Mean	95% CI		Proportion of Samples > 1
Flight to Flight	Maximum Temperature	1.26	1.14	1.39	1.00
	Minimum Temperature	0.83	0.75	0.91	0.00
	Wind Speed	0.84	0.76	0.92	0.00
Flight to Stationary		1.23	1.04	1.45	0.99
Feeding to Stationary	$\cos(\text{time})$	1.18	1.06	1.31	1.00
Stationary to Stationary		1.20	1.13	1.27	1.00
Walking to Stationary		1.22	1.13	1.32	1.00
Stationary to Stationary	$\sin(\text{time})$	1.17	1.11	1.24	1.00

Daily minimum and maximum temperatures did not affect the probability of remaining in flight the same way. Warmer than average daily minimum temperatures decreased the probability of remaining in flight (negative credible interval in top minimum temperature subplot of Figure 2.3b), while the warmer than average daily maximum temperatures increased the probability (positive credible interval in top maximum temperature subplot of Figure 2.3b). Increased wind speeds decreased the probability of remaining in flight and

transitioning to feeding from feeding and walking. We found no evidence of the weather variables effecting transition probabilities from stationary behavior. The probability of remaining in flight was significantly affected by minimum daily temperature, maximum daily temperature and wind speed, but not wind direction (Table 2.2).

2.4 Discussion

We provide a unified framework to connect variation in animal behavior with variation in habitat use and weather by propagating uncertainty in transitions using multiple imputation within a Bayesian Markov model with data from a long-distance migratory bird. Our approach is broadly applicable to other focal species and study systems across ecology. By analyzing data on the scale of frequency of collection, inferences are more intuitive and appropriate than aggregating to proportions. Importantly, our approach allows analysis of the behavior sequence with inherent temporal dependence and inference about covariate effects on behavior transition probabilities.

Transition matrices are well studied and can provide a wealth of inference and prediction beyond what is presented in this study including simulation of behavior sequences in different settings. Further, transition matrix models are used in ecological and evolutionary research beyond behavioral applications and Markov models. We extended our Markov model with a Pólya-Gamma sampling scheme, which will be useful for fast and automatic estimation of other complex ecological models that utilize the logistic link function.

We also implemented a unique approach to analyzing both ACC and GPS data from tracking devices. When we know the specific location of animals on the landscape, quantifying the effects of habitat on behavior transitions provides unprecedented information

regarding the differential rates of behavior in specific habitats. We made simple assumptions when we constructed our covariates by assigning values to ACC fixes from the most recent GPS fix, but there is the potential to link to existing animal movement models for interpolation of location data at times of ACC observations. McClintock et al. (2017) used a continuous time correlated random walk model to predict locations at a regular time interval (Johnson et al., 2008). The locations could be predicted to the time points of the ACC fixes and covariate values could reflect the prediction location or be imputed from the prediction distribution (Hooten et al., 2010; Scharf et al., 2017).

In the case of white-fronted geese during spring migration, different rates of behavior can be attributed to different habitats. Larger relative effects of habitats associated with food sources such as corn and soybeans on transition probabilities to feeding compared to effects of wetland habitats and open water aligns with previous knowledge of white-fronted goose ecology (Ely, 1992; Krapu et al., 1995). Open water consistently had lower transition probabilities than food habitats for transitions to feeding and flight. Our finding that white-fronted geese were less likely to transition to feeding or flight in open water compared to food habitats was consistent with our expectations. Also, we found significant effects of weather on behavior transition probabilities of white-fronted geese during spring migration. Higher winds decreased the odds of remaining feeding or transitioning to feeding from stationary or walking behavior. Duration of flight behavior (flight to flight transitions) was the most influenced by weather. The opposite effects of minimum and maximum daily temperature on flight durations may be indicative of more complex decision-making processes by white-fronted geese. An increase in flight duration with an increase in maximum daily temperature aligns with our *a priori* assumptions that these birds do not often migrate beyond the snow line during spring because food is relatively inaccessible under snow.

Furthermore, we were able to control for diurnal patterns in activity by using a continuous transformation rather than discretizing the time of day. The effects of time of day were consistent across from states. Specifically, the coefficients for $\cos(\text{time})$ all appeared to be significant in the same direction and magnitude in Figure 2.3a. In the future, it may be beneficial to relax the assumption that all covariate effects are transition specific and instead estimate a mix of transition-specific and behavior-specific coefficients. For example, Holsclaw et al. (2017) estimated a transition matrix for a Hidden Markov model with a transition-specific intercept and state specific coefficients for all covariates. They estimated weather covariates specific to rainfall states at weather stations in India to identify global effects of weather systems. For animal migration and behavior, the occurrence of certain weather systems or the time of day may always result in a specific decision regardless of the current behavior which suggests a model with behavior specific coefficients may be a better fit. Within the Bayesian framework, models with different parameter formulations can be compared using tools such as Bayes factors, information criteria, or prediction of hold out samples to test hypotheses about the behavior process (Hooten and Hobbs, 2015).

Our analysis was limited to a month-long subset of the spring 2018 migration and this pattern may become more clear with the inclusion of more years. Although much of the inference was verification of previous knowledge about white-fronted geese, the methodology allows us to infer about habitat use and behavior simultaneously using both the GPS and ACC information. Most importantly, we developed a detailed picture of time allocation in reference to the specific habitat types used by white-fronted geese during spring migration which is not addressed by traditional models used in previous activity budget analysis, movement trajectory prediction, or resource selection frameworks.

There is diminishing return on estimation and inference after a sufficiently large number

of imputation data sets are used (Scharf et al., 2017; McClintock, 2017). To investigate our choice of 200 data sets, we compared posterior inference among 200 imputation data sets, 100 imputation data sets and 1 data set corresponding to the most likely behavior classification. There was a general consensus among the three scenarios on the direction of the effects. Compared to the estimates based on the most likely classification, the estimates for the imputation scenarios were shrunk toward zero. Credible interval widths for the imputation scenarios appeared consistent or larger than the widths for coefficients estimated from the most likely behaviors (see Appendix A.6 for side-by-side comparisons).

Advances in animal tracking technologies continue to provide more frequently collected data for a greater duration of time. Thus, rich data sets are emerging at an unprecedented rate for ecologists and evolutionary biologists. Hence, there is an increasing need for development of models that appropriately handle the structure and volume of collected information for improved inference. Our Markov model framework provides much more capability for directly interpreting behavior patterns. In addition, the Pólya-Gamma latent variables facilitate for more efficient sampling and have yet to be used in the animal behavior and movement literature. If classification of behaviors is not feasible, the Pólya-Gamma sampling scheme can be incorporated into Bayesian estimation of transition probability matrices in a hidden Markov model framework (Holsclaw et al., 2017). The model directly handles temporal dependence in ACC data and learns about the behavior process from both ACC and GPS data. Our results suggest that new data sources coupled with appropriate modeling have unprecedented potential to provide a comprehensive understanding of complex ecological and evolutionary processes in animal movement.

Chapter 3

Hybrid Statistical and Deep Neural Models for Ecological Prediction

3.1 Introduction

The last two decades have shown a dramatic increase in ecological prediction (see the overview in Dietze, 2017b). As defined by Clark et al. (2001), ecological forecasting is defined as estimating the future states of ecosystems with fully specified uncertainties under explicit scenarios of environmental and anthropogenic factors. There is a growing number of iterative forecasting efforts for a variety of ecological processes and systems including phenology (Taylor and White, 2020), water quality (Beck and van Straten, 2012) and epidemiology (Estrada-Peña, 2001; Hobbs et al., 2015); see Dietze et al. (2018) for a review of several more examples.

In the context of predicting the future state of such complex processes, one must account for complex dynamics of the ecosystem processes of interest, as well as uncertainties in

our understanding of these systems and the associated data. Increasingly, a hierarchical statistical modeling framework has been used to represent such processes (Wikle, 2003; Clark and Gelfand, 2006; McMahon and Diez, 2007; Royle and Dorazio, 2008; Cressie et al., 2009; Hobbs and Hooten, 2015; Dietze, 2017a). Hierarchical modeling propagates uncertainty across many sources of variation. Aside from observational uncertainty, other sources of variability include process uncertainty, individual heterogeneity, and multiple data sources. The integration of observations, parameters, and latent processes in hierarchical modeling accounts for all of these uncertainties in a unified framework rather than using multi-stage frameworks with disjoint models. In ecology, hierarchical modeling is the foundation for integrated population models (Schaub and Abadi, 2011) and animal movement models (Hooten et al., 2017) to jointly analyze multiple data sets and/or individuals. The information and uncertainties in one part of the system is leveraged by other components allowing for more accurate estimates and uncertainty quantification.

Modeling complex processes that vary in space and time requires statistical or deterministic models that can accommodate the interaction of spatial components of variation through time (see the overview in Cressie and Wikle, 2011). For example, spatially-explicit diffusion models for ecological processes (Skellam, 1951; Cantrell and Cosner, 2004) and agent-based models (Grimm and Railsback, 2005) can be implemented within the process stage of a hierarchical Bayesian model (Hooten and Wikle, 2008; Hooten et al., 2010). Yet, in other cases the dynamics are less well-specified and one must resort to mechanistically motivated models (e.g., de Valpine and Hastings, 2002; Hooten and Wikle, 2010), or flexible latent process models, such as hidden Markov models (Zucchini et al., 2016) and generalized additive models (Fewster et al., 2000; Pedersen et al., 2019).

Another option for flexible process models is the class of neural models that have be-

come common in the last decade (Schmidhuber, 2015; Goodfellow et al., 2016). Such models are very powerful, especially for spatially (e.g., image recognition) and sequentially structured (e.g., natural language processing, machine translation, time series) data because they are able to learn multiple scales of features that can be used to approximate highly complex nonlinear functions. However, these methods are largely deterministic, with learning accomplished through stochastic gradient descent optimization and backpropagation. Although this leads to powerful prediction and classification algorithms, it severely limits formal uncertainty quantification. There have been various Bayesian and variational Bayesian implementations of neural models (Tran et al., 2019), but they tend to be very computationally expensive, limited to relatively shallow models, and/or fit with approximation algorithms rather than MCMC sampling (e.g., McDermott and Wikle, 2019a; see Polson et al., 2017 for a discussion of the history of Bayesian inference for deep learning).

The combination of neural models and hierarchical models, neural hierarchical models, aims to represent one component of a hierarchical model with a neural model. When coupling the two paradigms, often all of the model parameters are now learned by optimization of a custom loss function. For example, Joseph (2020) replaced the parameter model of classic hierarchical ecological models such as occupancy, capture-recapture, and N-mixture models with a neural network and optimized the negative log likelihood.

Recently, McDermott and Wikle (2019b) demonstrated that one could combine statistical uncertainty quantification and deep neural paradigms by using deep echo state networks (ESNs, a type of recurrent neural network (RNN); Jaeger, 2001) to generate time patterns (basis functions) that could be used in a regularized hierarchical regression. The advantage of the ESN in this framework over traditional RNNs or convolutional neural networks (CNNs) is that ESNs are a type of reservoir computing, in which the parameters (weights)

within the deep layers are not learned, but are simply sampled at random. Thus, the lower levels of the deep model can be calculated in an unsupervised manner, offline (and in parallel). The random initialization of weights greatly reduces the number of parameters and therefore data requirement typical of the RNN. As we show here, these levels provide multiscale basis functions that can be used in a frequentist or Bayesian generalized linear model (GLM) with shrinkage framework, much like a generalized additive mixed model (GAMM). Deep and wide ESNs were also proposed by Carmichael et al. (2018) with the main difference to our work being they estimate one set of output weights without uncertainty quantification of the random initialization of the weights.

We chose to demonstrate the forecasting performance of the deep ensemble ESN (D-EESN) and Bayesian D-EESN (BD-EESN) model using count data of settling patterns of mallard ducks in North America. The mallard settling pattern data has been used to demonstrate forecasting models for count data by McDermott et al. (2018) with a hierarchical analog model and Wu et al. (2013) with a Bayesian spatio-temporal Poisson model. Therefore, there is some established benchmarks on prediction performance for these data.

In the next section, we present the D-EESN model with weights estimated in a frequentist framework and the BD-EESN for generalized linear model output. In Section 3.3, we present an application of the methodology for long lead prediction of mallard settling patterns in North America. Finally, in Section 3.4, we summarize our findings and discuss future work.

3.2 Methodology

Following notation of McDermott and Wikle (2019b), suppose we observe a spatio-temporal process $\mathbf{Z}_t \equiv (Z_t(\mathbf{s}_1), \dots, Z_t(\mathbf{s}_{n_z}))'$ at the set of locations $\{\mathbf{s}_i \in D_z \subset \mathbb{R}^2 : i = 1, \dots, n_z\}$ for

time periods $\{t = 1, \dots, T\}$ and forecast the process with a lead time of τ , $\mathbf{Z}_{t+\tau}$. Additionally, assume we observe another temporal process, $\mathbf{x}_t \equiv (x_t(\mathbf{r}_1), \dots, x_t(\mathbf{r}_{n_x}))'$, for time periods $\{t' = 1, \dots, T'\}$, to be used as inputs (features, covariates, or predictors). The inputs may be indexed by space, $\{\mathbf{r}_d \in D_x \subset \mathbb{R}^2 : d = 1, \dots, n_x\}$, or may not, $\{\mathbf{r}_i \equiv r_d \subset R_d : d = 1, \dots, n_x\}$, depending on the application. Additionally, the temporal frequency of output observations may be less often than the input frequency (e.g., the output observations are annual and the inputs are monthly).

For observations of processes at many spatial locations, the spatial dimension of the process may be reduced through an appropriate linear dimension reduction technique for the observation type (e.g., empirical orthogonal functions (EOFs) for continuous spatial processes or non-negative matrix factorization (NMF) for non-negative spatial processes) such that $\mathbf{Z}_t \approx \mathbf{\Phi}\alpha_t$, where $\mathbf{\Phi}$ is an $n_z \times n_b$ matrix of n_b spatial basis functions and α_t is an n_b -dimensional vector of basis coefficients, indexed by $b = 1, \dots, n_b$ (e.g., Cressie and Wikle, 2011, Chapter 7). We then consider the basis coefficients, α_t , as the variables of interest and use the basis function matrices to transform back to the original spatial scale. Similarly, we may reduce the spatial dimension of the inputs in a similar manner and consider the basis coefficients as the predictors.

3.2.1 D-EESN with GLM

We generalize the output stage of the spatio-temporal D-EESN of (McDermott and Wikle, 2019b) to the generalized linear model for exponential family. For a deep ESN with L

layers,

$$\mathbf{Data\ Stage:} \quad \mathbf{Z}_t \approx \mathbf{\Phi} \boldsymbol{\alpha}_t, \quad (3.1)$$

$$\mathbf{Output\ Stage:} \quad \boldsymbol{\alpha}_t \sim \text{EF}(\boldsymbol{\mu}_t, \boldsymbol{\Theta}), \quad (3.2)$$

$$\mathbf{Link:} \quad g(\boldsymbol{\mu}_t) = \mathbf{V}_1 \mathbf{h}_{t,1} + \sum_{l=2}^L \mathbf{V}_l g_h(\tilde{\mathbf{h}}_{t,l}), \quad (3.3)$$

$$\mathbf{Shrinkage:} \quad \min_{\mathbf{V}_1, \dots, \mathbf{V}_L} \sum_t Q(\boldsymbol{\alpha}_t, g(\boldsymbol{\mu}_t)) + \lambda \|[\mathbf{V}_1, \dots, \mathbf{V}_L]\|_p, \quad (3.4)$$

$$\mathbf{Leaky\ Units\ } l : \quad \mathbf{h}_{t,l} = (1 - \zeta_l) \mathbf{h}_{t-1,l} + \zeta_l \hat{\mathbf{h}}_{t,l} \quad (3.5)$$

$$\mathbf{Hidden\ Stage\ } l : \quad \hat{\mathbf{h}}_{t,l} = f \left(\frac{\nu_l}{|\lambda_{W_l}|} \mathbf{W}_l \mathbf{h}_{t-1,l} + \mathbf{U}_l \tilde{\mathbf{h}}_{t,l+1} \right) \quad (3.6)$$

$$\mathbf{Reduction\ Stage\ } l + 1 : \quad \tilde{\mathbf{h}}_{t,l+1} \equiv \mathcal{T}(\mathbf{h}_{t,l+1}), \text{ for } l < L, \quad (3.7)$$

$$\mathbf{Input\ Stage:} \quad \mathbf{h}_{t,L} = f \left(\frac{\nu_L}{|\lambda_{W_L}|} \mathbf{W}_L \mathbf{h}_{t-1,L} + \mathbf{U}_L \tilde{\mathbf{x}}_t \right), \quad (3.8)$$

where $EF()$ is an exponential family with $E[\boldsymbol{\alpha}_t] = \boldsymbol{\mu}_t$, g is the link function for the exponential family, Q is the objective or loss function, λ is the shrinkage parameter for the penalty function $\|\cdot\|_p$ ($p = 1$ for LASSO and $p = 2$ for ridge regression), g_h and f are nonlinear activation functions, ζ_l is a leaking rate, λ_{W_l} is the spectral radius (largest eigenvalue of \mathbf{W}_l), ν_l is a scaling factor, and $\mathcal{T}(\cdot)$ is a dimension reduction function. The dimension reduction function maps the reservoir features to a lower-dimensional space for feedforward connections (i.e., $\mathcal{T} : \mathbb{R}^{n_{h,l}} \rightarrow \mathbb{R}^{n_{\tilde{h},l}}$ where $n_{h,l} \geq n_{\tilde{h},l}$). Typically \mathcal{T} is a linear basis dimension reduction technique with bases estimated by the temporal replicates (McDermott and Wikle, 2019b). The weight matrices correspond to recurrent connections, \mathbf{W}_l , feedforward connections, \mathbf{U}_l , and output connections, $\mathbf{V}_1, \dots, \mathbf{V}_L$. The function g_h matches the scale of the reduced features from the lower levels, $\tilde{\mathbf{h}}_{t,l}$ for $l \geq 2$, with the last

layer $\mathbf{h}_{t,1}$. We let $\tilde{\mathbf{x}}_t$ be a vector of q embeddings (lags) of the input process, \mathbf{x}_t , given by

$$\tilde{\mathbf{x}}_t = [\mathbf{x}'_t, \mathbf{x}'_{t-m}, \dots, \mathbf{x}'_{t-(q-1)m}]', \quad (3.9)$$

where \mathbf{x}_t is an n_x -dimensional vector, m is the lag of the embedding, and q is the number of lags of the input process. In the mallard settling pattern application, the temporal embedding lag of the input process is $m =$ one month and we used $q = 50$ lagged months. Both values were chosen based on the priors used by McDermott et al. (2018).

The elements of the square recurrent weight matrices, $\mathbf{W}_l = [w_{k_l, q_l}^{(l)}]$, and feedforward weight matrices, $\mathbf{U}_l = [u_{k_l, r_l}^{(l)}]$, of (3.6) are initialized randomly according to the following distributions

$$\begin{aligned} w_{k_l, q_l}^{(l)} &= \gamma_{k_l, q_l}^{w_l} \text{Unif}(-a_{w_l}, a_{w_l}) + (1 - \gamma_{k_l, q_l}^{w_l}) \delta_0, & \gamma_{k_l, q_l}^{w_l} &\sim \text{Bern}(\pi_{w_l}), \\ u_{k_l, r_l}^{(l)} &= \gamma_{k_l, r_l}^{u_l} \text{Unif}(-a_{u_l}, a_{u_l}) + (1 - \gamma_{k_l, r_l}^{u_l}) \delta_0, & \gamma_{k_l, r_l}^{u_l} &\sim \text{Bern}(\pi_{u_l}), \end{aligned} \quad (3.10)$$

where δ_0 is a Dirac measure at 0. The scaling of the recurrent weight matrices by their respective largest eigenvalue, λ_{W_l} , and $\nu_l \in (0, 1]$ ensures that the network is stable and gradually "forgets" over time (i.e., the echo state property; Jaeger, 2001) when the activation function, $f(\cdot)$, is tanh.

To account for uncertainty in the initialization of the random weights matrices, \mathbf{W}_l and \mathbf{U}_l , McDermott and Wikle (2017) introduced the ensemble ESN which randomly initializes a large number of reservoirs and estimates the reservoir specific output weights to create an ensemble of forecasts where each ensemble member acts as a weak learner (McDermott and Wikle, 2019b). The ensemble is created by repeating the following n_{res} times: randomly initializing a set of reservoir weight matrices, calculating the hidden units,

and estimating the output weights, $\mathbf{V}_1, \dots, \mathbf{V}_L$. Another useful interpretation of the ensembles is that one is simply generating a large set of potential predictors by stochastic, nonlinear transformation of the input. The output weights of D-EESN are typically estimated by ridge or LASSO regression because the dimension of the feature space is often larger than the output dimension (Lukoševičius and Jaeger, 2009). The estimated output weights are then used to forecast, giving an ensemble of forecasts which provide bootstrap estimates of uncertainty. We used the R package `glmnet` to estimate the output weights with a LASSO penalty (Friedman et al., 2010).

3.2.2 BD-EESN with GLM

The D-EESN is computationally efficient, but the lack of simultaneous output weight estimation does not allow for investigation of overall important features nor formal uncertainty quantification. A single regression model with shrinkage can learn feature selection among the ensemble members. In a Bayesian framework, the reservoir regression can be incorporated into a hierarchical model at the data, process, or parameter stage. An example of using the ensemble of ESN hidden features in the process stage of a Bayesian model (BD-EESN) is expressed as:

$$\text{Data Stage: } \mathbf{Z}_t | \alpha_t \sim \text{EF}(\boldsymbol{\mu}(\alpha_t), \boldsymbol{\Theta}), \quad (3.11)$$

$$\text{Process Stage: } g(\alpha_t) \sim \text{MVN}(\boldsymbol{\mu}_t, \sigma^2 I_{n_b}), \quad (3.12)$$

$$\boldsymbol{\mu}_t = \boldsymbol{\beta}_0 + \frac{1}{n_{res}} \sum_{j=1}^{n_{res}} \left[\boldsymbol{\beta}_1^{(j)} \mathbf{h}_{t,1}^{(j)} + \sum_{l=2}^L \boldsymbol{\beta}_l^{(j)} g_h(\tilde{\mathbf{h}}_{t,l}^{(j)}) \right], \quad (3.13)$$

where $\boldsymbol{\beta}_0$ is an n_b -dimensional vector of intercepts and α_t is considered to be a latent process with the features, $\mathbf{h}_{t,1}^{(j)}$ and $\tilde{\mathbf{h}}_{t,l}^{(j)}$ for all $j = 1, \dots, n_{res}$, calculated according to (3.5) - (3.8).

The coefficient matrices are defined for $l = 1, \dots, L$ and $j = 1, \dots, n_{res}$ as

$$\boldsymbol{\beta}_l^{(j)} \equiv \begin{bmatrix} \boldsymbol{\beta}_{l,1}^{(j)} \\ \vdots \\ \boldsymbol{\beta}_{l,n_b}^{(j)} \end{bmatrix} \quad (3.14)$$

where $\boldsymbol{\beta}_{l,b}^{(j)}$ for $b = 1, \dots, n_b$ is an $n_{h,1}$ -dimensional vector for $l = 1$ or $n_{\tilde{h},l}$ -dimensional vector for $l \geq 2$. The total number of regression coefficients is $n_{res} * n_b * (n_{h,1} + \sum_{l=2}^L n_{\tilde{h},l})$ which in general greatly exceeds the number of observations. A shrinkage prior on the coefficients is necessary to estimate the overparameterized model. McDermott and Wikle (2019b) demonstrated the estimation of the coefficients with a discrete mixture prior or two groups prior. We will use a continuous scale mixture prior, the horseshoe (HS) prior. The continuous mixture of the HS prior acts like Bayesian model averaging and thus, performs well on out of sample prediction problems (Carvalho et al., 2010). Specifically, we implemented the HS prior given by the following hierarchical distribution for $l = 1, \dots, L$ and $j = 1, \dots, n_{res}$:

$$\begin{aligned} \boldsymbol{\beta}_{l,b,k_l}^{(j)} | \sigma^2, \tau, \lambda_{l,b,k_l}^{(j)} &\sim \text{N}(0, \sigma^2 \tau^2 \lambda_{l,b,k_l}^{(j)2}), & \lambda_{l,b,k_l}^{(j)} &\sim \text{HalfCauchy}(0, 1), \\ \tau &\sim \text{HalfCauchy}(0, 1), & \sigma^2 &\propto 1/\sigma^2 \end{aligned} \quad (3.15)$$

where τ is the global shrinkage parameter and $\lambda_{l,b,k_l}^{(j)}$ is the local shrinkage parameter for $\boldsymbol{\beta}_{l,b,k_l}^{(j)}$. The global shrinkage controls the size of the model while the local shrinkage parameters determine which of the features are included (Carvalho et al., 2010). There is not an inclusion parameter in the horseshoe prior like there is in discrete mixture shrinkage priors such as the spike-and-slab or SSVS. Instead, the pseudo-inclusion probabilities are

$\kappa_{l,b,k_l}^{(j)} = 1/(1 + \tau^2 \lambda_{l,b,k_l}^{(j)^2})$ where if $(1 - \kappa_{l,b,k_l}^{(j)}) > 0.5$, then the coefficient is considered to be included in the model (Bhadra et al., 2019). The intercepts are not subject to shrinkage and a Jeffreys prior is used, $p(\boldsymbol{\beta}_0) \propto 1$.

3.2.2.1 MCMC

We implement the exact block algorithm of Johndrow et al. (2020) which blocks together updating of $(\boldsymbol{\beta}, \sigma^2, \tau^2)$ and samples the local shrinkage parameters using an exact rejection algorithm instead of the widely used slice sampler, which requires an additional set of latent variables. The details of conditional posteriors and updating steps can be found in Appendix B.2. In the application described in Section 3.3, we sampled 25000 iterations from the posterior discarding the first 5000 as burnin and thinning to keep every 10th iteration leaving 2000 posterior samples. We calculated the Gelman-Rubin statistics for a subset of the parameters to assess convergence based on two randomly initialized chains.

3.2.2.2 Hyperparameter Tuning

The size of the deep layers, $n_{h,l}$ for $l \geq 2$, magnitude of the non-zero weights, $a_{w,l}$ and $a_{u,l}$ for $l = 1, \dots, L$, and sparsity hyperparameters $\pi_{w,l}$ and $\pi_{u,l}$ for $l = 1, \dots, L$ were fixed as per recommendations by McDermott and Wikle (2019b). We trained all other hyperparameters for the ESN reservoirs on the training data with one-step-ahead prediction for the last training year using a genetic algorithm. We used the genetic algorithm in the R package GA with the default population size of 50, mutation probability 0.1, and crossover probability 0.8 (Scrucca et al., 2013). The ranges of the hyperparameters explored are in Appendix B.1 Table B.1.

All analysis was done with the programming language R (R Core Team, 2020). The

computation for the hyperparameter tuning and MCMC sampling was performed on the high performance computing infrastructure provided by Research Computing Support Services and in part by the National Science Foundation under grant number CNS-1429294 at the University of Missouri, Columbia MO. DOI: <https://doi.org/10.32469/10355/69802>.

3.3 Mallard Settling Patterns

3.3.1 Background and Data

The settling patterns of waterfowl in the North America is an indicator of population productivity used to guide conservation decision making (Hansen and McKnight, 1964). The main driver of stochasticity in settling patterns is environmental habitat conditions such as wetness of ponds (Johnson and Grier, 1988; Feldman et al., 2015). Drought and other North American climatic conditions strongly relates to the processes of La Niña and El Niño (e.g., Philander, 1990) as measured by Pacific Ocean sea surface temperature (SST). The relationship between settling patterns and SST is indirect, highly nonlinear, and difficult to express mechanistically (Hoerling et al., 1997). By using the D-EESN and BD-EESN, we build a rich set of features of SST with multiscale variability for forecasting settling patterns without *a priori* specifying the mechanism of the nonlinear relationship.

Long-lead forecasting of waterfowl settling patterns is advantageous for wildlife conservation because it potentially increases the time horizon for conservation planning such as setting hunting regulations and allocating habitat management resources. Since 1955, the U.S. Fish and Wildlife Service and Canadian Wildlife Service has annually conducted the

waterfowl breeding population and habitat survey (BPS) in May to guide decision making for fall regulations. The survey consists of counting individual males and breeding pairs of migratory waterfowl species observed along aerial transects (Smith, 1995). While a pilot flies the transect at 30-50m above ground, an in air observer counts waterfowl within 200m on either side. At some transects, the aerial counts are matched with ground surveys. The count data within transect is reported per 29km segment. In this application, we considered BPS count data for mallards (*Anas platyrhynchos*) from 1925 segments with a spatial extent of 85 – 165°W and 43 – 69°N for the years 1970 to 2014 obtained from (<https://migbirdapps.fws.gov>) to compare to previous works (Wu et al., 2013; McDermott et al., 2018).

As a predictor of mallard settling patterns, we used monthly Pacific Ocean SST from the National Oceanic and Atmospheric Administration (NOAA) Extending Reconstruction Sea Surface Temperature data for the years 1970 to 2013 (<http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/>). The spatial extent of the SST data was between 29°S–29°N latitude and 123.5 – 290.5°E longitude with a spatial resolution of $2 \times 2^\circ$ (i.e., 2229 spatial locations). The raw SST were transformed to anomalies by centering observations by the spatial means of monthly temperature (Wilks, 2011).

3.3.2 Forecasting Application Specifics

As mentioned in Section 3.2, the spatial dimension of the spatio-temporal process can be reduced by an appropriate dimension reduction technique. As in (McDermott et al., 2018), we reduced the spatial dimension of the BPS count data with non-negative matrix factorization (NMF; e.g., Lee and Seung, 2001) which allows for computation of spatial basis functions from the raw BPS data and ensures non-negativity in the projection back

to the original space. Briefly, NMF for the spatio-temporal data matrix $\mathbf{Z} = [\mathbf{Z}_1, \dots, \mathbf{Z}_T]$ is defined as:

$$\mathbf{Z} \approx \mathbf{\Phi}\mathbf{A}, \mathbf{\Phi} \geq 0, \mathbf{A} \geq 0, \quad (3.16)$$

where the inequality is applied element-wise to the $n_z \times n_b$ basis function matrix, $\mathbf{\Phi}$, and $n_b \times T$ basis coefficient matrix, $\mathbf{A} = [\alpha_1, \dots, \alpha_T]$. The number of basis functions, $n_b = 14$, for mallard settling patterns was chosen by (McDermott et al., 2018) to balance computational cost and approximation accuracy so we used the same number of NMF basis functions in our analyses. The estimation of the NMF basis function matrix and coefficients was done with the offset algorithm of (Badea, 2008) which includes an intercept term in an attempt to quantify a common baseline and the algorithm is initialized by non-negative double singular value decomposition (Boutsidis and Gallopoulos, 2008). We used the implementation of the algorithm and initialization in the R package NMF (Gaujoux and Seoighe, 2010).

We also reduced the spatial dimension of the SST anomalies using empirical orthogonal functions (EOFs). Empirical orthogonal functions (EOFs) are a dimension reduction technique analogous to principal component analysis for spatial and spatio-temporal data (see Cressie and Wikle, 2011). We kept 16 basis functions which accounted for 90.9% of the variation in the SST anomalies. We used an input embedding matrix of 50 consecutive months of SST EOFs lagged one year (i.e., $q = 50$ and $m = 1$ month in (3.9)). All ensembles consisted of 100 randomly initialized ESNs with the output weights of the D-EESN estimated using the glmnet package for Poisson regression with LASSO shrinkage.

For the BD-EESN, we assumed a Poisson distribution for the BPS data stage (3.11) with

a log link function:

$$\mathbf{Z}_t | \alpha_t \sim \text{Poisson}(\mathbf{\Phi} \alpha_t), \quad (3.17)$$

$$\text{Process Stage: } \log(\alpha_t) \sim \text{N}(\mu_t, \sigma^2 I_{n_b}), \quad (3.18)$$

$$\mu_t = \beta_0 + \frac{1}{100} \sum_{j=1}^{100} \left[\beta_1^{(j)} \mathbf{h}_{t,1}^{(j)} + \sum_{l=2}^L \beta_l^{(j)} g_h(\tilde{\mathbf{h}}_{t,l}^{(j)}) \right], \quad (3.19)$$

where we scaled the basis function matrix, $\mathbf{\Phi}$, by the respective variation in the basis coefficients to meet the assumption of homogeneous variances.

3.3.3 Results

We fit D-EESN and BD-EESN models with 2 to 7 layers for out-of-sample prediction of years 2009 and 2014. The forecasts were evaluated based on three measures: mean squared prediction error (MSPE), Pearson correlation, and continuous rank probability score (CRPS). The MSPE was calculated as the mean of the squared difference between the true counts and the median of the forecast samples. The Pearson correlation was calculated between the true counts and the median of the forecast samples. The CRPS was calculated for each spatial location based on the forecast samples using the empirical distribution function method found in the R package `scoringRules` (Jordan et al., 2019).

The D-EESN models were less variable across the layers and better performing than the BD-EESN models as indicated by the prediction measures in Table 3.1. Additionally, prediction performance was better for 2009 than 2014 even though the models for 2014 prediction used more data. D-EESN models with 2, 3, and 7 layers all performed similarly for predicting settling patterns in 2009. The best model for predicting 2014 was the D-EESN with 5 layers and the best BD-EESN model was the 3 layer model. There appears to be no

pattern in prediction performance as the number of layers increases.

We compared visually the performance of the models in the region bounded by 96 – 115°W longitude and 43 – 54°N latitude, because this is where the models performed most poorly and was the spatial extent of the data used in McDermott et al. (2018). The top D-EESN and BD-EESN models for both years had comparable median forecasts, but differed substantially in the standard deviations of the forecast distribution (Figures 3.1 and 3.2). The reason for the drastic difference was due to the longer tails of the forecast distributions from the BD-EESN models as seen by darker values in the maps for the 97.5th percentile from the BD-EESN models compared to the D-EESN models which motivated the comparisons of prediction performance on medians.

Table 3.1. The mean squared prediction error (MSPE), Pearson correlation, and spatial median of the continuous rank probability score (CRPS) for one-step-ahead forecast of mallard settling patterns in 2009 and 2014 with various architectures of the D-EESN and BD-EESN. Smaller indicates better prediction for MSPE and CRPS while larger correlations are better.

Year	Layers	D-EESN			BD-EESN		
		MSPE	Correlation	CRPS	MSPE	Correlation	CRPS
2009	2	38.06	79.48 %	1.49	50.81	70.62 %	3.04
	3	38.17	79.14 %	1.52	44.28	76.88 %	1.73
	4	40.16	79.36 %	1.50	42.35	78.04 %	1.57
	5	52.24	77.41 %	2.09	39.06	78.32 %	1.95
	6	50.71	77.29 %	2.06	46.35	77.26 %	2.18
	7	38.59	79.33 %	1.50	176.67	68.19 %	5.43
2014	2	46.97	82.17 %	2.06	57.40	83.41 %	1.90
	3	47.91	81.92 %	2.00	44.64	82.54 %	1.88
	4	44.84	82.47 %	1.94	50.79	83.13 %	1.99
	5	46.50	81.35 %	1.86	58.74	82.34 %	1.96
	6	44.92	82.26 %	1.91	47.81	81.06 %	2.19
	7	48.65	82.30 %	1.98	47.65	81.29 %	2.12

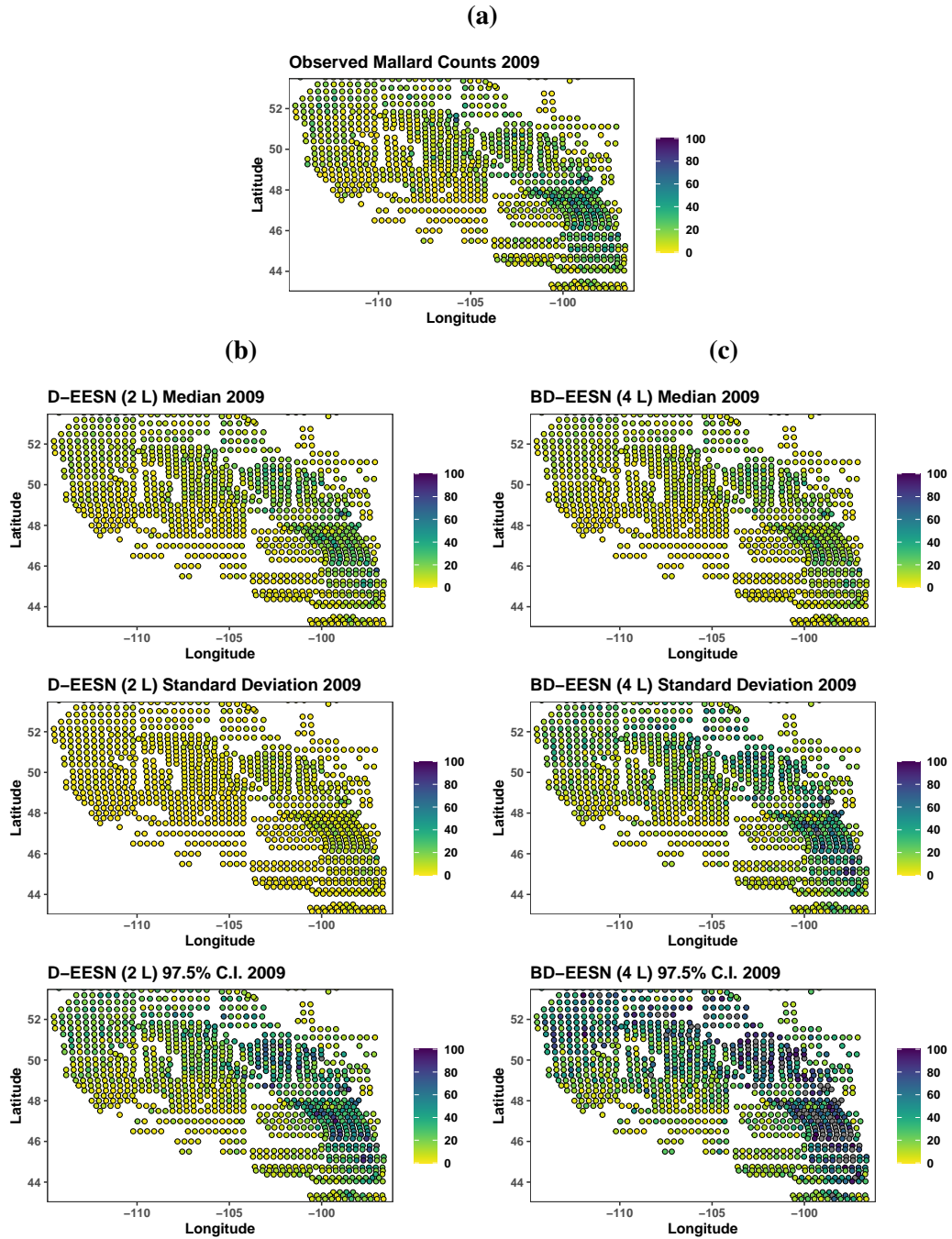


Fig. 3.1. Spatial maps of the true mallard settling pattern for 2009 (a) and side-by-side comparisons of forecast medians, standard deviations, and the 97.5th percentiles of the forecast distribution for the best D-EESN (b) and BD-EESN (c) model based on the CRPS found in Table 3.1. Grayed circles indicate values that exceed the legend threshold.

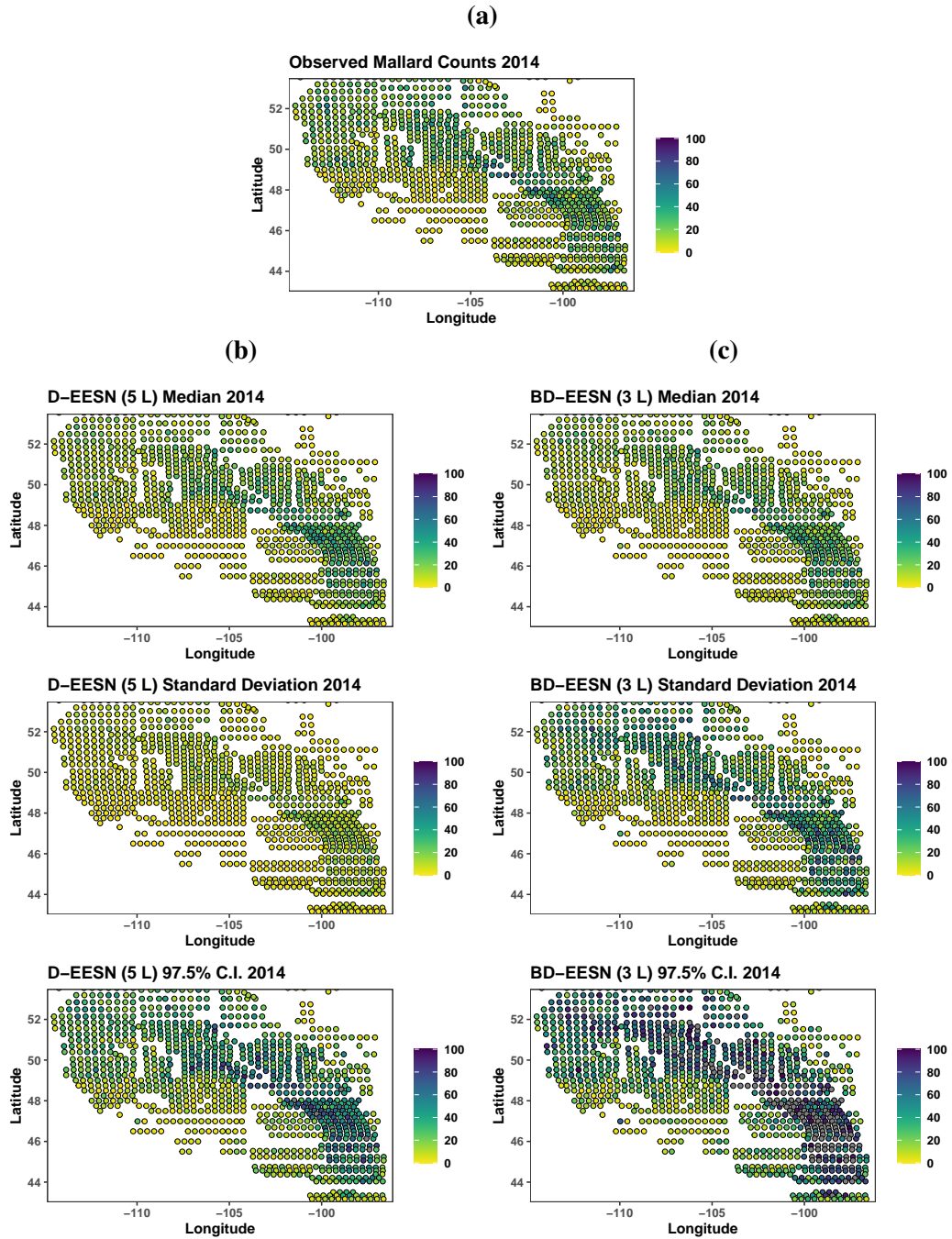


Fig. 3.2. Spatial maps of the true mallard settling pattern for 2014 (a) and side-by-side comparisons of forecast medians, standard deviations, and 97.5th percentiles of the forecast distribution for the best D-EESN (b) and BD-EESN (c) model based on the CRPS found in Table 3.1. Grayed circles indicate values that exceed the legend threshold.

In the BD-EESN models, the pseudo-inclusion probabilities calculated from the horseshoe priors did not indicate overly strong evidence for inclusion of any of the basis functions, with none of the values exceeding 0.5. For the best two BD-EESN models for the 2009 and 2014 forecasts, we investigated features with some evidence (pseudo-inclusion greater than 0.1). The top two BD-EESN model for the 2009 forecast had 5 and 2 features respectively while the top two 2014 forecasts had 2 and 4 features respectively which met this criterion (Figure 3.3). We can see the chosen features for the top models of both scenarios generally are smooth with peaks at varying time points (Figure 3.3a and 3.3b) and interestingly, it appears both sets of features have a dip around the middle time period. For the second best models, we see evidence of more multiscale features.

When compared to the prediction performance of the hierarchical analog models and Poisson spatio-temporal model of (McDermott et al., 2018), the spatial patterns of the prediction performance were similar, but the tails of our forecast distributions were larger in general. Therefore, we expect the prediction performance of our models in the subset of locations to be similar or potentially worse. Wu et al. (2013) forecasted mallard settling patterns across the entire spatial and temporal extent of the survey for the out-of-sample year 2010 using a Bayesian spatio-temporal Conway-Maxwell Poisson model with dynamic dispersion and reported MSPE with a range of 36.96 for the best model and 41.64 for the worst, which is slightly better but overlapping with our reported performance for other out-of-sample years.

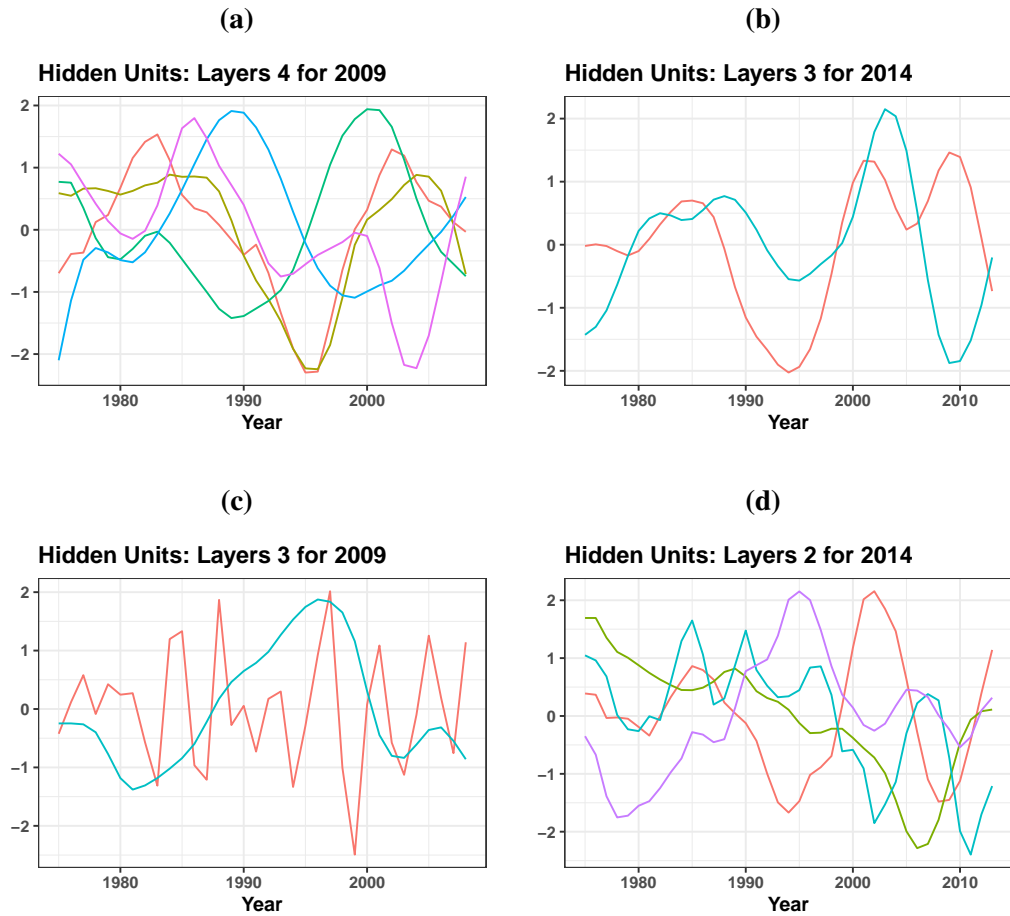


Fig. 3.3. Hidden units with a pseudo-inclusion probability greater than 0.1 for the top two BD-EESN models based on CRPS found in Table 3.1

3.4 Discussion

Ecological forecasting in nonlinear systems is challenging when the domain knowledge is not sufficient to develop a mechanistic model or adequate predictors. For example, the relationship between mallard settling patterns and North American climate as measured by Pacific Ocean sea surface temperature is apparent but the mechanisms are uncertain. By creating a nonlinear basis expansion of the input features offline, the ESN modeling

framework has the capacity to learn a complex, temporal relationship without specification by the modeler. To the best of our knowledge, we demonstrated the first example of a deep ESN with a GLM output for a machine learning task.

The structure of the deep ESN attempts to learn multiscale temporal interactions as we saw in the hidden units for the settling pattern prediction (Carmichael et al., 2019). Additionally, the use of the embedded input features assumes instantaneous dependence on some combination of the past. The ecological (antecedent) memory models of Ogle et al. (2015) are another example of a flexible framework which learns the temporal dependence between the input and output processes without complete specification by the modeler. One advantage of the ecological memory models is the sequential interpretability of the estimated temporal dependence. However, the disadvantage is the relationship between inputs and outputs are still assumed to be linear and therefore, we expect the ESN modeling framework to perform better for nonlinear processes.

As mentioned in McDermott and Wikle (2019b), the poor prediction performance of BD-EESN models compared to D-EESN models may be due to the tuning of hyperparameters in a D-EESN framework and then used in the subsequent BD-EESN models. This may especially be true in this application due to the LASSO penalty on the D-EESN allowing coefficients to be shrunk to zero, but the HS prior in the BD-EESN model does not allow for zeros (Bhadra et al., 2019). Furthermore, tuning was assessed on one-step-ahead prediction for one year and improvements may be made by instead considering validation measures based on performance across multiple prediction years. An additional component to the shrinkage prior to thin the amount of features used in the Bayesian model may improve performance especially in small data situations. The limited data size may also have been a contributor to the poor performance. The number of features, and therefore parameters,

greatly outnumbered the number of observations and was more extreme than the simulations used to test the HS algorithm in Johndrow et al. (2020). Lastly, in future implementations of the HS prior, it would be beneficial to group shrinkage by component of the basis coefficient vector which would allow for heterogeneous regression variances.

The time period for the mallard settling pattern data was relatively short and during feature initialization it appeared that under the given degree of sparsity some features degraded to zero in the feedforward connections. A future direction for ESN model development would be to study the behavior of features as more layers are added to the model. We hypothesize given a fixed degree of sparsity and weight magnitude there is a sufficient number of layers and ensembles needed to generate a diverse set of random basis functions dependent on the size of the data.

Although the presented results and previous studies by McDermott et al. (2018) and Wu et al. (2013) are not directly comparable due to different data extents, in general, prediction performance appears poor in the same spatial regions and the Bayesian Poisson methods tend to exhibit too much variation for locations with large counts. All of these works consider a spatial dimension reduction of the count data, but the spatial data is not equally spaced across the survey. One future extension could be to include information about transects which are nearby (or far apart). The survey groups together transects into strata and the strata information could be used to introduce random effects or potentially even estimation of strata specific basis functions in addition to global basis functions. Another point of investigation is evaluating the prediction performance of the Bayesian model with known basis function coefficients as it is possible the larger prediction errors could be accounting for the uncertainty in the loss of information due to the dimension reduction.

In conclusion, the ESN framework provides a competitive ecological forecasting model

for processes with multiscale nonlinearities and process interactions. The primary benefit of the ESN is the efficient basis expansion for learning the complexities of a system without complete domain knowledge.

Chapter 4

Bayesian Inverse Reinforcement Learning (IRL) for Collective Animal Movement

4.1 Introduction

Understanding individual animal decision-making processes in social groups is particularly elusive. Traditionally, agent-based models (ABMs) of individual interactions are used as building blocks for complex group dynamics (Vicsek et al., 1995; Couzin et al., 2002; McDermott et al., 2017). ABMs attempt to recreate what is observed in nature by defining a mechanistic model *a priori*. While the simple individual-based rules lead to complex group dynamics, ABMs suffer from automatic behavior after reaching some equilibrium, challenges to incorporate interactions with habitat, and no notion of memory (Ried et al., 2019). The goal of inverse modeling is to instead learn the underlying local rules from observations of sequential behavior decisions (Kangasrääsiö and Kaski, 2018; Yamaguchi

et al., 2018).

Parameters of ABMs in practice need to be tuned or learned by supervised learning (Ried et al., 2019). A recent alternative to supervised learning is reinforcement learning (RL). RL is goal-oriented learning from continuous interaction between an agent and its environment (Sutton and Barto, 1998). Therefore, many methods learn parameters controlling global behavior by trial and error experiments within the defined environment and local rules. The agent's learn preferences by receiving rewards from (or paying costs to) the environment and choose optimal behavior by minimizing the cumulative expected future costs (also referred to as costs-to-go). Similar to difficulty in tuning ABMs, defining the reward function to produce desired long term behavior is challenging (Ng and Russell, 2000; Arora and Doshi, 2018).

In systems where observations of behavior trajectories can be observed, inverse reinforcement learning (IRL) methods aim to learn the state rewards or values that governed the agents decisions. Ng and Russell (2000) introduced the first IRL algorithms including dynamic programming, which solves a system of equations based on the transition probabilities and a grid search method from trajectory samples. As surveyed by Arora and Doshi (2018), many more methods have since been developed or adapted to tackle problems of meaningful size and non-identifiability of the rewards. The methods can be broadly categorized as maximum margin optimization (Ratliff et al., 2006), entropy optimization (Ziebart et al., 2008), Bayesian IRL (Choi and Kim, 2011; Jin et al., 2017; Šošić et al., 2017b), and deep learning IRL (Wulfmeier et al., 2015), with majority of the methods being applied to Markov decision processes (MDPs). The benefit of Bayesian frameworks to the non-identifiability of the IRL problem is that they provide a distribution of values that can generate the observed expert behavior.

A big computational challenge to many of the aforementioned methods parameterized by the immediate state costs is the necessity to solve the forward MDP each iteration. This is especially challenging in multiagent MDPs, which describe collective animal movement due to calculations of distance between agents for describing an agent's state or perception of the environment. An alternative class of MDP, the linearly-solvable MDP (LMDP) introduced by Todorov (2009), is linear in its solution for the optimal policy and thus, less computationally costly for forward modeling. The LMDP is defined by a set of passive dynamics which describe an agent's state transitions in the absence of state costs or environmental feedback and then the optimal state transitions minimize costs-to-go. Moreover, IRL for LMDPs does not require the forward solution for each iteration as there is a linear relationship between the costs-to-go and immediate costs. Therefore, inference about immediate costs can be obtained by transformation of the estimated costs-to-go. As a special case, Dvijotham and Todorov (2010) showed the maximum entropy IRL is the solution to a LMDP with uniform passive dynamics. Kohjima et al. (2017) proposed a Bayesian IRL method for learning state values for LMDPs using variational approximation. As an extension, in this Chapter, we reduce the dimension of the state space with basis function approximation and compare variational approximation to MCMC sampling.

As argued by Ried et al. (2019), an MDP (or LMDP) for collective animal movement is a better model for the system than traditional self-propelled particle models (Vicsek et al., 1995). The MDP incorporates the internal processes of an animal by modeling the behavior as perception (state space), planning (state values), and action. Furthermore, the behavior is governed by feedback from the environment (which includes other agents) rather than assuming automatic interaction rules. Few applied examples of IRL for collective animal movement exist in the literature. Among those are the application of maximum entropy

IRL to flocking pigeons of Pinsler et al. (2018) and Bayesian policy estimation of the SPP and Ising models (Šošić et al., 2017a).

We present an application of IRL for collective animal movement using Bayesian learning of state values for an LMDP. We first demonstrate the modeling framework for a simulation of the Vicsek SPP model to illustrate the mechanisms of the LMDP framework. Then, in Section 4.4, we use the new methodology to estimate state costs-to-go for collective movement of guppies in a tank to infer trade offs between targeted motion and group cohesion. We discuss the findings and direction for future work in Section 4.5. Finally, in Section 4.5.1, we wrap up with a short discussion on the connection between the LMDP and current animal movement modeling.

4.2 IRL Methodology

4.2.1 LMDP

The LMDP is defined by the tuple $(S, \bar{\mathbf{P}}, \gamma, R)$ where S , γ , and R are defined as the state space, discount factor, and state rewards, and $\bar{\mathbf{P}}$ are a set of passive transition dynamics between the states (i.e., $\bar{\mathbf{P}}: S \times S \rightarrow [0, 1]$). For a discrete state space, $S = \{1, \dots, J\}$, $\bar{\mathbf{P}}$ is a $J \times J$ transition probability matrix with elements \bar{p}_{ij} for $i = 1, \dots, J$ and $j = 1, \dots, J$ corresponding to the transition from state i to state j .

The LMDP replaces the discrete action set, A , of an MDP with a set of continuous controls, $\mathbf{u} = (u_{ij})$ for $i = 1, \dots, J$ and $j = 1, \dots, J$, such that the controlled dynamics are expressed as:

$$p(s_t = j | s_{t-1} = i) = p_{ij}(u_{ij}) = \bar{p}_{ij} \exp(u_{ij}), \quad (4.1)$$

where $\{u_{ij} \in \mathbb{R}: \text{if } \bar{p}_{ij} = 0, \text{ then } u_{ij} = 0\}$. The controlled dynamics define the policy (e.g., how to choose the next state) of the LMDP. The controls, u_{ij} , are interpretable as the cost the controller is willing to pay to go against the passive dynamics (Todorov, 2009). For a given policy, the state control costs, $l(i, \mathbf{u})$, for a discrete state space LMDP are:

$$l(i, \mathbf{u}) = q(i) + KL(\mathbf{p}_i(\mathbf{u}) || \bar{\mathbf{p}}_i), \quad (4.2)$$

where $q(i) = -r(i)$ is the immediate state cost for states $i = 1, \dots, J$ where $r(i)$ is an element of the state reward vector, R , and $KL(\cdot)$ is the Kullback-Leibler (KL) divergence between the controlled transition probability, $\mathbf{p}_i(\mathbf{u}) = (p_{i1}(u_{i1}), \dots, p_{iJ}(u_{iJ}))'$, and passive transition probabilities, $\bar{\mathbf{p}}_i = (\bar{p}_{i1}, \dots, \bar{p}_{iJ})'$. The KL divergence penalty requires the agent to "pay" a larger price for behavior that strays further from the passive dynamics (Todorov, 2007).

The state costs-to-go, v_i for $i = 1, \dots, J$, are the discounted sum of future expected costs incurred from beginning in state i :

$$v_i = E[l(i, \mathbf{u}) + \gamma \sum_{t=1}^T l(j, \mathbf{u})] \quad (4.3)$$

where the expectation is with respect to the controlled transitions (4.1). The value of T determines whether the problem has finite- or infinite-horizon (e.g., $T < \infty$ or $T = \infty$). A finite-horizon LMDP can be modeled as an infinite-horizon LMDP by assuming the agent remains in the final observed state and incurs no future costs (Todorov, 2007). Costs-to-go can also be interpreted as relative time to goal completion where a smaller cost-to-go indicates that the agent can reach a desirable state more quickly by transitioning to that state than transitioning to a state with a higher cost-to-go. The Bellman optimality for the state

costs-to-go, v_i for $i = 1, \dots, J$, of an LMDP is:

$$v_i = \min_{\mathbf{u}} \left(l(i, \mathbf{u}) + \gamma \sum_{j=1}^J p_{ij}(u_{ij}) v_j \right), \quad (4.4)$$

where the summation is over the reachable states $j \in S$ as determined by the policy $p_{ij}(u_{ij})$ for all $j \in S$. The computational advantage of the LMDP is the Bellman optimality can be solved analytically using the method of Lagrange multipliers for the optimal transition probabilities:

$$p^*(s_t = j | s_{t-1} = i) = \frac{\bar{p}_{ij} \exp(-\gamma v_j)}{\sum_{k=1}^J \bar{p}_{ik} \exp(-\gamma v_k)}. \quad (4.5)$$

By substituting equation (4.5) into the Bellman optimality (4.4), the optimal costs-to-go are a solution to an eigenvector problem which can be solved using a power iteration method (Todorov, 2009), which we demonstrate in Section 4.3.

4.2.2 Inverse Reinforcement Learning (IRL)

Assume we observe a collection of sequences of optimal behavioral state trajectories, $\mathcal{S}_{nT_n} = \{s_{n0}, \dots, s_{nT}\}$, where s_{nt} is the observed state for individual n , $n = 1, \dots, N$, and time point t , $t = 0, 1, \dots, T$. Then, the observed state transitions are summarized into frequencies, $y_{ij} = \sum_n^N \sum_t^T I(s_{n,t} = j | s_{n,t-1} = i)$. We assume that each individual operates according to an LMDP with identical parameters, but that the state costs and therefore, costs-to-go, are

unknown. The likelihood of \mathcal{S} is:

$$\begin{aligned}
P(\mathcal{S}|\bar{\mathbf{P}}, \mathbf{v}) &= \prod_{n=1}^N \prod_{t=1}^T \prod_{i=1}^J \prod_{j=1}^J p^*(s_{n,t} = j | s_{n,t-1} = i), \\
&= \prod_{i=1}^J \prod_{j=1}^J \left(\frac{\bar{p}_{ij} \exp(-\gamma v_j)}{\sum_k \bar{p}_{ik} \exp(-\gamma v_k)} \right)^{y_{ij}},
\end{aligned} \tag{4.6}$$

for all individuals $n = 1, \dots, N$, times points $t = 0, 1, \dots, T$, and transitions from state $i \in S$ to state $j \in S$ and the second equality is based on the optimal transitions (4.5). We express the costs-to-go vector, $\mathbf{v} = (v_1, \dots, v_J)'$, as a linear combination of features in the $J \times n_b$ matrix \mathbf{X} with unknown weights $\boldsymbol{\beta}$ (e.g., $\mathbf{v} = \mathbf{X}\boldsymbol{\beta}$). We estimate the weights in a Bayesian framework by assuming the following hierarchical prior:

$$\begin{aligned}
\boldsymbol{\beta} &\sim \text{Normal}(0, 1/\tau), \\
\tau &\sim \text{Gamma}(0.1, 0.1),
\end{aligned} \tag{4.7}$$

and the parameters are estimated using MCMC sampling and variational approximation with the statistical platform STAN using the R package rstan (Carpenter et al., 2017; Stan Development Team, 2020). For the MCMC sampling, we used the Hamiltonian Monte Carlo with no-U-turn sampler (e.g, Hoffman and Gelman, 2014), which is the default algorithm. For variational inference, STAN assumes a Gaussian approximating distribution on a transformation of the parameters to a continuous domain (Kucukelbir et al., 2015). We provide brief definitions of the algorithms and the STAN modeling code in Appendix C.1. Note that the costs-to-go are only estimable up to a constant and therefore all resulting mean costs-to-go functions are shifted to have a minimum value of 0, which typically corresponds to a terminal state or a state in which an agent incurs no cost indefinitely (Todorov, 2009).

4.3 Vicsek LMDP

We illustrate the LMDP for collective movement using the Vicsek SPP model. The dynamics of the Vicsek model for agent $n = 1, \dots, N$ are defined for the direction $\theta_{n,t}$ and location $(x_{n,t}, y_{n,t})$ as:

$$\begin{aligned}\theta_{n,t+1} &= \langle \theta_{n,t} \rangle_\rho + \epsilon_{n,t}, \quad \epsilon_{n,t} \sim \mathcal{N}(0, \sigma^2), \\ x_{n,t+1} &= x_{n,t} + v_{n,t} * \cos(\theta_{n,t}), \\ y_{n,t+1} &= y_{n,t} + v_{n,t} * \sin(\theta_{n,t}),\end{aligned}\tag{4.8}$$

where the agent heads in the mean direction, $\langle \theta_{n,t} \rangle_\rho$, of other agents including itself within radius ρ with a speed of $v_{n,t}$. Šošić et al. (2017a) formulated the Vicsek model as an MDP with 13 discrete actions corresponding to turning angles $[-60^\circ, -50^\circ, \dots, 60^\circ]$, state observations of local misalignment, $s_{n,t} = \langle \theta_{n,t} \rangle_\rho - \theta_{n,t}$, and the optimal policy corresponds to the turning angle which minimizes the current local misalignment. In our simulation, we assume a constant velocity of 1, fixed interaction radius, $\rho = 0.1$, and turning angle standard deviation of 10 degrees, i.e., $\sigma = 10^\circ$.

We discretize the state space of local misalignment into 36 equally sized bins of 10 degrees and fix the state costs to:

$$q(s) = \begin{cases} 0 & \text{if } |s| \leq 5^\circ \\ 25 & \text{if } |s| \leq 15^\circ \\ 50 & \text{if } |s| \leq 25^\circ \\ 100 & \text{otherwise} \end{cases}.\tag{4.9}$$

where $s = 1, \dots, 36$ is the indicator of the discrete misalignment bins. The costs were chosen based on the results of Šošić et al. (2017a). The LMDP passive dynamics were

calculated by summing over the discrete actions. For the SPP model with asynchronous update, the change in local misalignment was assumed to be equivalent to the turning angle and therefore the passive dynamics are:

$$\bar{p}(s_t|s_{t-1}) = \sum_a p(\langle\theta_{n,t}\rangle_\rho|\langle\theta_{n,t-1}\rangle_\rho, a), \quad (4.10)$$

where the probabilities on the right are determined by a discretization of the normal CDF with $\sigma = 10^\circ$. Lastly, the true costs-to-go were calculated as the solution to the following system of equations:

$$\mathbf{z} = \mathbf{G}\bar{\mathbf{P}}\mathbf{z}, \quad (4.11)$$

where $\mathbf{z} = \exp(-\mathbf{v})$ is a J -dimensional vector referred to as the desirability function, $\mathbf{G} = \text{diag}(-q(\mathbf{s}))$, and $\bar{\mathbf{P}}$ is the $J \times J$ passive transition probability matrix. The system of equations is solved by initializing the vector \mathbf{z} to all ones, $\mathbf{z} = \mathbf{1}$, and repeatedly multiplying by $\mathbf{G}\bar{\mathbf{P}}$ until convergence. This method is referred to as Z -iteration in the LMDP literature (Todorov, 2009). The true costs-to-go appear to be a step function, where the steps represent states of equal costs and the steepness of the function between the steps encourages rapid transitions to the lowest cost state from nearby states (Figure 4.1).

We then simulated from the calculated optimal policy with 200 agents for 100 time points and calculated the state transition frequencies. We estimated the costs-to-go with full MCMC sampling and variational approximation. Additionally, we estimated the costs-to-go for each state separately, (e.g., $\mathbf{X} = \mathbf{I}$), and with Gaussian basis functions with centers on every other grid cell to reduce the state dimension by a factor of 2.

From Figure 4.1, it is evident that all modeling scenarios estimated the general trend of the true costs-to-go and the estimates from MCMC sampling have more uncertainty than

those from variational approximation. The costs estimated without basis function features exhibit a spike around -30° and 30° . This is due to a lack of observations to inform that grid cell. In general, all of the estimates appear to have underestimated the magnitude of the true costs. The underestimation is most likely due to the estimation being only possible up to an unknown constant, which may be larger in magnitude than the constant used (absolute value of the minimum of the means) to shift the functions to 0. It is likely that the many costs-to-go functions represented by the posterior distributions generate the same optimal behavior in the SPP as the costs subjectively chosen in (4.9).

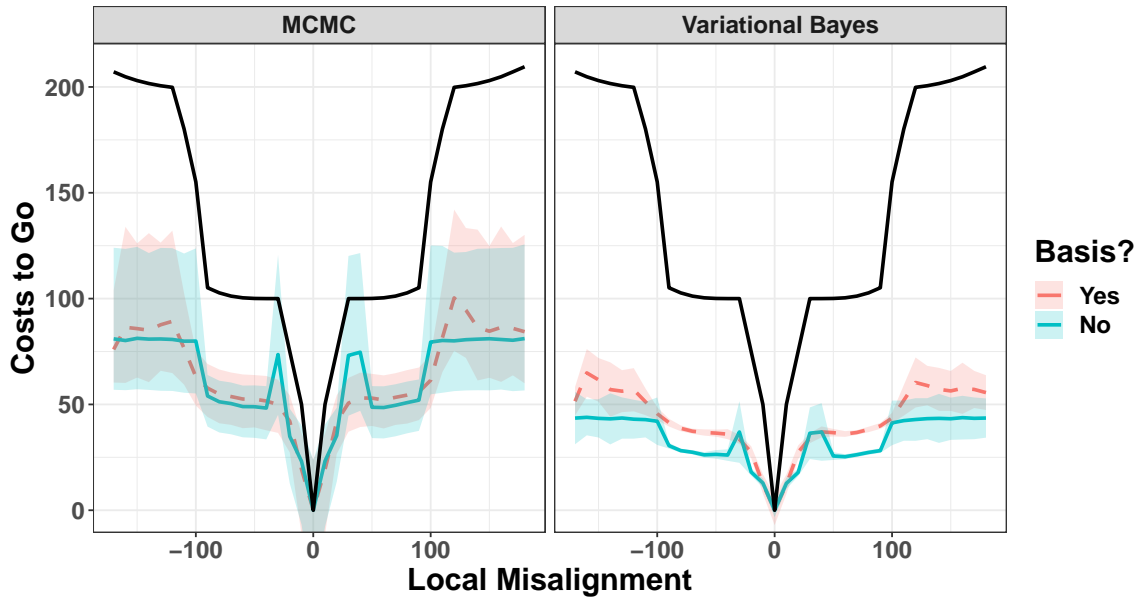


Fig. 4.1. Estimated costs-to-go for a Vicsek SPP model using LMDP IRL for Bayesian MCMC sampling and variational approximation under known passive dynamics. The models either used Gaussian basis functions (dashed lines) or independent state parameters (solid lines). The shaded regions correspond to the 95% C.I. The black curve is the true costs-to-go.

4.4 Guppy Application

We used the data available from Bode et al. (2012) on an experiment involving guppies (*Poecilia reticulata*) from a captive population. Groups of 10 same sex guppies were filmed from above in a square tank with one corner containing gravel and shade. The shaded corner provides shelter and is hypothesized to be attractive to the guppies. The guppies were released in the tank in the opposite corner. The data consists of movement trajectories clipped to start when all individuals were moving and to end when one guppy reached the shaded area. There were 26 experiments with 14 experiments consisting of all males and 12 of all females. Trajectories from all experiments were used to estimate the costs-to-go.

We defined an LMDP for the guppy trajectories by discretizing the state space of local misalignment as defined in Section 4.3 and target misalignment defined as the difference between the current heading and the direction to the target point. We rescaled all points to the unit square and calculated the local misalignment between an individual and all other individuals. The assumption of interaction with all other agents is reasonable as the movement was bounded and there were no visual obstructions outside the target area (Bode et al., 2012). The two misalignment states were discretized using the same 36 bins of length 10° as in the previous section resulting in a discretized grid of 36×36 . We assume a fixed discount factor of 1. Across the 26 experiments, there were 21,616 unique state transitions and an observation for almost every state. All estimation of parameters in the guppy application was done by variational approximation due to the size of the state space.

For the first set of estimated costs-to-go, we assumed the passive dynamics to be discrete uniform, (e.g., $\bar{p}_{ij} \propto 1$ for all $i, j = 1, \dots, 36$). The features, \mathbf{X} , considered were the identity matrix (referred to as "full" in Figure 4.2), radial basis functions with 50 centers estimated by k-means clustering, and 819 multiresolution bisquare basis functions generated uniformly

within the gridded state space by the R package FRK (Zammit-Mangion, 2020). The results shown in Figure 4.2 show a similar pattern across feature matrices, but the full estimation and bisquare basis functions appear to be more similar. In general, the results suggest the guppies perceived less cost for aligning with other guppies and there is more flexibility in target alignment as the low costs-to-go in yellow have more spread along the target misalignment axis.

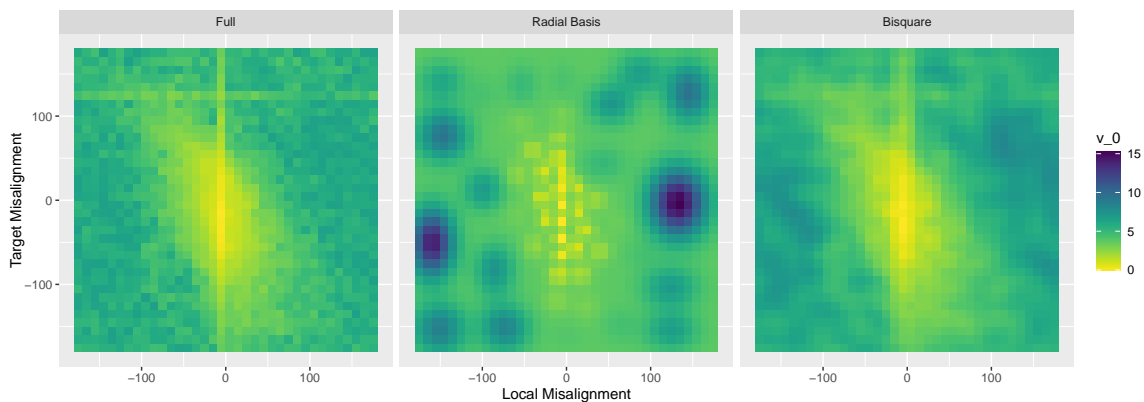


Fig. 4.2. Variational mean costs-to-go for the guppy experiments for a gridded state space of target and local misalignment across three sets of features: full (identity matrix), radial basis functions, and bisquare basis functions. The passive dynamics are assumed to be discrete uniform and the mean estimated costs have been shifted to have a minimum of 0. The yellow indicates states with lower costs-to-go and therefore states to which the guppies choose to transition.

Next, to assess the sensitivity to the assumed passive dynamics, we estimated the costs-

to-go under a set of passive dynamics corresponding to an independent, normal random walk on the gridded state space with standard deviation 90° . The standard deviation was chosen to be large enough to ensure all non-zero transition probabilities to use all of the observed data. The variational mean and standard deviation for the costs-to-go can be shown in Figure 4.3. Comparing to the previously estimated states, there appears to be more contrast between the costs-to-go for states along the diagonal than the off-diagonal. This sharper contrast is due to the higher action costs from the KL divergence penalty for transitions to states farther from the current state. For example, if a guppy is in a state with high costs-to-go, such as local and target misalignment around 100° , it will pay a larger action cost to transition against its passive dynamics to get to a lower costs-to-go state such as target and local misalignment of 0° than in the discrete uniform case. The standard deviations reflect the pattern of observed frequencies with states more frequently observed having smaller uncertainty.

The marginal costs-to-go estimated in Figure 4.3 are shown in Figure 4.5 where the costs-to-go are calculated as the mean across all values of the other state variable and shifted to have a minimum of 0. Interestingly, there appears to be evidence of collective alignment as shown in the local misalignment costs-to-go function due to the minimum cost occurring at 0° with gradual increase as misalignment increases in absolute value, but the function appears to be more smooth than the costs-to-go for the SPP model. Across the estimated costs-to-go functions there is a dip in costs-to-go for target alignment for bin $125 - 135^\circ$, which is due to the assigning of a heading of 0° to agents with a step length of 0. This suggests a better assumption would be to carry forward the turning angle from the last time period when the step length was greater than 0. Aside from this observed dip, the next least costly target alignment state is about -50° suggesting it is less costly to

VB Mean Costs-to-Go and Std. Dev.

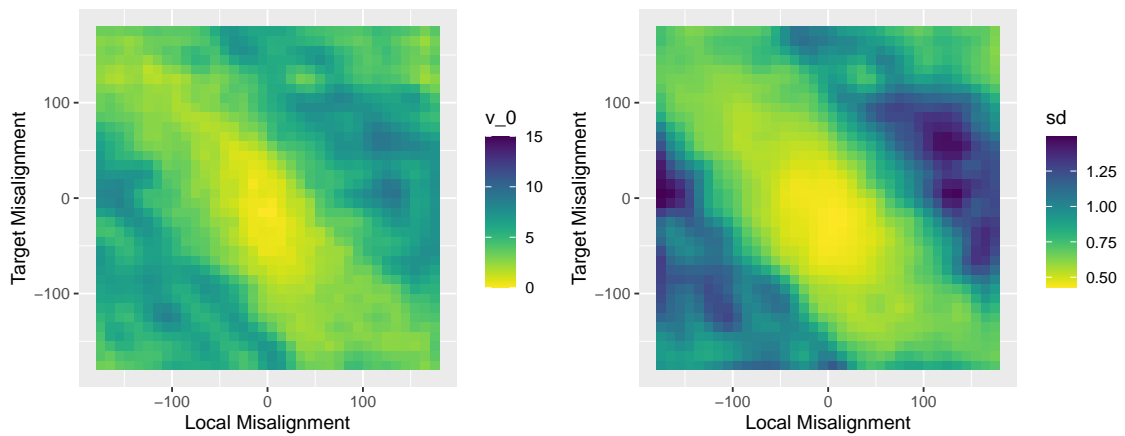


Fig. 4.3. Variational mean costs-to-go (left panel) and standard deviations (right panel) for the guppy experiments for a gridded state space of target and local misalignment with passive dynamics assumed to be a normal random walk and bisquare basis functions. The mean estimated costs-to-go have been shifted to have a minimum of 0. The yellow indicates states with lower costs-to-go and therefore states to which the guppies choose to transition.

approach the upper corner with the target 50° to the right. From inspection of the observed data found in Figure ??, it appears many of the guppies moved across the tank to the left first, which would require a right turn to decrease the target misalignment. A symmetry constraint could be applied to the costs-to-go by considering the absolute target alignment if it were assumed to be equally costly to approach from the right or left.

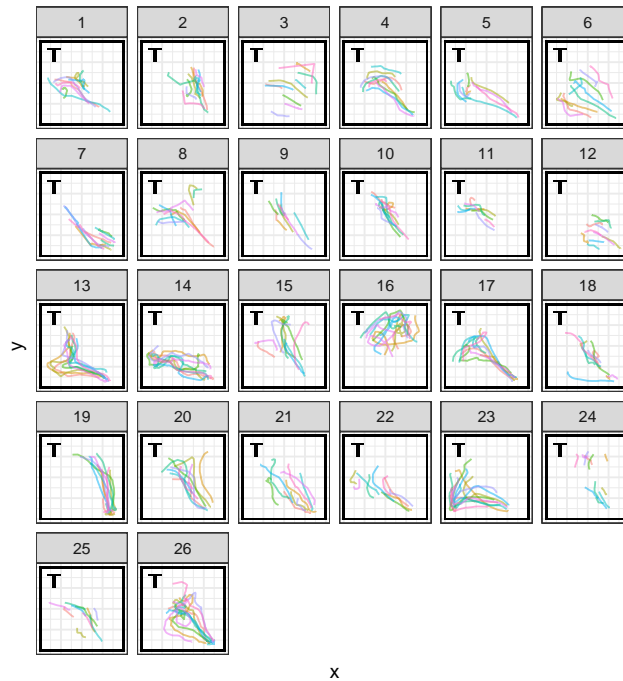


Fig. 4.4. Trajectories of all 26 experiments of groups of 10 guppies in a tank. The target is located at the point marked "T".

In Figures 4.2 and 4.3, there appears to be a diagonal pattern, which can be attributed to a couple of factors. First, it is likely the dip in target misalignment costs is a contributing factor, because $125 - 135^\circ$ target misalignment has low costs any trajectories transitioning through this state to the lowest cost states around -20° target and 0° local misalignment will have lower costs-to-go. Additionally, the corners appear far when plotted in the 2-D plane, but are close together in circular space so they have similar costs-to-go.

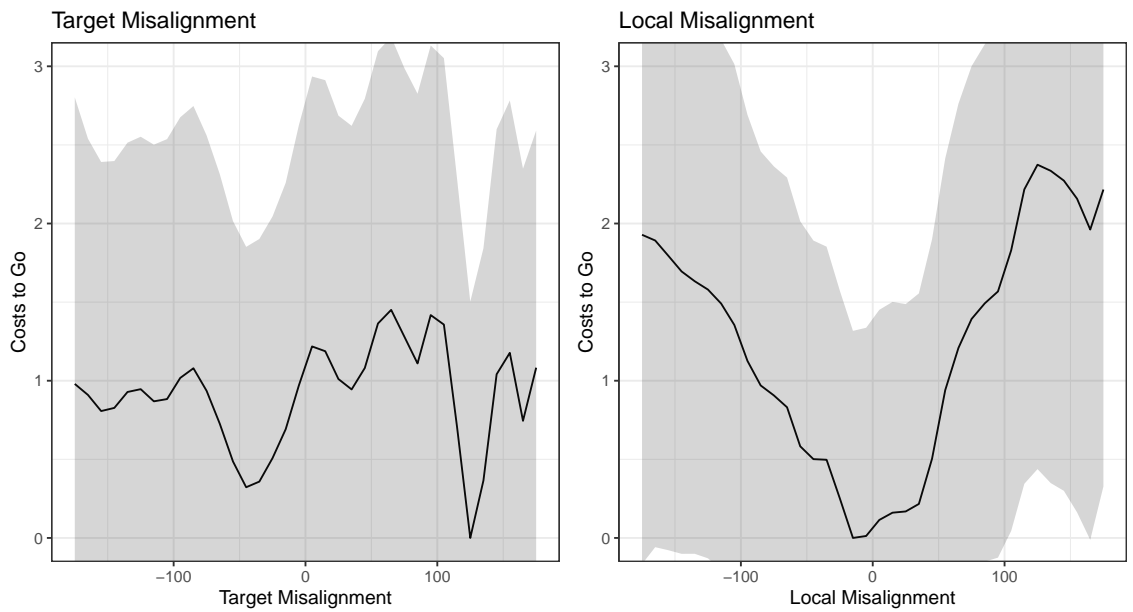


Fig. 4.5. Marginal costs-to-go of target and local misalignment for the guppy experiments for a gridded state space of target and local misalignment with passive dynamics assumed to be a normal random walk and bisquare basis functions. The mean estimated costs have been shifted to have a minimum of 0. The lower costs-to-go increase the probability of transitioning into that state.

4.5 Discussion

Collective motion from generative local interaction rules limit possible behavior, but the (L)MDP framework extends the definition of the agent to include perception and internal processes (Ried et al., 2019). By estimating the state costs-to-go or value functions, system specific local rules can be estimated.

As was shown in the estimation of the costs-to-go for the SPP model, it is difficult to estimate the true absolute values of the costs. Therefore, inference is constrained to relative differences in costs. This is similar to the estimation of relative selection probabilities in animal resource selection modeling (Hooten et al., 2017) and therefore IRL can still provide useful inference.

We can improve the inference for the guppy data by relaxing the assumptions, estimating passive dynamics, and expanding the state space to include other features. A meaningful passive dynamics model is necessary to make correct inference on cumulative costs-to-go due to the KL penalty. In future work, estimation of the passive dynamics parameters such as the random walk variance may be more appropriate. Additionally, the state space should include features based on physical distance and not simply heading to assess hypotheses about zonal collective movement which is a primary feature of collective movement ABMs (e.g., Couzin et al., 2002).

In the SPP simulation and guppy application, we assumed a discount factor of 1 which may be realistic for trajectories from such a short time frame. For observations from longer periods of time, it would be more realistic to assume there is some "forgetting" of past states which would correspond to a discount factor less than 1. Additionally, the discount factor can also be interpreted as the degree to which agents behave optimally (Choi and Kim, 2014). It might be expected that observations from animals in the wild are subject to more

stochasticity than experimental settings and therefore do not always behave optimally.

4.5.1 LMDP and Animal Movement Modeling

As a final discussion point, it is worth pointing out the similarities between the LMDP framework and traditional step selection models used in animal movement modeling (Hooten et al., 2017). Using the notation of Hooten et al. (2017), the step selection function is defined as:

$$[\boldsymbol{\mu}_t | \boldsymbol{\mu}_{t-1}, \boldsymbol{\beta}, \Delta_t, \boldsymbol{\theta}] \equiv \frac{g(\mathbf{x}(\boldsymbol{\mu}_t), \boldsymbol{\beta}) f(\boldsymbol{\mu}_t | \boldsymbol{\mu}_{t-1}, \Delta_t, \boldsymbol{\theta})}{\int_{\boldsymbol{\mu}} g(\mathbf{x}(\boldsymbol{\mu}), \boldsymbol{\beta}) f(\boldsymbol{\mu} | \boldsymbol{\mu}_{t-1}, \Delta_t, \boldsymbol{\theta}) d\boldsymbol{\mu}}, \quad (4.12)$$

where $g()$ is referred to as the selection function and $f()$ is the availability function. If we take the selection function as the exponential, assume regular discrete time, and discretize the state space, we have

$$[s_t | s_{t-1}, \boldsymbol{\beta}, \boldsymbol{\theta}] \equiv \frac{\exp(\mathbf{x}(s_t)' \boldsymbol{\beta}) f(s_t | s_{t-1}, \boldsymbol{\theta})}{\sum_s \exp(\mathbf{x}(s)' \boldsymbol{\beta}) f(s | s_{t-1}, \boldsymbol{\theta})}, \quad (4.13)$$

where s_t refers to the grid location in space, then we can see the likelihood is analogous to the form of (4.5). Therefore, a step selection analysis with the aforementioned assumptions is a LMDP with discretized spatial grid state space, passive dynamics governed by the availability function, $f()$, and state costs-to-go defined as $-\mathbf{x}(s)' \boldsymbol{\beta}$. There appears to be a lack of identifiability of the discount factor. However, if we assume the discount factor is 1, then all future states of the trajectory contribute to the state cost. In other words, the behavior of an animal moving under the dynamics of (4.13) are assuming the agent “never forgets” the costs incurred from past state visitations (Ried et al., 2019). Future work on selection functions in an LMDP framework with varying degrees of discount could provide novel insight into animal behavior. The discount factor may be estimable in systems where

multiple individuals are operating under the same long-term costs, but differ in length of memory. This may be individual differences, age differences, or seasonal differences.

Chapter 5

General Conclusions

We demonstrated the utility of linking alternative learning and statistical models for ecological inference of complex spatio-temporal processes. Chapter 1 highlighted the key features of the various methodology used for modeling the temporal evolution of such processes including a brief introduction to reinforcement learning (RL) with an illustrative example of RL for collective animal movement.

Chapter 2 presented a framework for incorporating the uncertainty from machine learning classification in a behavioral transition matrix. It is the first comprehensive example of Pólya-Gamma latent variables and multiple imputation within a Markov model (MM). Similarly, Chapter 4 estimated the value associated with of state transitions in a linearly-solvable Markov decision process (LMDP) with inverse reinforcement learning (IRL). The LMDP framework separated the transition process into two components: the passive dynamics and the internal processing of environmental information. Therefore, the environmental effects in Chapter 4 are interpretable as an agent's internal value system. Chapter 4 presented the first real data application of IRL with an LMDP for collective animal movement.

One benefit of the IRL framework of Chapter 4 is the insight into the mechanisms of animal decision-making. Similarly, Chapter 3 provided learning of the temporal evolution of an ecological process without *a priori* assumptions of the underlying mechanisms. The mechanisms were approximated by a deep echo state network (ESN). The ESN efficiently provided a nonlinear, stochastic basis expansion with long memory of the system by random weight initialization. The application to ecological forecasting of mallard settling patterns was the first example of the deep ensemble ESN (D-EESN) and Bayesian D-EESN with a generalized linear model output in a machine learning application.

Overall, advances in ecological modeling should continue to provide meaningful inference and predictions with full uncertainty quantification. We provided examples and guidance of exploiting the advantages of alternative learning for complex and nonlinear systems with statistical uncertainty quantification. Collaborative efforts among ecologists, statisticians, and computer scientists are key for continuing to advance learning the dynamics of ecological processes.

Appendix A

Supplement for Chapter 2

A.1 CropScape Habitat Categories

We combined CropScape categories obtained from the U.S. Department of Agriculture (CropScape; <https://nassgeodata.gmu.edu/CropScape/>) into fewer categories to reduce the number of estimable habitat coefficients. Table A.1 provides details on which categories were combined along with the frequency of each category associated with the ACC fixes.

Table A.1. Habitat categories used in the Bayesian Markov model with the associated CropScape categories from the USDA. CropScape categories were matched to GPS data for six greater white-fronted geese during March 2018.

Habitat Category	CropScape Categories	Frequency
Corn	Corn, Pop or Orn Corn	12852
Developed	Developed\High Intensity, Developed\Med Intensity, Developed\Low Intensity, Developed\Open Space	306
Fallow\Idle Cropland	Fallow\Idle Cropland	346
Grasses	Alfalfa, Grass\Pasture, Other Hay\Non Alfalfa, Switchgrass	7557
Herbaceous Wetlands	Herbaceous Wetlands	2242
Open Water	Open Water	12874
Other Crops	Dry Beans, Millet, Oats, Peas, Rye, Sorghum, Sugarbeets, Sunflower	760
Soybeans	Soybeans	5447
Wheat	Spring Wheat, Winter Wheat	571
Woody Wetlands	Woody Wetlands	589

A.2 Acceleration Classification

The training data consisted of 481 raw acceleration fixes with “ground-truthed” behavior classifications of flight(150), feeding (106), stationary (150), and walking (75). Each raw acceleration (m/s^2) fix was a multivariate time series collected at 10Hz for 3 seconds (i.e. 30 observations of acceleration in the x, y, and z direction). The fixes were summarized into the 52 features described in the appendix of Resheff et al. (2014) by using the associated web tool (<http://smell.huji.ac.il>). Additionally, the web tool fits 8 models: KNN, linear SVM, RBF kernel SVM, decision tree, random forest, naïve Bayes, LDA, and QDA. From the default settings, random forest performed the best and was chosen as the final classification method. The final random forest model was tuned offline in R. The number of trees was set to 500 (i.e., the default in R package randomForest). The number of variables to try was varied from 1 to 15 and chosen by repeating 10-fold cross validation 3 times. The final model chosen used 4 variables and had 96.5% accuracy on the training data set acquired by video recording. The top five variables based on importance as measured by the mean Gini index are the wave amplitude of the acceleration in the y axis, overall dynamic body acceleration (ODBA), dynamic body acceleration (DBA) in the x axis, the 25th percentile of the acceleration in the y axis, and the mean of the acceleration in the y axis (Figure A.2).

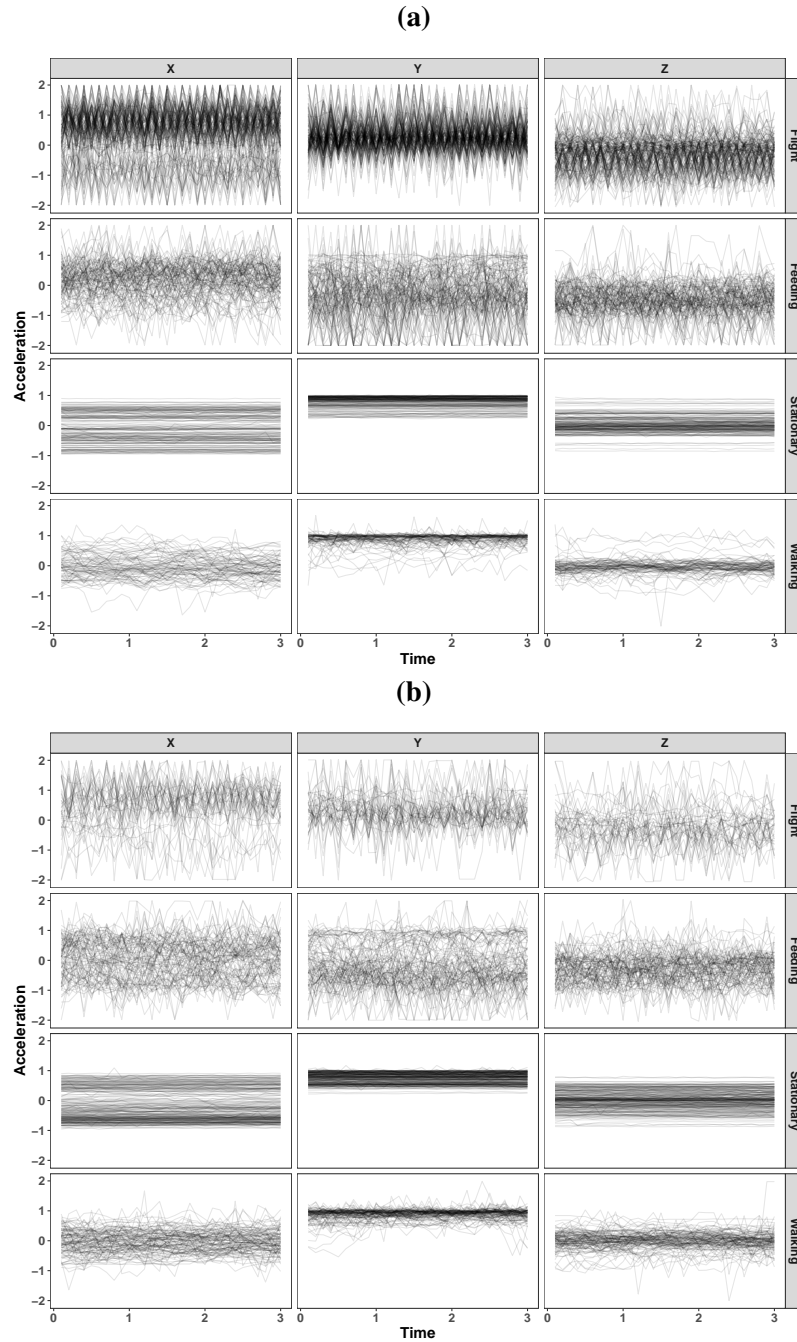


Fig. A.1. Raw acceleration time series for (a) the 481 training observations in the X, Y, and Z axis for the four behavior categories flight (150), feeding (106), stationary (150), and walking (75) and (b) a sample of 1% of the classified observations. In (b), the behaviors correspond to the most likely category based on the random forest predicted probabilities.

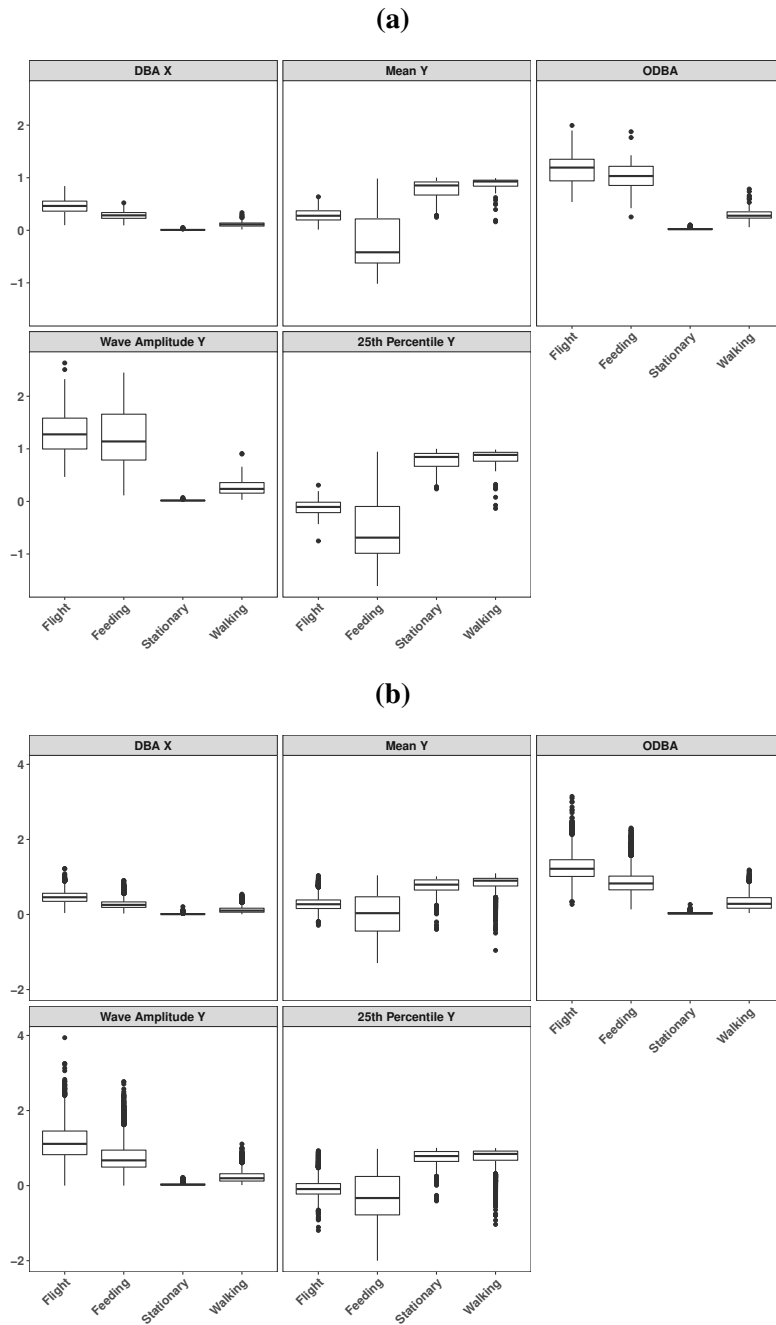


Fig. A.2. Box plots for the top five variables used in the classification random forest for the ground-truthed behaviors (a) and the classified observations (b). In (b), the behaviors correspond to the most likely category based on the random forest predicted probabilities.

A.3 Conditional Distributions

Our model contains J separate Bayesian multinomial logistic regressions with $T_i, i = 1, \dots, J$, observations where T_i is the number of transitions from category i . The covariate matrix, \mathbf{X}_i , is $T_i \times B$, with rows \mathbf{x}_{nt} and response vector \mathbf{Y}_{ij} is a $T_i \times 1$ vector with elements y_{nijt} where t corresponds to a transition from state i . The Gibbs updating steps are then:

$$\begin{aligned}\boldsymbol{\beta}_{ij} | \boldsymbol{\Omega}_{ij} &\sim \text{N}(\mathbf{m}_{ij}, \mathbf{V}_{ij}), \\ \omega_{nijt} | \boldsymbol{\beta}_{ij} &\sim \text{PG}(1, \eta_{nijt}),\end{aligned}\tag{A.1}$$

where $\mathbf{V}_{ij}^{-1} = \mathbf{X}_i' \boldsymbol{\Omega}_{ij} \mathbf{X}_i + \mathbf{V}_0^{-1}$ and $\mathbf{m}_{ij} = \mathbf{V}_{ij} (\mathbf{X}_i' (\mathbf{Y}_{ij} - 1/2) + \boldsymbol{\Omega}_{ij} \mathbf{C}_{ij}) + \mathbf{V}_0^{-1} \mathbf{m}_0$. Here, $\boldsymbol{\Omega}_{ij}$ is a $T_i \times T_i$ diagonal matrix with ω_{nijt} along the diagonal and \mathbf{C}_{ij} is a $T_i \times 1$ vector with elements C_{nijt} where t indexes transitions from category i . In this case, \mathbf{V}_0 and \mathbf{m}_0 correspond to the prior variance matrix and mean vector for $\boldsymbol{\beta}_{ij}$, respectively.

A.4 Model Validation

We validated the model by investigating the posterior distribution of the frequency of each behavior transition. For each iteration from the posterior distribution, we calculated the transition probabilities using the current values of the parameters and the covariates. Using the transition probabilities, we generated one sequence of behaviors for each individual. Then, we summarized each generated sequence into the behavior transition frequencies to build the posterior distribution. We also calculated the frequency of each transition for the 200 imputation data sets to build an "imputed" distribution. As shown in the figures on the next page, the posterior distribution covers the imputation distribution for every transition and individual suggesting the model is a good fit to the data.

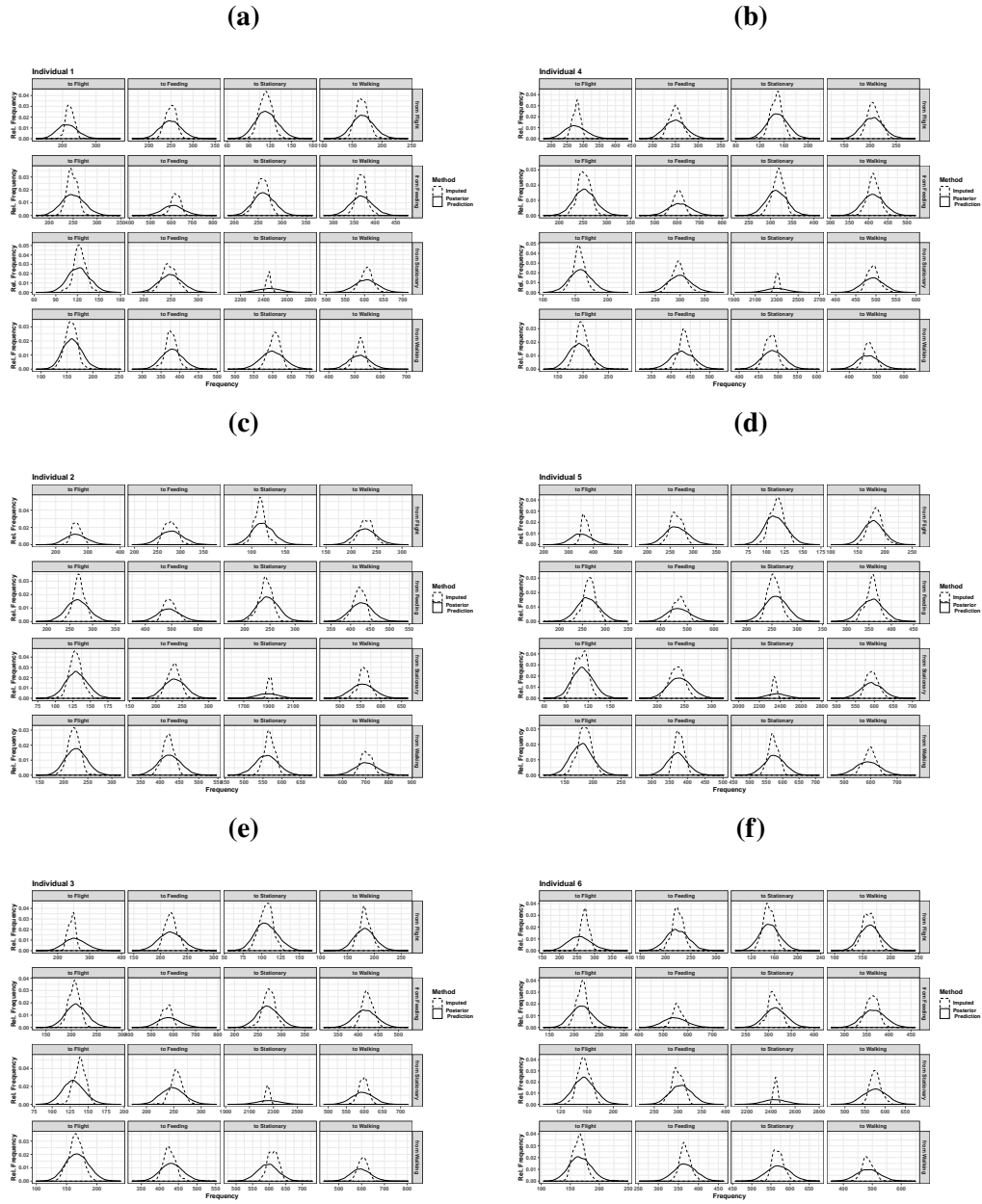
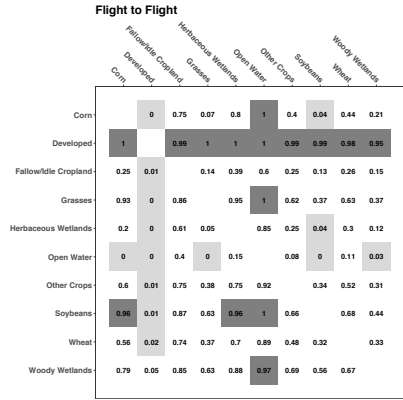


Fig. A.3. Comparisons of the imputed distribution of transition frequencies based on 200 imputation data sets and the posterior distribution of transition frequencies for all behavior transitions for six greater white-fronted geese in March 2018. The posterior covers the imputed distribution in all cases.

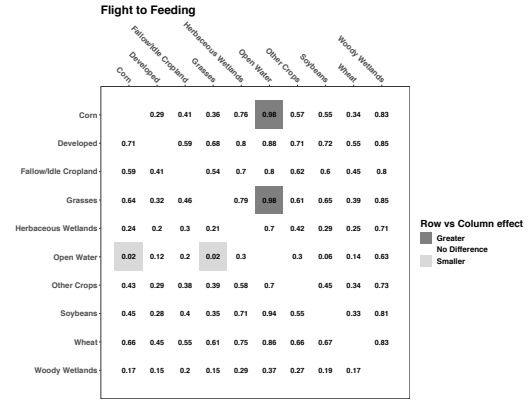
A.5 Habitat Multiple Comparison Matrices

The following matrices of pairwise comparisons between habitat coefficients visualize the proportion of samples from the posterior distribution in which the habitat coefficients differed in magnitude. The values indicate the proportion of samples in which the habitat coefficient down the row was greater than the coefficients along the column; the upper triangular values and lower triangular values sum to 1. For example in the first matrix, in the first row, corn, the proportion of times the estimate for the corn habitat coefficient was greater than the open water habitat coefficient was 1. The transition from grazing to flight was the only transition to not have any significant pairwise differences among habitats.

(a)

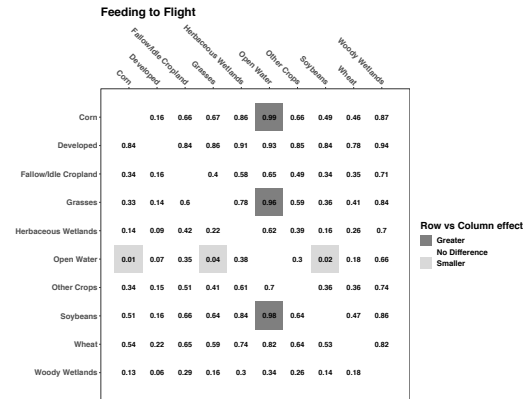
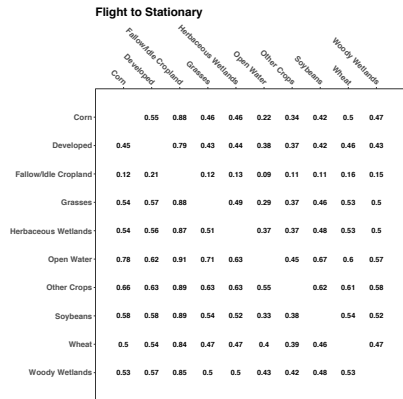


(b)



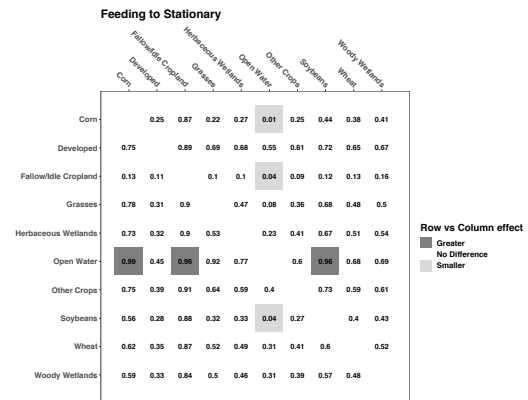
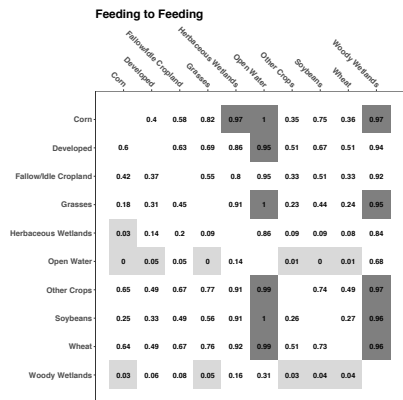
(c)

(d)



(e)

(f)



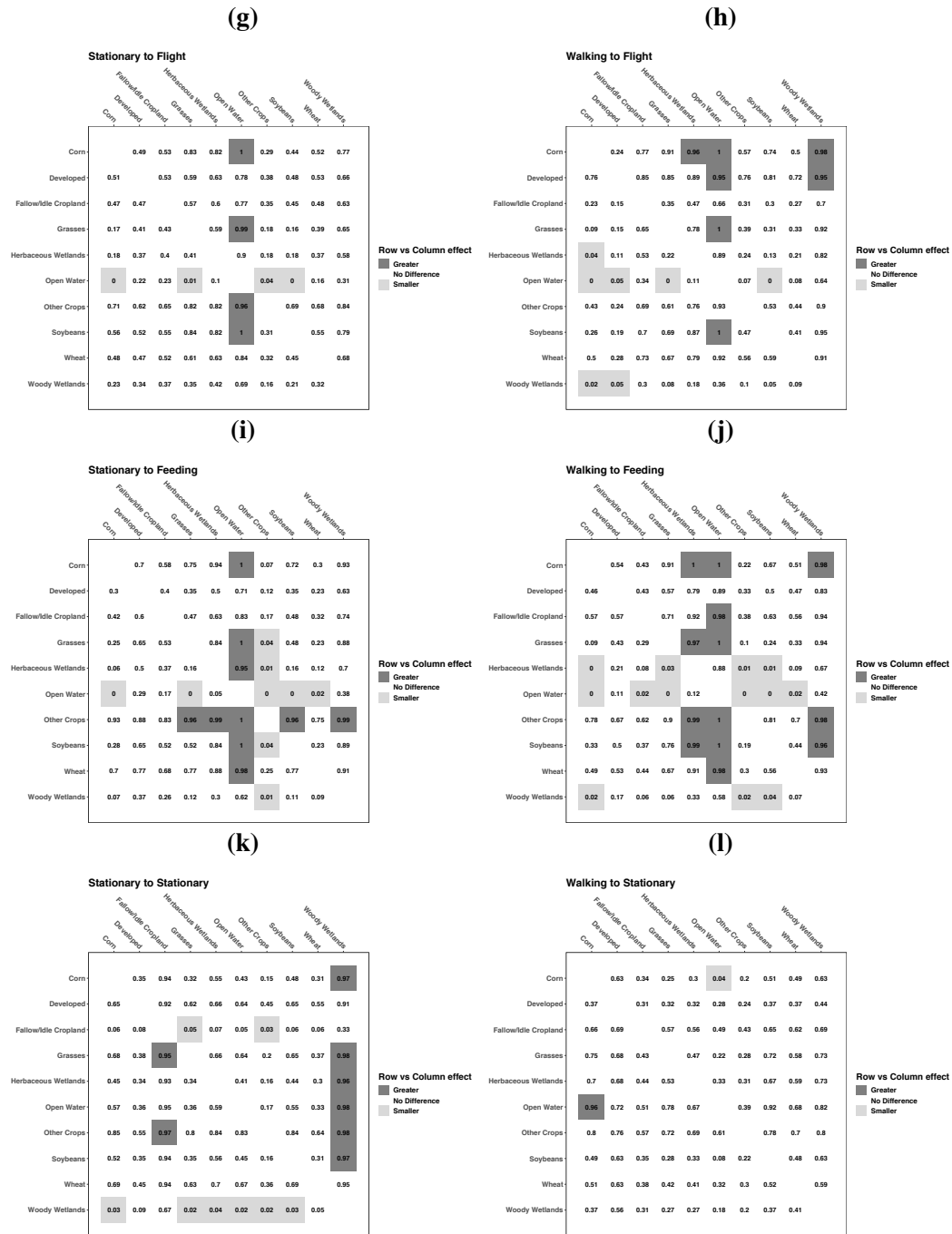


Fig. A.4. Matrix of pairwise comparisons between habitat coefficients for the all behavior transitions with estimated coefficients for six greater white-fronted geese in March 2018. The values indicate the proportion of posterior MCMC iterations in which the habitat coefficient down the row was greater than the coefficients along the column; the upper triangular values and lower triangular values sum to 1.

A.6 Sensitivity of Estimates

We assessed the sensitivity of estimates to the choice of number of imputation data sets. The following figures show the 95% credible intervals from the posterior distribution based on 200 imputation data sets, 100 imputation data sets, and 1 data set corresponding to the most likely behavior classification (i.e., category with largest predicted probability based on the random forest). We see general consensus among the three scenarios with a general shrinking of estimates to zero with multiple imputation compared to most likely behaviors (Figure 2.2 and Figure 2.3). Credible interval widths for the imputation scenarios appear to be consistent or larger than the widths for coefficients estimated from the most likely behaviors.

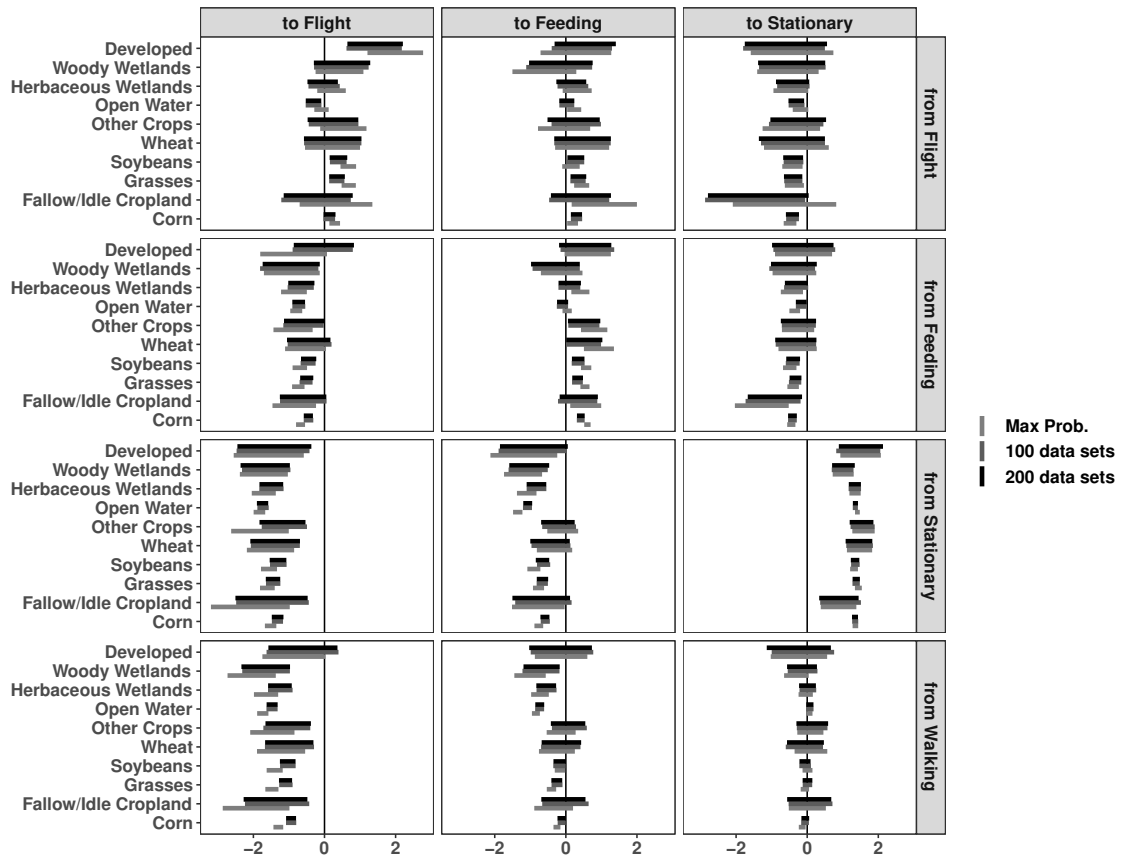


Fig. A.5. 95% credible intervals for habitat effects on the log-odds estimated in a Bayesian framework with either the behavior class with the largest classification probability (Max Prob.), 100 imputation data sets, or 200 imputation data sets by behavior transition for six greater white-fronted geese in March 2018.

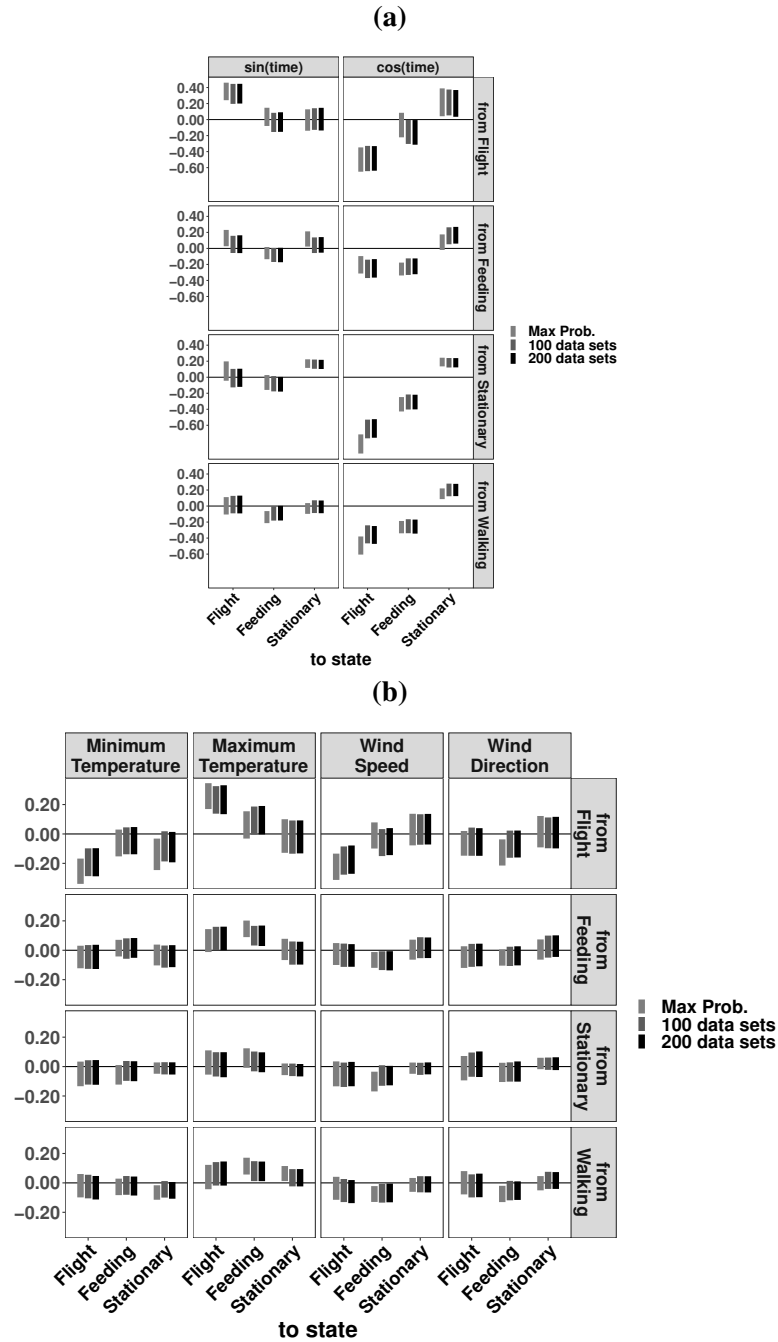


Fig. A.6. 95% credible intervals for diurnal (a) and weather (b) covariate effects on the log-odds estimated in a Bayesian framework with either the most likely behavior class (Max Prob.), 100 imputation data sets, or 200 imputation data sets by behavior transition for six greater white-fronted geese in March 2018.

Appendix B

Supplement for Chapter 3

B.1 Hyperparameters

The following table lists the hyperparameters associated with the D-EESN for fixed number of layers, L . The total number of parameters we optimized for a model of depth L was $3 * L + 1$.

Table B.1. A table of hyperparameters for the D-EESN. For all layer specific hyperparameters, values were tuned for each layer unless otherwise noted.

Hyperparameter	State	Description	Search space/fixed value
λ	Optimized	Shrinkage penalty parameter	(0, 0.01)
ζ_l	Optimized	Leaking rate	(0, 1)
ν_l	Optimized	Recurrent scaling parameter	(0, 1)
$n_{\tilde{h},l}$	Optimized	Number of reduced features for feedforward connections	{25, ..., 75}
$n_{h,1}$	Optimized	Size of reservoir 1	{6, ..., 20}
$n_{h,l}$ for $l \geq 2$	Fixed	Size of the deep reservoirs	115
$a_{w,l}$	Fixed	Magnitude of non-zero recurrent weights, \mathbf{W}_l	0.1
$a_{u,l}$	Fixed	Magnitude of non-zero feedforward weights, \mathbf{U}_l	0.1
$\pi_{w,l}$	Fixed	Probability of a non-zero recurrent weight, \mathbf{W}_l	0.1
$\pi_{u,l}$	Fixed	Probability of a non-zero feedforward weight, \mathbf{U}_l	0.1

B.2 MCMC Distributions

Below are the derivations of distributions used in the MCMC sampling of the Poisson BD-EESN. Let the total number of features from the ensemble of ESNs be $n_H = n_{h,1} + \sum_{l=2}^L n_{\tilde{h},l}$. We use the following notations for combining the coefficient matrices into a $n_H \times n_b$ -dimensional matrix and the hidden units into an n_H -dimensional row vector:

$$\mathbf{B} = \begin{bmatrix} \boldsymbol{\beta}_1^{(1)'} \\ \vdots \\ \boldsymbol{\beta}_L^{(1)'} \\ \vdots \\ \boldsymbol{\beta}_1^{(n_{res})'} \\ \vdots \\ \boldsymbol{\beta}_L^{(n_{res})'} \end{bmatrix}, \quad \mathbf{h}_t = \left[\mathbf{h}_{t,1}^{(1)'} \cdots \mathbf{h}_{t,L}^{(1)'} \cdots \mathbf{h}_{t,1}^{(n_{res})'} \cdots \mathbf{h}_{t,L}^{(n_{res})'} \right], \quad (\text{B.1})$$

and we have for each n_b -dimensional row vector of the log of the latent basis coefficients, $\log(\boldsymbol{\alpha}_t) \sim \text{MVN}(\boldsymbol{\beta}_0 + \mathbf{h}_t \mathbf{B}, \sigma^2 \mathbf{I}_{n_b})$. Additionally, define the $n_T \times n_b$ -dimensional matrix \mathbf{A} and the $n_T \times n_H$ -dimensional matrix \mathbf{H} as:

$$\mathbf{A} = \begin{bmatrix} \log(\boldsymbol{\alpha}_1) - \boldsymbol{\beta}_0 \\ \vdots \\ \log(\boldsymbol{\alpha}_T) - \boldsymbol{\beta}_0 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_T \end{bmatrix}. \quad (\text{B.2})$$

By properties of matrix normal, MN, since $\mathbf{A} = \mathbf{H}\mathbf{B} + \mathbf{E}$ where $\mathbf{E} \sim \text{MN}(\mathbf{0}, \mathbf{I}_{n_T}, \sigma^2 \mathbf{I}_{n_b})$, then by letting $\tilde{\mathbf{a}} = \text{vec}(\mathbf{A}')$, $\tilde{\mathbf{H}} = \mathbf{H} \otimes \mathbf{I}_{n_b}$, and $\boldsymbol{\beta} = \text{vec}(\mathbf{B}')$, $\tilde{\mathbf{a}} \sim \text{N}(\tilde{\mathbf{H}}\boldsymbol{\beta}, \sigma^2 \mathbf{I}_{n_T n_b})$ (Li et al., 2019).

B.2.1 Basis Coefficients

The latent basis coefficients were sampled by one-at-a-time MH on the log scale. The basis coefficients for MCMC iteration k were proposed by:

$$\log(\alpha_{b,t}^*) \sim \text{N}\left(\log\left(\alpha_{b,t}^{(k-1)}\right), \eta_{bt}\right), \quad (\text{B.3})$$

for each basis $b = 1, \dots, n_b$ and time point $t = 1, \dots, T$ with a MH step length of $\eta_{b,t}$. The step lengths were chosen by adaptive tuning (e.g., Andrieu and Thoms, 2008) with an intercept only model initialized at the NMF basis coefficients. Specifically, the step lengths were initialized to 1 then updated according to:

$$\log \eta_{b,t}^{(k)} = \log \eta_{b,t}^{(k-1)} + (r_{b,t}^{(k-1)} - r^*), \quad (\text{B.4})$$

where $r_{b,t}^{(k-1)}$ is the observed MH acceptance probability and r^* is the target acceptance ratio. We set the target acceptance ratio to 0.4 and after about 500 iterations the step lengths appeared to settle on appropriate values. We used the last values of the step lengths during all subsequent model fits.

The MH acceptance probability for the basis coefficients is $r_{b,t}^{(k)} = \min(a_{b,t}^{(k)}, 1)$ where:

$$a_{b,t}^{(k)} = \frac{p(\mathbf{Z}_t | \alpha_{b,t}^*, \boldsymbol{\alpha}_{(-b),t}^{(k-1)}) p(\log(\alpha_{b,t}^*) | \beta_{b,0}, \boldsymbol{\beta}, \sigma^2)}{p(\mathbf{Z}_t | \alpha_{b,t}^{(k-1)}, \boldsymbol{\alpha}_{(-b),t}^{(k-1)}) p(\log(\alpha_{b,t}^{(k-1)}) | \beta_{b,0}, \boldsymbol{\beta}, \sigma^2)}, \quad (\text{B.5})$$

with the first distribution being the Poisson data likelihood and the second the Normal regression prior on the log of the coefficients.

B.2.2 Intercept Vector

We assumed a Jeffreys prior for the intercept vector, $p(\boldsymbol{\beta}_0) \propto 1$. Therefore, the updating step for the intercept vector is Gibbs with the following conditional posterior:

$$[\boldsymbol{\beta}_0 | \cdot] \sim \text{MVN} \left(\frac{1}{T} \sum_{t=1}^T \log \alpha_t - \mathbf{h}_t \mathbf{B}, \frac{\sigma^2}{n_T} \mathbf{I}_{n_b} \right) \quad (\text{B.6})$$

B.2.3 Regression Coefficients and HS parameters

Following the notation of Johndrow et al. (2020), define $\xi = \tau^{-2}$ and $\eta_{l,b,k_l}^{(j)} = \lambda_{l,b,k_l}^{(j)-2}$:

$$\begin{aligned} \mathbf{D} &= \text{diag} \left(\eta_{l,b,k_l}^{(j)-1} \right), \quad \mathbf{M}_\xi = \mathbf{I}_{n_T n_H} + \xi^{-1} \tilde{\mathbf{H}} \mathbf{D} \tilde{\mathbf{H}}', \\ p(\xi | \boldsymbol{\eta}) &= |\mathbf{M}_\xi|^{-1/2} \left(\frac{1}{2} \tilde{\mathbf{a}}' \mathbf{M}_\xi^{-1} \tilde{\mathbf{a}} \right)^{(n_T n_H)/2} \frac{1}{\sqrt{\xi} (1 + \xi)}. \end{aligned} \quad (\text{B.7})$$

Johndrow et al. (2020) derived an exact rejection sampler for the (inverse) local shrinkage parameters instead of the traditional latent variable method of slice sampling. The update rules for targeting the exact horseshoe posterior using their proposed blocked Metropolis-within-Gibbs are:

1. sample $\boldsymbol{\eta} \sim p(\boldsymbol{\eta} | \tau, \boldsymbol{\beta}, \sigma^2) \propto \prod_{j=1}^{n_{res}} \prod_{b=1}^{n_b} \prod_{k_l=1}^{n_{\tilde{h},l}} \prod_{l=1}^L \frac{1}{1 + \eta_{l,b,k_l}^{(j)}} e^{-\frac{\xi \beta_{l,b,k_l}^{(j)2} \eta_{l,b,k_l}^{(j)}}{2\sigma^2}}$,
2. propose $\log \xi^* \sim \text{N}(\log \xi, s)$ and accept with probability $\frac{p(\xi^* | \boldsymbol{\eta}) \xi^*}{p(\xi | \boldsymbol{\eta}) \xi}$,
3. sample $\sigma^2 | \boldsymbol{\eta}, \xi \sim \text{InvGamma}(\frac{n_T n_H}{2}, \frac{\tilde{\mathbf{a}}' \mathbf{M}_\xi^{-1} \tilde{\mathbf{a}}}{2})$,
4. sample $\boldsymbol{\beta} | \boldsymbol{\eta}, \xi, \sigma^2 \sim \text{N}((\tilde{\mathbf{H}}' \tilde{\mathbf{H}} + (\xi^{-1} \mathbf{D})^{-1})^{-1} \tilde{\mathbf{H}}' \tilde{\mathbf{a}}, \sigma^2 (\tilde{\mathbf{H}}' \tilde{\mathbf{H}} + (\xi^{-1} \mathbf{D})^{-1})^{-1})$.

We used the MH step size, s , of 0.8 for proposing ξ as recommended by Johndrow et al. (2020). In large problems, the computation of the inverse conditional variance-covariance matrix is very expensive. We used the computational technique for sampling from a large multivariate Normal proposed by (Bhattacharya et al., 2016). The steps of the algorithm are:

1. sample $\mathbf{u} \sim \text{N}(0, \xi^{-1} \mathbf{D})$ and $\mathbf{f} \sim \text{N}(0, \mathbf{I}_{n_T n_H})$ independently,
2. set $\mathbf{v} = \tilde{\mathbf{H}} \mathbf{u} + \mathbf{f}$ and compute $\mathbf{v}^* = \mathbf{M}_\xi^{-1}(\tilde{\mathbf{a}} - \mathbf{v})$,
3. set $\boldsymbol{\beta} = \sigma(\mathbf{u} + \xi^{-1} \mathbf{D} \tilde{\mathbf{H}}' \mathbf{v}^*)$.

Appendix C

Supplement for Chapter 4

C.1 STAN Algorithms and Code

We implemented the MCMC sampling and variational approximation for Bayesian inference in Chapter 4 using the modeling program STAN (Carpenter et al., 2017).

The default MCMC sampling algorithm used by STAN is the Hamiltonian Monte Carlo (HMC) with no-U-Turn sampler (NUTS). The HMC algorithm is an efficient algorithm for proposing parameters for exploring the posterior distribution (Betancourt and Girolami, 2015). Then, the proposed parameters are accepted in a Metropolis update. The parameters are proposed by moving along the gradient of the conditional posterior and the size of the move is determined by auxiliary momentum variables. Inbetween proposals, the moves are repeated a fixed number of times (leapfrog steps) and the final value of the parameters are evaluated for acceptance. The NUTS provides automatic tuning of the number of leapfrog steps needed for adequate exploration of the posterior (Hoffman and Gelman, 2014). In

Chapter 4, we sampled four chains of 6000 iterations with the first 2000 discarded as burnin and checked diagnostics using the tools provided by the R package rstan (Stan Development Team, 2020)

Variational Bayesian inference is an optimization problem that minimizes the Kullback-Leibler divergence between an approximate posterior distribution and the true posterior distribution (Tzikas et al., 2008). The true posterior distribution is unknown and thus, the KL divergence cannot be computed exactly but it can be bounded. This bound is referred to as the evidence lower bound (ELBO). The parameters of the approximate posterior distribution are updated iteratively using stochastic gradient ascent until the change in the ELBO is sufficiently small. STAN implements the algorithm of Kucukelbir et al. (2015) and stops when the change in ELBO is less than 0.01.

C.1.1 STAN Model Code

The following code is used for both the MCMC sampling and variational inference with STAN.

```
data {
  int<lower=1> N; // number of states
  int<lower=1> Nv; // number of v parameters
  int<lower=1> Ni; // number of individuals
  int<lower=1> Nobs; // number of observations of n_ijk
  int<lower=1> Nb; //number of basis functions
  int<lower=1> maxcount; // max frequency of counts
  real<lower=0, upper=1> discount; // MDP discount factor
  int<lower=0,upper=maxcount> n[Nobs]; // vector of counts
  int<lower=1,upper=N> J[Nobs]; // vector of from state
    indices
  int<lower=1,upper=N> K[Nobs]; // vector of to state
    indices
  int<lower=1,upper=Ni> I[Nobs]; // vector of individual
    indices
  matrix[Nv,Nb] Fmat; // matrix of features, if identity
    than slopes model
  row_vector<lower=0,upper=1>[N] Aw[Nv]; //adjacency matrix
    of transition indices map for w
}
parameters {
```

```

    vector[Nb] w; // basis weight parameters
    real<lower=0> alpha; // precision
}
transformed parameters{
    vector[Nv] v;
    v = Fmat*w;
}
model {
    real tmp;
    alpha ~ gamma(0.1,0.1);
    w ~ normal(0,sqrt(1/alpha));
    for(i in 1:Nobs){
        tmp = Aw[J[i],]*exp(-discount*v);
        target += n[i]*(log(Aw[J[i],K[i]])+(-discount*v[K[i]])-
            log(tmp));
    }
}

```

Bibliography

- Albert, J. H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association*, 88(422):669–679.
- Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, 49(3-4):227–266.
- Andrieu, C. and Thoms, J. (2008). A tutorial on adaptive mcmc. *Statistics and computing*, 18(4):343–373.
- Arora, S. and Doshi, P. (2018). A survey of inverse reinforcement learning: Challenges, methods and progress. *arXiv preprint arXiv:1806.06877*.
- Badea, L. (2008). Extracting gene expression profiles common to colon and pancreatic adenocarcinoma using simultaneous nonnegative matrix factorization. In *Biocomputing 2008*, pages 267–278. World Scientific.
- Baldassarre, G. A. (2014). *Ducks, Geese, and Swans of North America*, volume 1. JHU Press.
- Beck, M. B. and van Straten, G. (2012). *Uncertainty and forecasting of water quality*. Springer Science & Business Media.

- Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2):157–166.
- Betancourt, M. and Girolami, M. (2015). Hamiltonian Monte Carlo for hierarchical models. *Current Trends in Bayesian Methodology with Applications*, 79(30):2–4.
- Bhadra, A., Datta, J., Polson, N. G., Willard, B., et al. (2019). Lasso meets horseshoe: A survey. *Statistical Science*, 34(3):405–427.
- Bhattacharya, A., Chakraborty, A., and Mallick, B. K. (2016). Fast sampling with Gaussian scale mixture priors in high-dimensional regression. *Biometrika*, 103(4):985–991.
- Bode, N. W., Franks, D. W., Wood, A. J., Piercy, J. J., Croft, D. P., and Codling, E. A. (2012). Distinguishing social from nonsocial navigation in moving animal groups. *The American Naturalist*, 179(5):621–632.
- Boutsidis, C. and Gallopoulos, E. (2008). SVD based initialization: A head start for nonnegative matrix factorization. *Pattern Recognition*, 41(4):1350–1362.
- Bradley, J. R., Wikle, C. K., and Holan, S. H. (2019). Spatio-temporal models for big multinomial data using the conditional multivariate logit-beta distribution. *Journal of Time Series Analysis*, 40(3):363–382.
- Breed, M. D. and Moore, J. (2015). *Animal Behavior*. Academic Press.
- Broekhuis, F., Grünewälder, S., McNutt, J. W., and Macdonald, D. W. (2014). Optimal hunting conditions drive circalunar behavior of a diurnal carnivore. *Behavioral Ecology*, 25(5):1268–1275.

- Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455.
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, 85(7):1771–1789.
- Buckland, S. T., Newman, K. B., Fernández, C., Thomas, L., and Harwood, J. (2007). Embedding population dynamics models in inference. *Statistical Science*, 22(1):44–58.
- Cantrell, R. and Cosner, C. (2004). *Spatial Ecology via Reaction-Diffusion Equations*. Wiley Series in Mathematical & Computational Biology. Wiley.
- Carmichael, Z., Syed, H., Burtner, S., and Kudithipudi, D. (2018). Mod-DeepESN: modular deep echo state network. *arXiv preprint arXiv:1808.00523*.
- Carmichael, Z., Syed, H., and Kudithipudi, D. (2019). Analysis of wide and deep echo state networks for multiscale spatiotemporal time series forecasting. In *Proceedings of the 7th Annual Neuro-Inspired Computational Elements Workshop, NICE '19*, New York, NY, USA. Association for Computing Machinery.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1).
- Carvalho, C. M., Polson, N. G., and Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480.
- Chakravarty, P., Cozzi, G., Ozgul, A., and Aminian, K. (2019). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, 10(6):802–814.

- Chen, J., Zhu, J., Wang, Z., Zheng, X., and Zhang, B. (2013). Scalable inference for logistic-normal topic models. In *Advances in Neural Information Processing Systems*, pages 2445–2453.
- Choi, J. and Kim, K.-E. (2011). Map inference for Bayesian inverse reinforcement learning. In *Advances in Neural Information Processing Systems*, pages 1989–1997.
- Choi, J. and Kim, K.-E. (2013). Bayesian nonparametric feature construction for inverse reinforcement learning. In *Twenty-Third International Joint Conference on Artificial Intelligence*.
- Choi, J. and Kim, K.-E. (2014). Hierarchical Bayesian inverse reinforcement learning. *IEEE Transactions on Cybernetics*, 45(4):793–805.
- Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., Lodge, D. M., Pascual, M., Pielke, R., Pizer, W., Pringle, C., Reid, W. V., Rose, K. A., Sala, O., Schlesinger, W. H., Wall, D. H., and Wear, D. (2001). Ecological forecasts: An emerging imperative. *Science*, 293(5530):657–660.
- Clark, J. S. and Gelfand, A. (2006). *Hierarchical Modelling for the Environmental Sciences: Statistical Methods and Applications*. Oxford University Press, Inc., USA.
- Cochran, M. E. and Ellner, S. (1992). Simple methods for calculating age-based life history parameters for stage-structured populations: Ecological archives m062-002. *Ecological monographs*, 62(3):345–364.
- 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):1–11.

- Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19(3):553–570.
- Cressie, N. and Wikle, C. K. (2011). *Statistics for Spatio-Temporal Data*. John Wiley & Sons, Hoboken, NJ.
- de Valpine, P. and Hastings, A. (2002). Fitting population models incorporating process noise and observation error. *Ecological Monographs*, 72(1):57–76.
- Dietze, M. C. (2017a). *Ecological Forecasting*. Princeton University Press.
- Dietze, M. C. (2017b). Prediction in ecology: A first-principles framework. *Ecological Applications*, 27(7):2048–2060.
- Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lurch, C. K., Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C., and White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*, 115(7):1424–1432.
- Dvijotham, K. and Todorov, E. (2010). Inverse Optimal Control with Linearly-Solvable MDPs. In *ICML*, pages 335–342.
- Ely, C. R. (1992). Time allocation by Greater white-fronted geese: Influence of diet, energy reserves and predation. *The Condor*, 94(4):857–870.
- Estrada-Peña, A. (2001). Forecasting habitat suitability for ticks and prevention of tick-borne diseases. *Veterinary Parasitology*, 98(1-3):111 – 132.

- Fehlmann, G., O’Riain, M. J., Hopkins, P. W., O’Sullivan, J., Holton, M. D., Shepard, E. L., and King, A. J. (2017). Identification of behaviours from accelerometer data in a wild social primate. *Animal Biotelemetry*, 5(1):6.
- Feldman, R. E., Anderson, M. G., Howerter, D. W., and Murray, D. L. (2015). Where does environmental stochasticity most influence population dynamics? an assessment along a regional core-periphery gradient for prairie breeding ducks. *Global Ecology and Biogeography*, 24(8):896–904.
- Fewster, R. M., Buckland, S. T., Siriwardena, G. M., Baillie, S. R., and Wilson, J. D. (2000). Analysis of population trends for farmland birds using generalized additive models. *Ecology*, 81(7):1970–1984.
- Fox, A. D., Glahder, C. M., and Walsh, A. J. (2003). Spring migration routes and timing of Greenland white-fronted geese—results from satellite telemetry. *Oikos*, 103(2):415–425.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22.
- Frühwirth-Schnatter, S. and Frühwirth, R. (2007). Auxiliary mixture sampling with applications to logistic models. *Computational Statistics & Data Analysis*, 51(7):3509–3528.
- Frühwirth-Schnatter, S. and Frühwirth, R. (2010). Data augmentation and MCMC for binary and multinomial logit models. In *Statistical Modelling and Regression Structures*, pages 111–132. Springer.
- Funahashi, K.-i. and Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6(6):801–806.

- Gaujoux, R. and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics*, 11(1):367.
- Gelman, A., Hill, J., and Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5(2):189–211.
- Gilks, W. R., Richardson, S., and Spiegelhalter, D. (1995). *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC.
- Glynn, C., Tokdar, S. T., Howard, B., Banks, D. L., et al. (2019). Bayesian analysis of dynamic linear topic models. *Bayesian Analysis*, 14(1):53–80.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- Grimm, V. and Railsback, S. F. (2005). *Individual-based Modeling and Ecology*, volume 8. Princeton University Press, Princeton.
- Hansen, H. A. and McKnight, D. E. (1964). Emigration of drought-displaced ducks to the arctic. In *Transactions of the North American Wildlife and Natural Resources Conference*, volume 29, pages 119–127.
- Harrison, X. A., Blount, J. D., Inger, R., Norris, D. R., and Bearhop, S. (2011). Carry-over effects as drivers of fitness differences in animals. *Journal of Animal Ecology*, 80(1):4–18.
- Heurich, M., Hilger, A., Küchenhoff, H., Andrén, H., Bufka, L., Krofel, M., Mattisson, J., Odden, J., Persson, J., Rauset, G. R., et al. (2014). Activity patterns of Eurasian lynx are modulated by light regime and individual traits over a wide latitudinal range. *PLoS One*, 9(12):e114143.

- Hobbs, N. and Hooten, M. (2015). *Bayesian Models: A Statistical Primer for Ecologists*. Princeton University Press.
- Hobbs, N. T., Geremia, C., Treanor, J., Wallen, R., White, P. J., Hooten, M. B., and Rhyan, J. C. (2015). State-space modeling to support management of brucellosis in the Yellowstone bison population. *Ecological Monographs*, 85(4):525–556.
- Hoerling, M. P., Kumar, A., and Zhong, M. (1997). El niño, la niña, and the nonlinearity of their teleconnections. *Journal of Climate*, 10(8):1769–1786.
- Hoffman, M. D. and Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15(47):1593–1623.
- Holmes, C. C. and Held, L. (2006). Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis*, 1(1):145–168.
- Holsclaw, T., Greene, A. M., Robertson, A. W., and Smyth, P. (2017). Bayesian nonhomogeneous Markov models via Pólya-Gamma data augmentation with applications to rainfall modeling. *Ann. Appl. Stat.*, 11(1):393–426.
- Hooten, M. B. and Hobbs, N. T. (2015). A guide to Bayesian model selection for ecologists. *Ecological Monographs*, 85(1):3–28.
- Hooten, M. B., Johnson, D. S., Hanks, E. M., and Lowry, J. H. (2010). Agent-based inference for animal movement and selection. *Journal of Agricultural, Biological and Environmental Statistics*, 15(4):523–538.
- Hooten, M. B., Johnson, D. S., McClintock, B. T., and Morales, J. M. (2017). *Animal Movement: Statistical Models for Telemetry Data*. CRC Press.

- Hooten, M. B., Scharf, H. R., Hefley, T. J., Pearse, A. T., and Weegman, M. D. (2018). Animal movement models for migratory individuals and groups. *Methods in Ecology and Evolution*, 9(7):1692–1705.
- Hooten, M. B., Scharf, H. R., and Morales, J. M. (2019). Running on empty: Recharge dynamics from animal movement data. *Ecology Letters*, 22(2):377–389.
- Hooten, M. B. and Wikle, C. K. (2008). A hierarchical Bayesian non-linear spatio-temporal model for the spread of invasive species with application to the Eurasian Collared-Dove. *Environmental and Ecological Statistics*, 15(1):59–70.
- Hooten, M. B. and Wikle, C. K. (2010). Statistical agent-based models for discrete spatio-temporal systems. *Journal of the American Statistical Association*, 105(489):236–248.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558.
- Jaeger, H. (2001). The “echo state” approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148(34):13.
- Jin, M., Damianou, A., Abbeel, P., and Spanos, C. (2017). Inverse reinforcement learning via deep Gaussian process. In *Conference on Uncertainty in Artificial Intelligence*. UAI.
- Johndrow, J., Orenstein, P., and Bhattacharya, A. (2020). Scalable approximate MCMC algorithms for the horseshoe prior. *Journal of Machine Learning Research*, 21(73):1–61.
- Johnson, D. H. and Grier, J. W. (1988). Determinants of breeding distributions of ducks. *Wildlife Monographs*, pages 3–37.

- Johnson, D. S., London, J. M., Lea, M.-A., and Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, 89(5):1208–1215.
- Jordan, A., Krüger, F., and Lerch, S. (2019). Evaluating probabilistic forecasts with scoringRules. *Journal of Statistical Software*, 90(12):1–37.
- Joseph, M. B. (2020). Neural hierarchical models of ecological populations. *Ecology Letters*, 23(4):734–747.
- Kangasrääsiö, A. and Kaski, S. (2018). Inverse reinforcement learning from summary data. *Machine Learning*, 107(8-10):1517–1535.
- Kohjima, M., Matsubayashi, T., and Sawada, H. (2017). Generalized inverse reinforcement learning with linearly solvable MDP. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 373–388. Springer.
- Krapu, G. L., Reinecke, K. J., Jorde, D. G., and Simpson, S. G. (1995). Spring-staging ecology of midcontinent greater white-fronted geese. *The Journal of wildlife management*, pages 736–746.
- Kucukelbir, A., Ranganath, R., Gelman, A., and Blei, D. (2015). Automatic variational inference in Stan. In *Advances in Neural Information Processing Systems*, pages 568–576.
- Lee, D. D. and Seung, H. S. (2001). Algorithms for non-negative matrix factorization. In *Advances in Neural Information Processing Systems*, pages 556–562.
- Leos-Barajas, V., Photopoulou, T., Langrock, R., Patterson, T. A., Watanabe, Y. Y., Murgatroyd, M., and Papastamatiou, Y. P. (2017). Analysis of animal accelerometer data using hidden Markov models. *Methods in Ecology and Evolution*, 8(2):161–173.

- Leslie, P. H. (1945). On the use of matrices in certain population mathematics. *Biometrika*, 33(3):183–212.
- Li, Y., Datta, J., Craig, B. A., and Bhadra, A. (2019). Joint mean-covariance estimation via the horseshoe with an application in genomic data analysis. *arXiv preprint arXiv:1903.06768*.
- Linderman, S., Johnson, M., and Adams, R. P. (2015). Dependent multinomial models made easy: Stick-breaking with the Pólya-Gamma augmentation. In *Advances in Neural Information Processing Systems*, pages 3456–3464.
- Lukoševičius, M. and Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149.
- Ma, Q., Shen, L., and Cottrell, G. W. (2017). Deep-ESN: A multiple projection-encoding hierarchical reservoir computing framework. *arXiv preprint arXiv:1711.05255*.
- McClintock, B. T. (2017). Incorporating telemetry error into hidden Markov models of animal movement using multiple imputation. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3):249–269.
- McClintock, B. T., London, J. M., Cameron, M. F., and Boveng, P. L. (2017). Bridging the gaps in animal movement: Hidden behaviors and ecological relationships revealed by integrated data streams. *Ecosphere*, 8(3):e01751.
- McDermott, P. L. and Wikle, C. K. (2017). An ensemble quadratic echo state network for non-linear spatio-temporal forecasting. *Stat*, 6(1):315–330.
- McDermott, P. L. and Wikle, C. K. (2019a). Bayesian recurrent neural network models for forecasting and quantifying uncertainty in spatial-temporal data. *Entropy*, 21(2):184.

- McDermott, P. L. and Wikle, C. K. (2019b). Deep echo state networks with uncertainty quantification for spatio-temporal forecasting. *Environmetrics*, 30(3):e2553.
- McDermott, P. L., Wikle, C. K., and Millspaugh, J. (2017). Hierarchical nonlinear spatio-temporal agent-based models for collective animal movement. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3):294–312.
- McDermott, P. L., Wikle, C. K., and Millspaugh, J. (2018). A hierarchical spatiotemporal analog forecasting model for count data. *Ecology and evolution*, 8(1):790–800.
- McMahon, S. M. and Diez, J. M. (2007). Scales of association: hierarchical linear models and the measurement of ecological systems. *Ecology Letters*, 10(6):437–452.
- Michélot, T. and Blackwell, P. G. (2019). State-switching continuous-time correlated random walks. *Methods in Ecology and Evolution*, 10(5):637–649.
- Michélot, T., Langrock, R., Bestley, S., Jonsen, I. D., Photopoulou, T., and Patterson, T. A. (2017). Estimation and simulation of foraging trips in land-based marine predators. *Ecology*, 98(7):1932–1944.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529.
- 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.
- Morihiro, K., Isokawa, T., Nishimura, H., and Matsui, N. (2006). Emergence of flocking

- behavior based on reinforcement learning. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pages 699–706. Springer.
- Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., and Getz, W. M. (2012). Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *Journal of Experimental Biology*, 215(6):986–996.
- Ng, A. Y. and Russell, S. J. (2000). Algorithms for inverse reinforcement learning. In *ICML*, pages 663–670.
- Ogle, K., Barber, J. J., Barron-Gafford, G. A., Bentley, L. P., Young, J. M., Huxman, T. E., Loik, M. E., and Tissue, D. T. (2015). Quantifying ecological memory in plant and ecosystem processes. *Ecology Letters*, 18(3):221–235.
- Patterson, T. A., Parton, A., Langrock, R., Blackwell, P. G., Thomas, L., and King, R. (2017). Statistical modelling of individual animal movement: an overview of key methods and a discussion of practical challenges. *ASIA Advances in Statistical Analysis*, 101(4):399–438.
- Pedersen, E. J., Miller, D. L., Simpson, G. L., and Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, 7:e6876.
- Philander, S. (1990). El niño, la niña, and the southern oscillation, 293 pp. *San Diego, CA: Academic*.
- Pinsler, R., Maag, M., Arenz, O., and Neumann, G. (2018). Inverse reinforcement learning of bird flocking behavior. *ICRA Swarms Workshop*.

- Polson, N. G., Scott, J. G., and Windle, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. *Journal of the American Statistical Association*, 108(504):1339–1349.
- Polson, N. G., Sokolov, V., et al. (2017). Deep learning: a Bayesian perspective. *Bayesian Analysis*, 12(4):1275–1304.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramachandran, D. and Amir, E. (2007). Bayesian inverse reinforcement learning. In *IJCAI*, pages 2586–2591.
- Ratliff, N. D., Bagnell, J. A., and Zinkevich, M. A. (2006). Maximum margin planning. In *Proceedings of the 23rd International Conference on Machine Learning*, pages 729–736.
- Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., and Nathan, R. (2014). AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Movement Ecology*, 2(1):27.
- Ried, K., Müller, T., and Briegel, H. J. (2019). Modelling collective motion based on the principle of agency: General framework and the case of marching locusts. *PLOS ONE*, 14(2):1–21.
- Royle, J. A. and Dorazio, R. M. (2008). *Hierarchical Modeling and Inference in Ecology: the Analysis of Data from Populations, Metapopulations and Communities*. Elsevier.
- Rubin, D. B. (2004). *Multiple Imputation for Nonresponse in Surveys*, volume 81. John Wiley & Sons.

- Rugg, D. J. and Buech, R. R. (1990). Analyzing Time Budgets with Markov Chains. *Biometrics*, 46(4):1123–1131.
- Scharf, H., Hooten, M. B., and Johnson, D. S. (2017). Imputation approaches for animal movement modeling. *Journal of Agricultural, Biological and Environmental Statistics*, 22(3):335–352.
- Schaub, M. and Abadi, F. (2011). Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, 152(1):227–237.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61:85–117.
- Scrucca, L. et al. (2013). GA: A package for genetic algorithms in R. *Journal of Statistical Software*, 53(4):1–37.
- Skellam, J. G. (1951). Random dispersal in theoretical populations. *Biometrika*, 38(1/2):196–218.
- Smith, G. W. (1995). *A critical review of the aerial and ground surveys of breeding waterfowl in North America*, volume 5. US Department of the Interior, National Biological Service.
- Šošić, A., KhudaBukhsh, W. R., Zoubir, A. M., and Koepl, H. (2017a). Inverse reinforcement learning in swarm systems. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pages 1413–1421.
- Šošić, A., Zoubir, A. M., and Koepl, H. (2017b). A Bayesian approach to policy recognition and state representation learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(6):1295–1308.

- Stan Development Team (2020). RStan: the R interface to Stan. R package version 2.19.3.
- Strandburg-Peshkin, A., Farine, D. R., Crofoot, M. C., and Couzin, I. D. (2017). Habitat and social factors shape individual decisions and emergent group structure during baboon collective movement. *Elife*, 6:e19505.
- Sung, M., Soyer, R., and Nhan, N. (2007). Bayesian analysis of non-homogeneous Markov chains: Application to mental health data. *Statistics in Medicine*, 26(15):3000–3017.
- Sutton, R. S. and Barto, A. G. (1998). *Introduction to Reinforcement Learning*, volume 2. MIT press Cambridge.
- Taylor, S. D. and White, E. P. (2020). Automated data-intensive forecasting of plant phenology throughout the United States. *Ecological Applications*, 30(1):e02025.
- Todorov, E. (2007). Linearly-solvable Markov decision problems. In *Advances in Neural Information Processing Systems*, pages 1369–1376.
- Todorov, E. (2009). Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences*, 106(28):11478–11483.
- Tran, M.-N., Nguyen, N., Nott, D., and Kohn, R. (2019). Bayesian deep net GLM and GLMM. *Journal of Computational and Graphical Statistics*, pages 1–17.
- Tzikas, D. G., Likas, A. C., and Galatsanos, N. P. (2008). The variational approximation for Bayesian inference. *IEEE Signal Processing Magazine*, 25(6):131–146.
- Van Hasselt, H., Guez, A., and Silver, D. (2016). Deep reinforcement learning with double Q-learning. In *Thirtieth AAAI Conference on Artificial Intelligence*.

- 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.
- Wikle, C. K. (2003). Hierarchical models in environmental science. *International Statistical Review*, 71(2):181–199.
- Wilks, D. S. (2011). *Statistical Methods in the Atmospheric Sciences*, volume 100. Academic press.
- Wu, G., Holan, S. H., and Wikle, C. K. (2013). Hierarchical bayesian spatio-temporal conway–maxwell poisson models with dynamic dispersion. *Journal of Agricultural, Biological, and Environmental Statistics*, 18(3):335–356.
- Wulfmeier, M., Ondruska, P., and Posner, I. (2015). Deep inverse reinforcement learning. *arXiv preprint arXiv:1507.04888*.
- Yamaguchi, S., Naoki, H., Ikeda, M., Tsukada, Y., Nakano, S., Mori, I., and Ishii, S. (2018). Identification of animal behavioral strategies by inverse reinforcement learning. *PLoS Computational Biology*, 14(5):e1006122.
- Zammit-Mangion, A. (2020). *FRK: Fixed Rank Kriging*. R package version 0.2.2.1.
- Ziebart, B. D., Maas, A. L., Bagnell, J. A., and Dey, A. K. (2008). Maximum entropy inverse reinforcement learning. In *Aaai*, volume 8, pages 1433–1438. Chicago, IL, USA.
- Zucchini, W., MacDonald, I. L., and Langrock, R. (2016). *Hidden Markov Models for Time Series: An Introduction using R*. Chapman and Hall/CRC.

VITA

Toryn Leanne Jake Schafer was born in Ottawa, Illinois, on July 27, 1992. Daughter of Crystal Edwards, she was raised in Illinois and graduated from Zion-Benton Township High School. She earned her B.S. from Colorado State University (2014) in Fish, Wildlife and Conservation Biology and Statistics. She completed her Bachelor's research under the supervision of Dr. Stewart W. Breck and Dr. Larissa Bailey. She moved to Columbia, Missouri to pursue a doctoral degree in Statistics at the University of Missouri under Dr. Christopher K. Wikle. She completed the Statistics M.A. degree requirements (2018). She has accepted a postdoctoral research position in the Department of Statistics and Data Science at Cornell University under supervision of Dr. David S. Matteson.