The Annals of Applied Statistics
2021, Vol. 15, No. 2, 813–830
https://doi.org/10.1214/20-AOAS1421

A BAYESIAN SEMIPARAMETRIC JOLLY–SEBER MODEL WITH INDIVIDUAL HETEROGENEITY: AN APPLICATION TO MIGRATORY MALLARDS AT STOPOVER

BY GUOHUI WU[1], SCOTT H. HOLAN[2], ALEXIS AVRIL[3,*] AND JONAS WALDENSTRÖM[3,†]

[1]*SAS Institue Inc., raywu2014@gmail.com*
[2]*Department of Statistics, University of Missouri, holans@missouri.edu*
[3]*Centre for Ecology and Evolution in Microbial Model Systems, Linnaeus University, *alexis.avril@gmail.com; †jonas.waldenstrom@lnu.se*

We propose a Bayesian hierarchical Jolly–Seber model that can accommodate a semiparametric functional relationship between external covariates and capture probabilities, individual heterogeneity in departure due to an internal time-varying covariate and the dependence of arrival time on external covariates. Modelwise, we consider a stochastic process to characterize the evolution of the partially observable internal covariate that is linked to departure probabilities. Computationally, we develop a well-tailored Markov chain Monte Carlo algorithm that is free of tuning through data augmentation. Inferentially, our model allows us to make inference about stopover duration and population sizes, the impacts of various covariates on departure and arrival time and to identify flexible yet data-driven functional relationships between external covariates and capture probabilities. We demonstrate the effectiveness of our model through a motivating dataset collected for studying the migration of mallards (*Anas platyrhynchos*) in Sweden.

1. Introduction.

Migration is a common phenomenon in birds, especially in areas with pronounced seasonal variation. However, in most species, migration is not conducted as a single flight from the breeding area to the nonbreeding area; rather it is broken down into shorter legs interspersed with stopovers of variable length at suitable sites where energy spent during migration can be replenished (e.g., see Newton (2010), and the references therein). Mostly determined by the time spent at stopover sites, the overall speed of migration is tightly linked to behaviors at stopover sites, and the distribution and quality of stopover sites impacts the success and survival of birds during migration. A key to stopover duration analysis rests on understanding various species-specific stopover behaviors and how intrinsic factors, such as internal condition (e.g., body condition), and external factors, such as environmental conditions, contribute to these behaviors. For this reason, the primary objectives in stopover studies include estimating the timing of arrival and departure, stopover duration (i.e., the length of stay at a stopover site), stopover population sizes at different levels and understanding the role of internal and external factors.

Capture–recapture models are still the gold standard for studying population level processes, such as entrance, capture and departure, and to estimate parameters related to these processes (Williams, Nichols and Conroy (2002), Royle and Dorazio (2008)), which makes them good candidates for stopover duration analysis (e.g., see King et al. (2010), Matechou (2010), Pledger et al. (2009), and the references therein). To exemplify this, Kaiser (1995), Dinsmore and Collazo (2003) and Rice et al. (2007) demonstrate applications of the Cormack–Jolly–Seber (CJS) model to achieve stopover duration estimation. Importantly, the

---

Received December 2019; revised November 2020.

Key words and phrases. Capture–recapture, individual heterogeneity, low-rank thin-plate splines, Ornstein–Uhlenbeck process, stopover duration analysis.

resulting estimate of the stopover duration from the CJS model can be biased due to the conditional nature of the model and unknown arrival time (Pledger et al. (2009)). Building upon the Schwarz and Arnason (1996) formulation of the Jolly–Seber (JS) model (Jolly (1965), Seber (1965)), Pledger et al. (2009) develop a flexible stopover model to allow capture probability and stopover retention probability (i.e., the probability of staying at a study site) to depend on an unknown time since arrival. Apart from deriving indirect estimate of the mean stopover duration, they also consider modeling the stopover retention curve to examine different stopover departure patterns. To extend the model by Pledger et al. (2009), Matechou et al. (2014) develop a mixture model where captured individuals do not need to be correctly and distinctly marked. In other words, data for such an extended model consists of counts of individuals captured in each sampling period rather than the encounter history of uniquely marked individuals. Subsequently, Lyons et al. (2016) develop a Bayesian stopover model that accommodates both encounter histories of uniquely marked individuals and counts of unmarked individuals. Their model allows for the estimation of capture and stopover retention probabilities, entrance probabilities, stopover population sizes and stopover duration.

Particularly for individual heterogeneity, Bonner and Schwarz (2009) develop a JS model that can account for individual heterogeneity in capture and survival probabilities due to a time-varying continuous individual covariate using a two-step Bayesian approach. The primary disadvantage of this two-step approach is that the entrance probability does not appear in the likelihood, disallowing us to link entrance probabilities to external covariates or to estimate stopover duration. Schofield and Barker (2011) present a general framework for a variety of open population models (the JS model included) with individual heterogeneity and demonstrate how freely available software programs, such as BUGS (Lunn et al. (2000)), can be used for Bayesian estimation of these models. In principle, the authors provide a convenient framework that can be adapted for stopover analysis; however, in practice, their approach is limited to smaller datasets due to computational inefficiency. Recently, Matechou et al. (2016) develop a stopover model by extending the JS model to allow individuals to arrive in different groups and, hence, their model accounts for heterogeneity in departure due to a group effect.

Despite the usefulness of the aforementioned stopover models, many real-world applications require development of an application-specific model. For our mallard migration studies, key objectives are to understand migratory decisions and stopover behaviors and their contributing factors and to estimate stopover duration and population sizes at different scales. Specifically, we need to relate various type of covariates to the processes that are tightly coupled with the population dynamics in order to test hypothesis and to identify the drivers for certain migratory behaviors. On the one hand, it remains to be known how mallards adjust individual departure behavior according to their refueling rates at the stopover site and their current body condition in terms of fat loads (Berthold (2001)). As a result, our model should be able to address individual heterogeneity in departure decisions due to the internal factor—body condition—which varies with both time and individual. On the other hand, weather has known link to bird migration during departure and also aloft. In general, birds prefer initiating a flight when winds provide flight assistance, that is, tailwinds, and under other conditions favorable for flying, such as under low rainfalls (Berthold (2001)). Consequently, we need to link the timing of arrival and capture probabilities to some external weather covariates to understand how weather conditions affect mallard's migratory decisions. Last, we need to allow flexible modeling of covariate effects on capture outcome.

Motivated by mallard migration studies, we develop a stopover model using data augmentation (Tanner and Wong (1987)) under the Bayesian hierarchical state-space framework. The methodological contributions can be described as follows. First, our model accounts for individual heterogeneity in departure due to a time-varying continuous individual covariate.

Since the values of such a covariate are observable only when an individual is captured, a stochastic process is proposed to characterize the change of this covariate over time. Second, our model links the arrival time to external factors and, hence allows us to draw inference about their impacts on the time of arrival. Furthermore, our model allows for a data-driven functional relationship between the capture probabilities and external factors through the use of smoothing splines, which enables us to detect nonlinear covariate effects. More importantly, we develop a well-tailored Markov chain Monte Carlo (MCMC) algorithm for our model to achieve greater computational efficiency.

This paper is organized as follows. Section 2 introduces our motivating dataset from the mallard migration studies. Section 3 provides details about our proposed model and two goodness-of-fit criteria for model assessment. A simulated example, illustrating the effectiveness of our modeling approach, is presented in Section 4. Section 5 demonstrates the application of our methodology through a stopover duration analysis for our motivating data collected by the Ottenby Bird Observatory in Sweden. Discussion is provided in Section 6. Further details surrounding the full conditional distributions and our customized MCMC sampling algorithm are provided in the Supplementary Material (Wu et al. (2021)).

2. The mallard migration data.

The mallard (*Anas platyrhynchos*), is the most common and widespread dabbling duck in the Northern hemisphere and an important model species for studies of ecological processes (Gunnarsson et al. (2012)), harvest management (Nichols et al. (2007)) and epidemiology of bird borne infections (Latorre-Margalef et al. (2009), van Toor et al. (2018a)). It is a partial migrant, where southerly populations in the distribution range tend to be resident and the northernmost obligatory migrants, and in other populations a mix of resident and migrants (Cramp and Simmons (1977)). The mallard is a medium-sized bird with heavy wing loading where migration is energetically costly. From ringing and telemetry studies it is clear that migratory mallards break up their journey into shorter flights and spend a large proportion of their migration time at stopover sites, replenishing resources and preparing for the next leg of migration (Gunnarsson et al. (2012)). Thus, stopover sites play a key role in successful migration and survival of mallards, and a priority for sustainable mallard management is to better characterize the ecology of birds at stopover. This includes assessing the timing of migration and densities of birds at specific stopover sites and to what extent internal and external factors (e.g., body condition and weather) affect stopover behaviors.

The long-term capture of mallards has been carried out at Ottenby Bird Observatory on the Swedish island of Öland in the Baltic Sea (56°12′N, 16°24′E) (see Figure 1). The southernmost part of this island is an attractive stopover site for mallards within the Northwest European Flyway, offering habitats for both roosting and foraging (Bengtsson et al. (2014)). Mallards that utilize our study site—Ottenby—mainly originate from mainland Sweden, Estonia, Finland and Russia (Gunnarsson et al. (2012)). After leaving Öland these mallards migrate to wintering areas in Northwestern Europe, predominantly in southern Denmark, northern Germany and the Netherlands (Gunnarsson et al. (2012)).

Data collection started in 2002 and, originally, aimed for monitoring presence of influenza A virus in birds, but the data of banded individuals over time is also extremely suitable for addressing stopover ecology questions. To collect data, Ottenby Bird Observatory used a stationary trap at the study site to catch mallards for ringing and epidemiological studies. In particular, mallards were attracted by bait grain and by the presence of a few (normally around 10) domestic ducks kept in a compartment of the trap. Traps were inspected daily during the field seasons, and any wild duck captured was ringed and measured for structural size (i.e., the distance from the tip of the bill to the back of the head) and body mass and subsequently released. This data-collection process, over the course of a stopover season, results in the

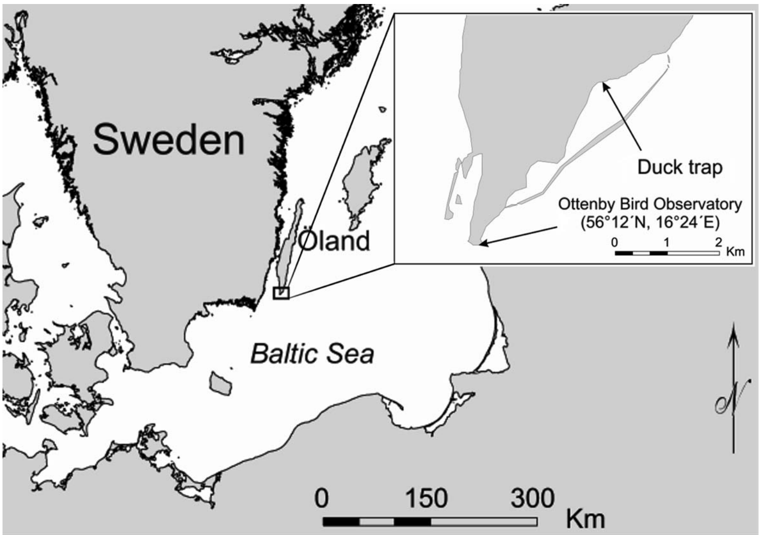

FIG. 1. *Plot of study site for monitoring mallards in the Swedish island of Öland.*

capture–recapture data. Data available for our analysis was collected from 2004–2011 during the autumn migration season, which begins on August 1st and ends on December 16th of each year.

Motivated by the mallard data at hand, our primary goal is to develop a model that accomplishes three important research objectives. The first objective is to determine whether there is individual heterogeneity in mallards' departure due to the internal factor—body condition (i.e., body mass corrected by the structural size). The second objective is to estimate stopover duration, stopover population sizes and to detect whether there is a temporal trend for daily stopover population sizes. The third objective is to understand how external factors, such as different weather measures, relate to the timing of arrival and capture outcome for mallards at our study site.

3. Methodology.

3.1. *Parameters and notation.* Consider a capture–recapture experiment spanning over T sampling occasions at distinct times $t_1, t_2, \dots, t_T$ across a certain study area targeting a population $\mathcal{P}$ regarding a particular species of interest. Let the total number of individuals of $\mathcal{P}$ that visited the study area during the entire experiment be N, which is unknown and needs to be estimated. For $k = 1, 2, \dots, T - 1$, let $\delta_k = t_{k+1} - t_k$ denote the time interval between two consecutive sampling occasions k and $k + 1$. Without loss of generality, we assume $t_1 < t_2 < \cdots < t_T$; that is, $\delta_k > 0$ for $k = 1, 2, \dots, T - 1$.

Denote $z_{i,t}$ as a binary latent variable to indicate the state of an individual $i \in \mathcal{P}$ at time t, where the term "state" refers to two statuses of an individual. That is, for $i = 1, 2, \dots, N$ and $1 \le t \le T$, $z_{i,t} = 1$ if individual i is alive and present in the study area at time t; and $z_{i,t} = 0$ if individual i has not entered the study area or is dead at time t. Moreover, we assume that a total of n individuals were caught during the study. For each individual being caught, denote a row vector $\boldsymbol{y}_i = (y_{i,1}, y_{i,2}, \dots, y_{i,T})$ as the corresponding capture history, where $y_{i,t}$ is a binary variable indicating if individual i is caught at occasion t; that is, $y_{i,t} = 1$ if individual i is caught at occasion t and 0 otherwise for $i = 1, 2, \dots, n$ and $t = 1, 2, \dots, T$. Often, the observed capture history is summarized using a $n \times T$ matrix $\mathbf{y}_{\text{obs}}$, whose ith row consists of $\boldsymbol{y}_i$. Upon the capture of each individual, measurements on a set of individual covariates are taken and recorded.

Motivated by the mallard data, we consider the single covariate case and allow the individual covariate X to be continuous and to vary with both individual and time, for example, body condition. The value of $X_{i,t}$ is observable when an individual i is captured at time t and unobservable otherwise. As a result, for each individual i we need to model the evolution of $X_{i,t}$ over time t.

3.2. *Modeling time-varying continuous covariates.* Let $X(t)$ be a continuous variable at time $t \in \mathcal{T} = [0, T]$. We assume that $X(t)$ follows an Ornstein–Uhlenbeck (OU) process; that is, $X(t)$ satisfies a stochastic differential equation of the form

$$dX(t) = \tau\bigl(\alpha - X(t)\bigr)\,dt + \sigma\,dW(t), \tag{3.1}$$

where $\sigma > 0$ controls the noise variance, $\tau > 0$ describes the rate of mean reversion, α is the long-term (or asymptotic) mean and $W(t)$ is a standard Wiener process on $t \in \mathcal{T}$. It is straightforward to see that (3.1) reduces to the von Bertalanffy growth equation (von Bertalanffy (1938)) when $\sigma = 0$. When $\sigma \neq 0$, the extra random noise term in the OU process provides increased flexibility, accounting for random noise resulting from several factors, for example, measurement error and/or random variation due to changes in the environmental conditions (Filipe et al. (2010)). Denote $X_t = X(t)$ for $t \in \mathcal{T}$, the OU process is *stationary* (i.e., $(X_{t_1}, X_{t_2}, \ldots, X_{t_s})$ and $(X_{t_1+h}, X_{t_2+h}, \ldots, X_{t_s+h})$ are identically distributed), *Markovian* (i.e., $P(X_{t_s} \le x | X_{t_1}, X_{t_2}, \ldots, X_{t_{s-1}}) = P(X_{t_s} \le x | X_{t_{s-1}})$), and $(X_{t_1}, X_{t_2}, \ldots, X_{t_s})$ follows a multivariate *Gaussian* distribution (see Finch (2004), and the references therein) for $t_1 < t_2 < \cdots < t_s$ and $h > 0$.

For $t_{k-1} < t_k$, it follows that the transition distribution takes the following form:

$$X(t_k) | \alpha, \tau, \sigma^2, X(t_{k-1}) = x_{k-1} \sim \mathrm{N}\bigl(\mu(x_{k-1}, \tau, \delta_{k-1}, \alpha), V\bigl(\sigma^2, \tau, \delta_{k-1}\bigr)\bigr),$$

where

$$\mu(x_{k-1}, \tau, \delta_{k-1}, \alpha) = \exp(-\tau\delta_{k-1})x_{k-1} + \bigl\{1 - \exp(-\tau\delta_{k-1})\bigr\}\alpha,$$

$$V\bigl(\sigma^2, \tau, \delta_{k-1}\bigr) = \frac{\sigma^2\{1 - \exp(-2\tau\delta_{k-1})\}}{2\tau}$$

(see Filipe et al. (2010), and the references therein). Compared with the diffusion process used by Bonner and Schwarz (2006, 2009) and Schofield and Barker (2011), the OU process we consider provides estimates for the rate parameter τ and long-term mean α.

We characterize the "growth" of an individual in terms of the measurement $X(t)$ on a continuous time-varying covariate X at time $t \in \mathcal{T}$ by assuming that $X(t)$ satisfies the OU process defined by (3.1). Denote $X_{i,t}$ as the measurement on covariate X for an individual i at time t. At discrete sampling times $t = 2, \ldots, T$, the conditional distribution of $X_{i,t}$ takes the following form:

$$X_{i,t} | \alpha, \tau, \sigma^2, X_{i,t-1} = x_{i,t-1} \sim \mathrm{N}\bigl(\mu(x_{i,t-1}, \tau, \delta_{t-1}, \alpha), V\bigl(\sigma^2, \tau, \delta_{t-1}\bigr)\bigr),$$

where $X_{i,1} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(\alpha, \sigma_0^2)$ is assumed for the initial condition.

3.3. *Semiparametric Jolly–Seber model with individual heterogeneity.* Our proposed model is characterized by a state model, an observation model and a parameter model. The state model describes how the latent state $z_{i,t}$ evolves over time, whereas the observation model describes the capture outcome $y_{i,t}$ conditioning on the latent state $z_{i,t}$. The parameter model describes how capture probabilities, departure probabilities and entrance probabilities are linked to some internal and external factors.

Since the dimension of latent variable $z_{i,t}$ varies with the unknown parameter N, the number of parameters is not fixed, which will cause some computational disadvantages. To maintain a constant number of parameters, the *data augmentation* technique is often utilized (e.g., see Royle and Dorazio (2008)). At the core of data augmentation, we first introduce a parameter $M > N$ and then augment the observed capture history $\mathbf{y}_{\text{obs}}$ by an artificial zero matrix $\mathbf{y}_{\text{aug}}$ of size $(M - n) \times T$, each row of which represents the capture history of individuals that were never captured during the study. As will be shown, this data augmentation also allows us to formulate our model in a hierarchical manner.

3.3.1. *State model.* Following Royle and Dorazio (2008), the *state model* can be defined by

$$z_{i,1} | \pi_1^c \sim \text{Bernoulli}(\pi_1^c), \tag{3.2}$$

$$z_{i,t+1} | z_{i,t}, \phi_{i,t}, \pi_{t+1}^c \sim \text{Bernoulli}(\phi_{i,t} z_{i,t} + \pi_{t+1}^c R_{i,t}), \tag{3.3}$$

where $R_{i,t} = \prod_{s=1}^{t} 1(z_{i,s} = 0)$ indicates whether an individual i can enter the population right after time t for $i = 1, 2, \ldots, M$ and $t = 1, 2, \ldots, T - 1$. Here, $1(z = a)$ is the indicator function that takes value 1 if $z = a$ and 0 otherwise. In addition, $\phi_{i,t}$ refers to survival probability (or stopover retention probability in a stopover model), that is, the probability that an individual i will remain in the study area at time $t + 1$ given its presence in the study area at time t for $t = 1, 2, \ldots, T - 1$. Moreover, π_t^c denotes the conditional entrance probability at time t, given that an individual has not entered the study area; that is,

$$\pi_{t+1}^c = \frac{\beta_t}{\sum_{j=t}^{T-1} \beta_j}, \quad t = 0, 1, \ldots, T - 1,$$

where β_t is the entrance probability at time t, that is, the probability that an individual enters the study area between time t and $t + 1$. By definition, it follows that $\sum_{t=1}^{T} \beta_{t-1} = 1$.

The interpretation of the state model described in (3.2) and (3.3) is straightforward. First, (3.2) indicates that an individual i is subject to entrance with probability π_1^c at time $t = 1$. In (3.3) we see that if an individual i has not entered the study area right before time $t + 1$ (i.e., $R_{i,t} = 1$), it is subject to entrance with probability π_{t+1}^c. Second, if an individual i has already entered and is present in the study area at time t, it will remain in the study area at time $t + 1$ with probability $\phi_{i,t}$. Moreover, if an individual i is dead at time t (i.e., $z_{i,t} = 0$), it will remain dead (i.e., $z_{i,t^*} = 0$ for $t^* = t + 1, \ldots, T$).

3.3.2. *Observation model.* The *observation model* is given by (Royle and Dorazio (2008))

$$w_i | \Psi \sim \text{Bernoulli}(\Psi), \tag{3.4}$$

$$y_{i,t} | w_i, z_{i,t}, p_t \sim \text{Bernoulli}(w_i z_{i,t} p_t), \tag{3.5}$$

for $i = 1, 2, \ldots, M$ and $t = 1, 2, \ldots, T$. Here, p_t is the probability of capturing an individual that is alive at time t. The binary variable w_i in (3.4) is introduced as a result of data augmentation to indicate if an individual i is a member of $\mathcal{P}$. Specifically, $w_i = 1$ if an individual i is a member of $\mathcal{P}$ and 0 otherwise. For an individual i that is captured at least once during the study (i.e., $\boldsymbol{y}_i \neq \mathbf{0}$), it is clear that $w_i = 1$ is implied. According to (3.5), an individual i is subject to capture at time t only if it is alive and present in the study area (i.e., $z_{i,t} = 1$). In addition, the augmented capture history $\boldsymbol{y}_i = \mathbf{0}$ for $i = n + 1, \ldots, M$ can be accommodated by (3.5).

An important feature of building the JS model from the "individual" up is that it enables us to estimate certain quantities that are important in stopover duration analysis fairly easily.

For example, the total number of animals that visited the site during the entire study period, N, can be estimated as $N = \sum_{i=1}^{M} w_i$. It then follows from (3.4) that $N \sim \text{Binomial}(M, \Psi)$. As a result, if M is specified to be close to N, Ψ will approach 1, which would suggest that M needs to be increased. On the other hand, the stopover population size at time t, N_t can be estimated as $N_t = \sum_{i=1}^{M} w_i z_{i,t}$. Moreover, we can estimate the mean stopover duration averaged over all captured individuals as (Lyons et al. (2016))

$$S = \frac{\sum_{i=1}^{n} \sum_{t=1}^{T} z_{i,t}}{n}.$$

The number of individuals alive at both times t_1 and t_2, say N_{t_1,t_2}, can be estimated as $N_{t_1,t_2} = \sum_{i=1}^{M} w_i z_{i,t_1} z_{i,t_2}$.

3.3.3. *Parameter model.* The *parameter model* links capture probabilities p_t, departure probabilities $d_{i,t} = 1 - \phi_{i,t}$ and entrance probabilities β_t with various covariates as follows:

$$\log\left(\frac{p_t}{1-p_t}\right) = \boldsymbol{g}_t'\boldsymbol{\zeta} + \sum_{k=1}^{K} u_k |o_t - \kappa_k|^3, \quad t = 1, 2, \ldots, T, \tag{3.6}$$

$$\log\left(\frac{1-\phi_{i,t}}{\phi_{i,t}}\right) = \boldsymbol{x}_{i,t}'\boldsymbol{\theta} = \theta_0 + \theta_1 x_{i,t}, \quad t = 1, 2, \ldots, T-1, \tag{3.7}$$

$$\log\left(\frac{\beta_{t-1}}{\beta_{T-1}}\right) = \boldsymbol{\Lambda}_t'\boldsymbol{\gamma}, \quad t = 1, 2, \ldots, T-1, \tag{3.8}$$

where $\boldsymbol{g}_t = (g_{1t}, g_{2t}, \ldots, g_{Qt})'$ is a $Q \times 1$ vector that consists of values for external covariates $g_1, g_2, \ldots, g_Q$ at time t, for example, trap and weather related covariates; $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \ldots, \zeta_Q)'$ denotes a $Q \times 1$ vector of regression coefficients, and $\boldsymbol{u} = (u_1, u_2, \ldots, u_K)' \sim N(\mathbf{0}, \sigma_u^2 \boldsymbol{\Omega}^{-1})$, where $\boldsymbol{\Omega}$ is a matrix whose (k, l)th entry is $|\kappa_k - \kappa_l|^3$ for $1 \le k, l \le K$ and K is the number of knot points. The fixed knot κ_k is chosen to be the sample quantile of the o_t's corresponding to probability $k/(K+1)$ for $k = 1, 2, \ldots, K$. Moreover, $\boldsymbol{\Lambda}_t = (\Lambda_{1t}, \Lambda_{2t}, \ldots, \Lambda_{Pt})'$ denotes a $P \times 1$ vector that consists of the values of external covariates $\Lambda_1, \Lambda_2, \ldots, \Lambda_P$ at time t, for example, weather conditions, and $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_P)'$ is a $P \times 1$ vector of regression coefficients.

The semiparametric model for the capture probabilities in (3.6) has two parts: (1) a parametric part that posits a linear relationship between covariates $g_1, g_2, \ldots, g_Q$ and the logit of p_t, and (2) a nonparametric part that accommodates a nonlinear effect of the covariate o on the logit of p_t. Albeit many basis functions can be used for the nonparametric part of the model (e.g., see Ruppert, Wand and Carroll (2003)), low-rank thin-plate splines are preferred over other basis functions since the former leads to better mixing of the MCMC chains in Bayesian analysis (Crainiceanu, Ruppert and Wand (2005)). Let $\mathbf{O}$ be a $T \times K$ matrix with the (t, k)th element being $|o_t - \kappa_k|^3$ for $1 \le t \le T$ and $1 \le k \le K$, (3.6) can be reparameterized as

$$\log\left(\frac{p_t}{1-p_t}\right) = \boldsymbol{g}_t'\boldsymbol{\zeta} + \boldsymbol{Z}_t\boldsymbol{b},$$

where $\boldsymbol{b} = \boldsymbol{\Omega}^{\frac{1}{2}}\boldsymbol{u} \sim N(\mathbf{0}, \sigma_u^2 \mathbf{I}_K)$ and $\boldsymbol{Z}_t$ is the tth row of the matrix $\mathbf{Z} = \mathbf{O}\boldsymbol{\Omega}^{\frac{1}{2}}$.

As shown in (3.7), the departure probabilities are linked to an internal covariate X that varies with both time and individual (such as body condition) to account for individual heterogeneity. Strictly speaking, departures can arise from three outcomes—start of a migratory flight, relocation to another habitat that is not covered by traps and death. When the sampling period is relatively short, as it is the case in our motivating mallard example, death between two consecutive sampling periods is almost negligible. Consequently, the term departure primarily refers to start of another migratory flight or relocation to another habitat. The realization of $X_{i,t}$ for an individual i at time t, $x_{i,t}$, is only observable when the individual is caught. As a result, we assume that the "growth" of an individual in terms of the internal covariate X follows the OU process discussed in Section 3.2 so that missing values of $X_{i,t}$ can be "imputed" by conditioning on the observed data.

In (3.8), we link the entrance probabilities to external factors to infer the impacts of these factors on the timing of arrival. Due to the implied restriction $\sum_{t=1}^{T} \beta_{t-1} = 1$, (3.8) is equivalent to the following:

$$\beta_t = \begin{cases} \dfrac{\exp(\mathbf{\Lambda}_t' \boldsymbol{\gamma})}{1 + \sum_{j=0}^{T-2} \exp(\mathbf{\Lambda}_j' \boldsymbol{\gamma})} & \text{if } t = 0, 1, \ldots, T-2, \\ \dfrac{1}{1 + \sum_{j=0}^{T-2} \exp(\mathbf{\Lambda}_j' \boldsymbol{\gamma})} & \text{if } t = T-1. \end{cases}$$

3.4. *Priors and posteriors.* To complete the specification of our model, we need to assign prior distributions for the model parameters and derive the full conditional distributions. Denote $\boldsymbol{w} = \{w_i : i = 1, 2, \ldots, M\}$ and $\mathbf{z} = \{z_{i,t} : i = 1, 2, \ldots, M; t = 1, 2, \ldots, T\}$, the set of parameters in the model we propose is $\mathbf{\Theta} = \{\boldsymbol{\zeta}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \Psi, \boldsymbol{w}, \mathbf{z}, \sigma_0^2, \alpha, \tau, \sigma^2, \sigma_u^2\}$. Denote $\mathrm{IG}(A, B)$ as the inverse gamma distribution with shape parameter A and scale parameter B. We assign prior distributions as follows: $\boldsymbol{\zeta} \sim \mathrm{N}(\boldsymbol{\mu}_\zeta, \mathbf{\Sigma}_\zeta)$; $\boldsymbol{b} \sim \mathrm{N}(\mathbf{0}, \sigma_b^2 \mathbf{I}_K)$; $\boldsymbol{\theta} \sim \mathrm{N}(\boldsymbol{\mu}_\theta, \mathbf{\Sigma}_\theta)$; $\boldsymbol{\gamma} \sim \mathrm{N}(\boldsymbol{\mu}_\gamma, \mathbf{\Sigma}_\gamma)$; $w_i \overset{\text{i.i.d.}}{\sim} \mathrm{Bernoulli}(\Psi)$ for $i = 1, 2, \ldots, M$; $\Psi \sim \mathrm{Beta}(a_\Psi, b_\Psi)$; $\sigma_0^2 \sim \mathrm{IG}(q_0, r_0)$; $\alpha \sim \mathrm{N}(\mu_\alpha, \sigma_\alpha^2)$; $\sigma^2 \sim \mathrm{IG}(q_\sigma, r_\sigma)$; $\tau \sim \mathrm{Uniform}(q_\tau, r_\tau)$; and $\sigma_u^2 \sim \mathrm{IG}(q_u, r_u)$. In our implementation we choose vague priors that are noninformative relative to the scale of data.

Let $\mathbf{Y} = \mathbf{y}_{\text{obs}} \cup \mathbf{y}_{\text{aug}}$ denote the augmented capture history. Assuming conditional independence, the joint posterior distributions of the model parameters $[\mathbf{\Theta}|\mathbf{Y}]$ can be derived as

$$\begin{aligned}[\mathbf{\Theta}|\mathbf{Y}] \propto \Bigg\{ \prod_{i=1}^{M} \Bigg(\prod_{t=1}^{T-1} [z_{i,t+1}|z_{i,t}, \boldsymbol{\gamma}, x_{i,t}, \boldsymbol{\theta}][x_{i,t+1}|x_{i,t}, \alpha, \tau, \sigma^2] \Bigg) [z_{i,1}|\boldsymbol{\gamma}][x_{i,1}|\alpha, \sigma_0^2] \\ \times \Bigg(\prod_{t=1}^{T} [y_{i,t}|\omega_i, z_{i,t}, \boldsymbol{\zeta}, \boldsymbol{b}] \Bigg) [w_i|\Psi] \Bigg\} [\boldsymbol{\theta}][\sigma_0^2][\alpha][\tau][\sigma^2][\boldsymbol{\gamma}][\boldsymbol{b}|\sigma_u^2][\boldsymbol{\zeta}][\Psi][\sigma_u^2],\end{aligned}$$

where the square bracket [ ] denotes a probability density function.

3.5. *Model assessment.* An extremely important aspect of Bayesian modeling is to evaluate goodness-of-fit for the model being considered. In the context of capture–recapture models, the Bayesian p-value is often considered (e.g., see King et al. (2010), and the references therein). Roughly speaking, the Bayesian p-value is a posterior probability that measures the similarity between the data generated from the posterior predictive distribution under a specified model and the observed data. To calculate the Bayesian p-value, we first define a discrepancy function $h(\boldsymbol{D}, \mathbf{\Theta})$, where $\boldsymbol{D}$ and $\mathbf{\Theta}$ denote the data and the parameters for the model being considered, respectively. Then, we calculate the value of the discrepancy function for both the observed data $\boldsymbol{D}^\star$ and the simulated data $\boldsymbol{D}'$, which is generated conditioning on the posterior distribution of model parameters. Finally, the Bayesian p-value is defined as the percentage of times that values of the discrepancy function for $\boldsymbol{D}^\star$ exceeds those of the discrepancy function for $\boldsymbol{D}'$. Mathematically, the definition of the Bayesian p-value, P_b, can be formulated as $P_b = p(h(\boldsymbol{D}^\star, \boldsymbol{\Theta}) > h(\boldsymbol{D}', \boldsymbol{\Theta})|\boldsymbol{D}^\star)$. As a rule of thumb, a Bayesian p-value close to 0 or 1 indicates that the model being considered does not provide a good fit to the data and that there is inconsistence between the model and data (Gelman (2003)).

For the model we propose, goodness-of-fit requires the assessment of two components. On the one hand, we need to assess the goodness-of-fit for the overall JS model to the data. On the other hand, we need to evaluate the goodness-of-fit for the OU process to the time-varying continuous individual covariate X. Consequently, it suffices to calculate the Bayesian p-values P_b^{JS} for the JS model and P_b^{OU} for modeling the internal covariate X using the OU process. Among the many choices of discrepancy functions (e.g., see Brooks, Catchpole and Morgan (2000)), we used the complete log-likelihood function for P_b^{JS}; that is, $h^{\mathrm{JS}}(\boldsymbol{D}, \boldsymbol{\Theta}) = \ell(\mathbf{Y}, \boldsymbol{z}|\boldsymbol{\Theta}_{-\boldsymbol{z}}, \boldsymbol{D})$, where $\ell(\mathbf{Y}, \boldsymbol{z}|\boldsymbol{\Theta}, \boldsymbol{D})$ is the complete log-likelihood function of $\mathbf{Y}$, $\boldsymbol{z}$, given all model parameters excluding $\boldsymbol{z}$ (i.e., $\boldsymbol{\Theta}_{-\boldsymbol{z}}$) and the data $\boldsymbol{D}$. Similar to King and Brooks (2002), we consider the complete log-likelihood function as a discrepancy measure primarily due to its tractability and computational advantages. Different from Bonner and Schwarz (2009), for P_b^{OU}, we compare the observed and expected value of the individual covariate for each capture rather than recapture and consider the discrepancy function to be

$$h^{\mathrm{OU}}(\boldsymbol{D}, \boldsymbol{\Theta}) = \frac{1}{n_c} \sum_{i \in \mathcal{P}} \sum_{t: y_{i,t}=1} \left\{ \frac{x_{i,t} - E(x_{i,t}|x_{i,t-1})}{\sigma(x_{i,t}|x_{i,t-1})} \right\}^2,$$

where $n_c = \sum_{i \in \mathcal{P}} \sum_{t=1}^{T} y_{i,t}$ is the total number of captures over T sampling occasions and $\sigma(x_{i,t}|x_{i,t-1})$ denotes the standard deviation for the conditional distribution of $x_{i,t}|x_{i,t-1}$.

Our adoption of Bayesian p-values is intended only to provide an additional layer of posterior predictive checks. Nevertheless, given our particular choices of discrepancy measures, these Bayesian p-values alone should not be used to prove or disprove the legitimacy of our proposed model. To this end, we acknowledge two major limitations of using Bayesian p-value for assessing a model fit. First, different choices of discrepancy measures can lead to different conclusions about goodness-of-fit. Second, it can be challenging to propose discrepancy measures with high power for detecting misfit of a model to the data. Consequently, our use of Bayesian p-values is meant as an indication of model adequacy.

4. Simulated example. To evaluate the performance of our proposed model, we consider a simulated example where the exact model specification is chosen for illustration. For this simulation we set $N = 3000$ and $T = 138$. The parameter model is given by (3.6), (3.7) and (3.8). For the parametric part of (3.6), we consider $Q = 3$ and $\boldsymbol{\zeta} = (1.0, -0.9, 0.6)'$ with covariates simulated according to $g_{1t}, g_{2t}, g_{3t} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, 1)$ for $t = 1, 2, \ldots, T$. For the nonparametric part of (3.6), we set $o_t = \frac{t}{T}$ for $t = 1, 2, \ldots, T$, $\sigma_u^2 = 0.25$, and $K = 10$. In (3.7) we set $\boldsymbol{\theta} = (-1.8, 0.3)'$ and simulate values of the internal covariate X from the OU process with $\sigma_0^2 = 1.37$, $\alpha = 0.20$, $\tau = 0.19$, $\sigma^2 = 0.36$, and $\delta_k \equiv 1.0$ (for $k = 1, 2, \ldots, T-1$). For (3.8) we consider $P = 3$, $\boldsymbol{\gamma} = (1.2, -0.8, 0.6)'$, and $\Lambda_{1t}, \Lambda_{2t}, \Lambda_{3t} \overset{\text{i.i.d.}}{\sim} \mathrm{N}(0, 1)$ for $t = 1, 2, \ldots, T-1$. This particular simulation setup ensures that the resulting encounter history is neither too dense nor too sparse.

In terms of the prior specification, we set $M = 6000$ and $a_\Psi = b_\Psi = 1.0$. For regression coefficients $\boldsymbol{\theta}$, $\boldsymbol{\zeta}$ and $\boldsymbol{\gamma}$, the prior distributions are given by: $\boldsymbol{\theta} \sim \mathrm{N}(\mathbf{0}, 100\mathbf{I}_2)$, $\boldsymbol{\zeta} \sim \mathrm{N}(\mathbf{0}, 100\mathbf{I}_3)$ and $\boldsymbol{\gamma} \sim \mathrm{N}(\mathbf{0}, 100\mathbf{I}_3)$. For variance parameter σ_u^2, we consider $\sigma_u^2 \sim \mathrm{IG}(q_u, r_u)$ with $q_u = 2.1$

and $r_u = 1.1$. For parameters related to the OU process, we consider the prior specification as: $\sigma_0^2 \sim \mathrm{IG}(q_{\sigma_0}, r_{\sigma_0})$ with $q_{\sigma_0} = 2.1$ and $r_{\sigma_0} = 1.1$; $\alpha \sim \mathrm{N}(\mu_\alpha, \sigma_\alpha^2)$ with $\mu_\alpha = 5$ and $\sigma_\alpha^2 = 100$; $\tau \sim \mathrm{Uniform}(q_\tau, r_\tau)$ with $q_\tau = 0.01$ and $r_\tau = 1.0$; $\sigma^2 \sim \mathrm{IG}(q_\sigma, r_\sigma)$ with $q_\sigma = 2.1$ and $r_\sigma = 1.1$. Our prior specification reflects vague prior distributions relative to scale of the simulated data.

For the MCMC implementation, we run three chains in parallel each with a total of 150,000 iterations. For each Markov chain we discard the first 100,000 iterations as burn-in and draw inference based on every fifth remaining samples. The convergence of the Markov chain to the stationary distribution is assessed by visualizing trace plots of the sample chains and by Gelman and Rubin's diagnostic (Brooks and Gelman (1998)) using the `coda` package (Plummer et al. (2006)) in R (R Core Team (2018)). In this case, visual inspection of the trace plots do not suggest lack of convergence for any model parameters. For Gelman and Rubin's diagnostic the upper 95% confidence limits of the potential scale reduction factors for all model parameters are less than 1.07.

Table 1 provides posterior summary statistics for selected model parameters along with the corresponding true values. It can be seen that the 95% credible intervals (CIs) of parameters related to the parameter model and the OU process cover the true values in all cases. In addition, we can see from Table 1 that the true values for mean stopover duration S and total stopover population size N are recovered. For entrance probabilities β_{t-1} and time-specific population size N_t, Figure 2 graphically presents their 95% CIs along with the corresponding true values, from which we can conclude that all true values are recovered. Figure 3 plots the 95% CIs of capture probabilities p_t along with their corresponding true values, showing that true values of p_t are recovered.

For goodness-of-fit assessment, the Bayesian p-value for the JS model and the OU process is 0.44 and 0.47, respectively. Hence, these p-values do not suggest any lack-of-fit for either the JS model or the use of OU process. To summarize, this simulation suggests that we are able to correctly estimate parameters in our proposed model.

TABLE 1

Posterior summary statistics for selected parameters in the semiparametric JS model with individual heterogeneity for the simulated example. Note that S is the mean stopover duration averaged over all captured individuals

Parameter	μ_{post}	σ_{post}	$Q_{0.025}$	$Q_{0.50}$	$Q_{0.975}$	Truth
ζ_1	0.99554	0.02622	0.94417	0.99554	1.04677	1.0
ζ_2	−0.90057	0.02552	−0.95114	−0.90054	−0.85193	−0.9
ζ_3	0.59768	0.02170	0.55443	0.59761	0.64088	0.6
θ_0	−1.75835	0.03171	−1.82208	−1.75778	−1.69716	−1.8
θ_1	0.31160	0.04431	0.22601	0.31093	0.40151	0.3
γ_1	1.01647	0.25393	0.52347	1.01589	1.51425	1.2
γ_2	−1.02342	0.27839	−1.56250	−1.02357	−0.46578	−0.8
γ_3	0.47463	0.27251	−0.05276	0.47317	1.00610	0.6
σ_0^2	1.43562	0.16827	1.14986	1.41439	1.80354	1.37
α	0.17367	0.13093	−0.09156	0.17416	0.44537	0.2
τ	0.18815	0.01471	0.15982	0.18725	0.21899	0.19
σ^2	0.36191	0.02886	0.31157	0.35991	0.42282	0.36
σ_u^2	0.56438	0.38001	0.18654	0.46427	1.58137	0.25
S	7.05	0.04	6.98	7.05	7.13	7.13
N	3032	77.53	2871	3033	3175	3000

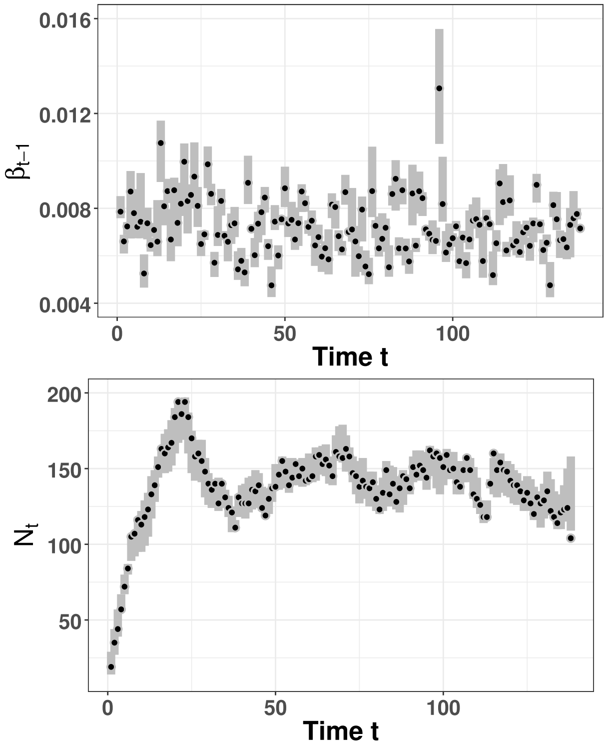

FIG. 2. *Plot of pointwise* 95% *credible intervals and true values for entrance probabilities β_{t-1} and population sizes N_t for $t = 1, 2, \dots, T$ in the simulated example. Note that the upper and lower value of each gray vertical line denotes the* 2.5*th and* 97.5*th percentiles of posterior samples, respectively. Also, the solid black circle on each gray line denotes the true value.*

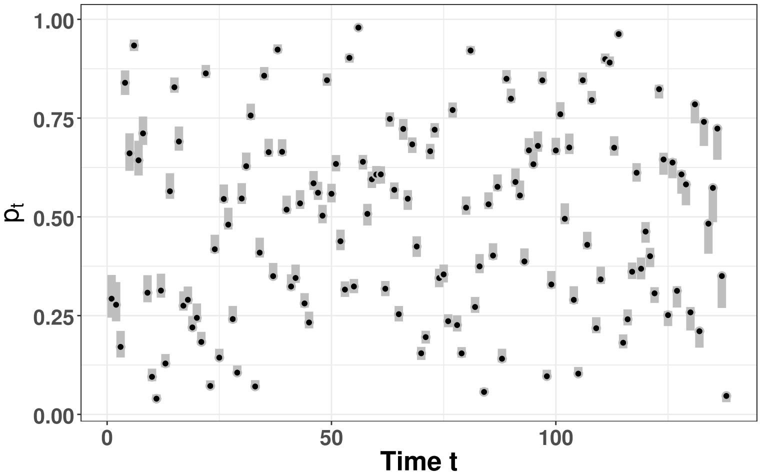

FIG. 3. *Plot of pointwise* 95% *credible intervals and true values for capture probabilities p_t in the simulated example. Note that the upper and lower value of each gray vertical line denotes the* 2.5*th and* 97.5*th percentiles of posterior samples, respectively. Also, the solid black circle on each gray line denotes the true value.*

5. Stopover duration analysis for mallard. We apply the model we propose to the mallard, *Anas platyrhynchos*, stopover duration analysis. The mallard data was collected daily between August 1st and December 16th each year from 2004 to 2011. Each day when a mallard was captured, measurements on body mass and structural size were taken, based on which body condition is calculated (as the ratio of body mass to structural size). For the purpose of illustration, we only present the results from our model using the mallard data collected in 2011, which includes 739 individual mallards caught over $T = 138$ days.

To address three research objectives of our mallard study, we consider the parameter model as

$$\log\left(\frac{p_t}{1-p_t}\right) = \mathbf{\Lambda}_t'\boldsymbol{\zeta} + \sum_{k=1}^{K} u_k |t - \kappa_k|^3, \quad t = 1, 2, \dots, T,$$

$$\log\left(\frac{1-\phi_{i,t}}{\phi_{i,t}}\right) = \theta_0 + \theta_1 \mathtt{Bcond}_{i,t}, \quad t = 1, 2, \dots, T-1,$$

$$\log\left(\frac{\beta_{t-1}}{\beta_{T-1}}\right) = \mathbf{\Lambda}_t'\boldsymbol{\gamma}, \quad t = 1, 2, \dots, T-1.$$

Here, $\mathbf{\Lambda}_t = (\Lambda_{1t}, \Lambda_{2t}, \Lambda_{3t}, \Lambda_{4t}, \Lambda_{5t})'$ and the regressors are related to five daily weather measures—wind direction, wind speed, atmospheric pressure, temperature and rainfall. In particular, Λ_1 reflects variation in rainfall with small values suggesting low precipitation; Λ_2 describes the orthogonal wind (along the NNW-SSE axis) with negative (positive) values suggesting strong NNW (SSE) winds; Λ_3 represents dominant wind (along the WSW-ENE axis) with negative (positive) values suggesting strong WSW (ENE) winds; Λ_4 represents temperature deviation with positive (negative) values suggesting temperature higher (lower) than the seasonal norms; Λ_5 reflects atmospheric pressure with smaller values suggesting higher atmospheric pressure. Moreover, $\mathtt{Bcond}_{i,t}$ denotes the body condition for an individual mallard i at day t. Since body condition varies with both individual and time, its change over time is modeled via the OU process discussed in Section 3.2.

In terms of the MCMC implementation, the number of iterations, burn-in and thinning and the prior specification are similar to those in Section 4, except that we use the prior $\sigma_0^2 \sim \mathrm{IG}(q_{\sigma_0}, r_{\sigma_0})$ with $q_{\sigma_0} = 3.0$ and $r_{\sigma_0} = 2.0$ to achieve better mixing for the nuisance parameter σ_0^2. To evaluate the robustness of our results to different choices of K and M, we consider a sensitivity analysis. On the one hand, we consider four different choices of K: $K = 10$, $K = 15$, $K = 20$ and $K = \max\{20, \min(150, \frac{T}{4})\} = 34$. On the other hand, we consider two different scenarios where $M = 3000$ and $M = 5000$. For these choices of K and M, the results are similar. Nevertheless, the mixing of the MCMC chains for parameters related to OU process is better for $K = 15$ than for three other choices of K. As a result, we present the results from our model with $K = 15$ and $M = 3000$. Regarding convergence assessment, visual inspection of the trace plots does not suggest lack of convergence for any model parameters. For the Gelman and Rubin's diagnostic, the upper 95% confidence limits of the potential scale reduction factors for all model parameters are less than 1.07.

For the purpose of interpretation, we conclude that a parameter is significant if its 95% CIs do not cover 0. Table 2 provides posterior summary statistics for some selected model parameters. According to Table 2, the regression coefficients ζ_2 and ζ_5 are significant. Since ζ_2 is positive significant, this suggests that, when holding other weather conditions constant, the chance of capturing a mallard is larger on days with strong SSE winds than on days with strong NNW winds. Moreover, ζ_5 is negative significant, suggesting that, when holding other weather conditions constant, the chance of capturing a mallard is larger on days with low atmospheric pressure than on days with high atmospheric pressure. For entrance probabilities, it is found that Λ_5 has a negative significant impact on the timing of arrival for mallards

TABLE 2

Posterior summary statistics for selected parameters in the semiparametric JS model with individual heterogeneity for the mallard data. Note that S is the mean stopover duration averaged over all captured mallards

Parameter	μ_{post}	σ_{post}	$Q_{0.025}$	$Q_{0.50}$	$Q_{0.975}$
ζ_1	0.01094	0.03558	−0.05949	0.01068	0.07988
ζ_2	0.15518	0.04108	0.07430	0.15499	0.23528
ζ_3	0.03289	0.04389	−0.05321	0.03289	0.11756
ζ_4	−0.12164	0.08199	−0.28170	−0.12137	0.03762
ζ_5	−0.08804	0.04125	−0.16958	−0.08807	−0.00737
θ_0	−0.92251	0.56468	−2.03495	−0.92337	0.20819
θ_1	−0.14132	0.05758	−0.25633	−0.14048	−0.02860
γ_1	−0.42482	0.30720	−1.16999	−0.37924	0.04364
γ_2	−0.13546	0.11436	−0.35944	−0.13399	0.08951
γ_3	0.03264	0.08606	−0.13415	0.03266	0.20231
γ_4	0.08025	0.06702	−0.04833	0.08068	0.21259
γ_5	−0.19530	0.07362	−0.34222	−0.19412	−0.05330
σ_0^2	1.21477	0.74716	0.32119	1.04971	3.30401
α	10.02102	0.24533	9.61002	9.99905	10.57637
τ	0.08509	0.00902	0.06720	0.08483	0.10354
σ^2	0.20948	0.02180	0.17505	0.20635	0.26394
σ_u^2	0.12737	0.04535	0.06647	0.11873	0.23666
S	12.78	0.19	12.43	12.77	13.18
N	1050	63	925	1052	1173

because the 95% CIs of its coefficient are entirely negative. This suggests that, when other weather conditions are held constant, the chance of a mallard arriving at the stopover site is larger on days with low atmospheric pressure than on days with high atmospheric pressure. For the total stopover population size, N, our results suggest that there were about 1050 mallards that used our study area as a stopover site between August 1st and December 16th in 2011, with the corresponding 95% CIs being [925, 1173]. For average stopover duration, S, our results suggest that, on average, mallards spent 12.78 days at our study site during migration, with the corresponding 95% CIs of S being [12.43, 13.18].

From Table 2 we conclude that there is a significant negative impact of body condition on departure probabilities since the 95% CIs of θ_1 are entirely negative. This result lends support to the necessity of incorporating body condition into the model for departure probabilities to account for individual heterogeneity in mallards' departure. In terms of the impact of body condition on departure decisions, our results suggest that mallards in low body condition have a higher propensity to depart than birds in better body condition. This result is somehow opposite to what is expected during stopover, whereby birds refuel fat stores (and increase body condition) in preparation for the next flight (Berthold (2001)). One potential explanation of such a finding could be that, among mallards stopping over at our study site, those experiencing poor refueling opportunities (e.g., due to insufficient food supply or competition), leave the site soon in search of better refueling places so as to not waste more time in a site that is not rewarding (e.g., see Schaub, Jenni and Bairlein (2008)). Our finding surrounding the departure behaviors of mallards corroborates similar findings for migratory birds; for example, see Kuenzi, Moore and Simons (1991) and Yong and Moore (1993).

Figure 4 provides pointwise posterior summary statistics for p_t, the nonparametric part $\boldsymbol{Z}_t\boldsymbol{b}$ and daily population sizes N_t. According to Figure 4, we can conclude that the nonparametric part of our parameter model for capture probability is needed because the 95% CIs for $\boldsymbol{Z}_t\boldsymbol{b}$ exclude 0. Moreover, we note that both p_t and $\boldsymbol{Z}_t\boldsymbol{b}$ demonstrate the same nonlinear trend

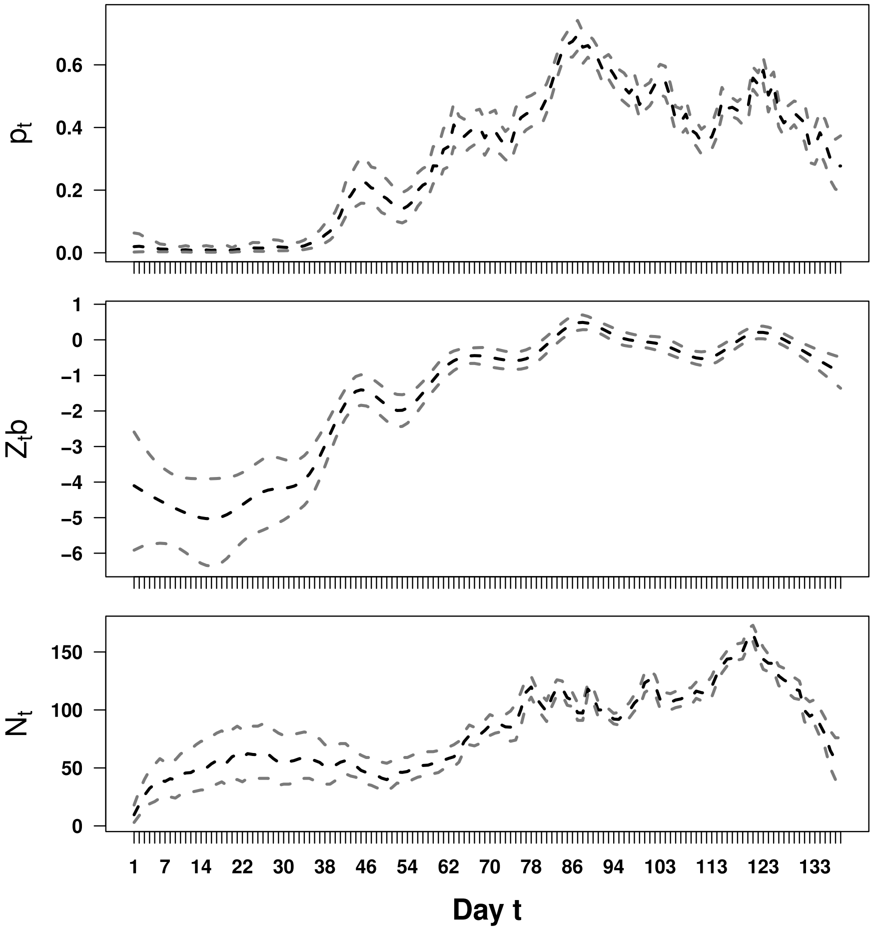

FIG. 4. *Plots of pointwise summary statistics of capture probabilities* p_t, $\mathbf{Z}_t b$ *and daily population sizes* N_t *for mallards. Note that the gray dashed lines are the pointwise* 95% *credible intervals*; *the black dashed line is the posterior mean.*

with respect to time, indicating that capture probability is dominated by the nonparametric part $\boldsymbol{Z}_t\boldsymbol{b}$. This nonlinear time effect on capture probability is probably due to heterogeneous daily sampling efforts that are unaccounted for. For daily stopover population size, N_t, we can see an overall upward trend starting from August to the end of November, when daily stopover population sizes were peaked. Starting in December, there is an overall downward trend in N_t, suggesting that daily stopover population sizes decrease. The upward trend in N_t from August to November is attributed to the arrival of birds at our stopover site. The decrease in daily population sizes in December is driven by the downward trend in entrance probabilities shown in Figure 5 and the departure of birds from our stopover site. The abrupt downward trend in N_t starting in December, contrary to a more gradual upward trend in N_t from August to the end of November, reflects an effect of the phenology of migration, where mallards departed from our stopover site at a faster rate in December than early in the season. This is because mallards cannot afford to stay long at our stopover site to replenish as winter approaches in December.

Lastly, to assess goodness-of-fit for our proposed model, we compute the Bayesian p-values for both the JS model to the mallard data and for the modeling of body condition via the OU process. Using the complete log-likelihood function as the discrepancy function, the Bayesian p-value for the JS model is 0.32 and the Bayesian p-value for the OU process is 0.61. Hence, these p-values do not suggest any lack of fit for either the JS model we propose

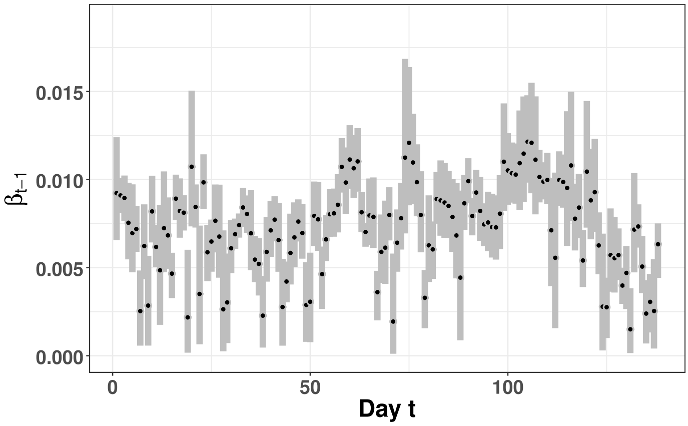

FIG. 5. *Plots of pointwise summary statistics of daily entrance probabilities β_{t-1} for mallards. Note that the upper and lower value of each gray vertical line denotes the* 2.5*th and* 97.5*th percentiles of posterior samples, respectively. Also, the solid black circle on each gray line denotes the posterior mean.*

for mallard data or for the OU process that is used to model the evolution of body conditions over time.

6. Discussion. Of particular importance to strategic management and conservation planning is to understand the impacts of both internal and external factors on the migratory behavior of wildlife. Motivated by mallard migration studies, we propose a stopover model that can accommodate individual heterogeneity in departure, dependence of arrival time on covariates, and nonlinear functional relationships between capture probabilities and external covariates. In particular, we account for individual heterogeneity in departure via a time-varying continuous individual covariate that is assumed to follow an OU process. The semiparametric regression for capture probabilities, using low rank thin-plate splines, allows us to identify both linear and nonlinear covariate effects on capture probabilities. Collectively, these developments provide a framework with increased biological relevance that can be applied to various migration studies of wildlife. This has big promise for migration and movement ecology, as the length and behavior of birds during stopover are instrumental for overall speed of migration with consequences at both the individual and the population level.

Regarding three important research objectives associated with mallard migration studies, our results confirm that there is individual heterogeneity in mallards' departure due to body condition. Specifically, mallards in low body condition are more likely to depart than birds in better body condition, which may indicate that mallards tend to leave the site as soon as possible if they cannot refuel efficiently, likely in search for better refueling places, but also to avoid wasting too much time in a stopover area that does not allow them to replenish energy reserves quickly. Second, our results suggest that there were about 1050 (95% CIs: $[925, 1173]$) mallards that visited our stopover site from August 1st to December 16th in 2011. On average, mallards spent about 13 days at our stopover site for energy replenishment before leaving our stopover site to continue their migratory flights. Our results also suggest a temporal trend in daily stopover population sizes, with an upward trend from August 1st to the end of November and a downward trend in December.

For the impacts of external weather covariates, our results indicate that mallards are more likely to be captured on days with strong SSE winds and low atmospheric pressure. This is expected, as general migration directions from the area lay in the S to W sectors (van Toor et al. (2018b)) and strong SSE winds then are opposite to tailwinds and provide no flight assistance for mallards. Furthermore, falling atmospheric pressure is also a predictor of harsh or inclement weather, under which mallards already present in the stopover site are less likely to depart and, hence, are more likely to be captured. Regarding the timing of arrival, our results suggest that mallards are more likely to arrive at our stopover site on days with low atmospheric pressure than on days with high atmospheric pressure. As low atmospheric pressure is linked to bad weather conditions, migrating mallards passing the stopover site would likely be more prone to temporarily cease migration and land in the study area than to continue migration, again increasing likelihood of being captured.

As shown by our results, it is advantageous to consider a semiparametric regression model for capture probabilities due to the dominant nonlinear time effect after accounting for weather covariates. This flexible modeling of nonlinear time effects on capture outcomes allows us to address possible heterogeneous daily sampling efforts that are unaccounted for. Despite the flexibility of semiparametric regression, different basis functions can be used. For example, Gimenez et al. (2006) consider Bayesian penalized splines that utilize truncated polynomial basis to model survival probabilities in the CJS model. Similarly, Bonner, Thomson and Schwarz (2009) consider a Bayesian semiparametric regression for survival probability in the CJS model that features B-spline basis functions. Instead of fixing the number and location of knot points, they consider a Bayesian adaptive approach where a reversible jump MCMC algorithm is employed to explore splines with different knot configurations. In addition, Bonner and Schwarz (2011) develop a model for the data that arises from a Peterson-type mark-recapture experiment, where B-spline basis functions are considered for smoothing population size estimates. Among these choices we apply Bayesian low-rank thin-plate splines to smooth capture probabilities in the JS model to achieve better mixing of the MCMC chains (Crainiceanu, Ruppert and Wand (2005)).

In principle, our model can be estimated using freely available software programs, such as BUGS, following the general framework presented by Schofield and Barker (2011). Nevertheless, Bayesian estimation of our model using these freely available software programs can be inefficient due to lack of well-crafted MCMC algorithms. In contrast, the MCMC algorithm that we develop for our model does not require any tedious tuning, making our methodology more computationally efficient, especially in the case of large datasets. Although our model is application-specific, it is worth pointing out that our well-tailored sampling algorithm, for example, the block updating scheme for latent state variables, can be applied to other variants of the JS model with straightforward adaption.

Acknowledgments. The authors would like to thank the Editor, Thomas Brendan Murphy, Associate Editor and anonymous referees for providing valuable comments that have helped strengthen this manuscript.

Funding. This research was partially supported by the U.S. National Science Foundation (NSF) under NSF Grant SES-1853096. The duck surveillance data was supported by the Swedish Research Council (2011-3568, 2015-03877). This is contribution number 318 from Ottenby Bird Observatory.

SUPPLEMENTARY MATERIAL

Supplement A (DOI: 10.1214/20-AOAS1421SUPPA; .pdf). Additional details of full conditional distributions and the well-tailored MCMC sampling algorithm for selected model parameters.

Supplement B (DOI: 10.1214/20-AOAS1421SUPPB; .zip). Code and data sets.

REFERENCES

BENGTSSON, D., AVRIL, A., GUNNARSSON, G., ELMBERG, J., SÖDERQUIST, P., NOREVIK, G., TOLF, C., SAFI, K., FIEDLER, W. et al. (2014). Movements, home-range size and habitat selection of mallards during autumn migration. *PLoS ONE* **9** e100764. https://doi.org/10.1371/journal.pone.0100764

BERTHOLD, P. (2001). *Bird Migration*: *A General Survey*, 2nd ed. Oxford Univ. Press, New York.

BONNER, S. J. and SCHWARZ, C. J. (2006). An extension of the Cormack–Jolly–Seber model for continuous covariates with application to *Microtus pennsylvanicus*. *Biometrics* **62** 142–149. MR2226567 https://doi.org/10.1111/j.1541-0420.2005.00399.x

BONNER, S. J. and SCHWARZ, C. J. (2009). Bayesian estimation of abundance for open populations with covariate dependent capture and survival probabilities draft. Available at http://www.simon.bonners.ca/PDFS/cts_covariate_js_0809.pdf.

BONNER, S. J. and SCHWARZ, C. J. (2011). Smoothing population size estimates for time-stratified mark-recapture experiments using Bayesian P-splines. *Biometrics* **67** 1498–1507. MR2872401 https://doi.org/10.1111/j.1541-0420.2011.01599.x

BONNER, S., THOMSON, D. and SCHWARZ, C. (2009). Time-varying covariates and semi-parametric regression in capture–recapture: An adaptive spline approach. In *Modeling Demographic Processes in Marked Populations*, *Environmental and Ecological Statistics* 3 (D. Thomson, E. Cooch and M. Conroy, eds.) 657–675. Springer, New York.

BROOKS, S. P., CATCHPOLE, E. A. and MORGAN, B. J. T. (2000). Bayesian animal survival estimation. *Statist*. *Sci*. **15** 357–376. MR1847773 https://doi.org/10.1214/ss/1009213003

BROOKS, S. P. and GELMAN, A. (1998). General methods for monitoring convergence of iterative simulations. *J*. *Comput*. *Graph*. *Statist*. **7** 434–455. MR1665662 https://doi.org/10.2307/1390675

CRAINICEANU, C. M., RUPPERT, D. and WAND, M. P. (2005). Bayesian analysis for penalized spline regression using Win BUGS. *J*. *Stat*. *Softw*. **14** 1–24.

CRAMP, S. and SIMMONS, K. (1977). *The Birds of the Western Palearctic*. *Handbook of the Birds of Europe*, *the Middle East and North Africa*. *Ostrich to Ducks* **1**. Oxford Unversity Press, Oxford.

DINSMORE, S. J. and COLLAZO, J. A. (2003). The influence of body condition on local apparent survival of spring migrant sanderlings in coastal North Carolina. *Condor* **105** 465–473.

FILIPE, P. A., BRAUMANN, C. A., BRITES, N. M. and ROQUETE, C. J. (2010). Modelling animal growth in random environments: An application using nonparametric estimation. *Biom*. *J*. **52** 653–666. MR2757011 https://doi.org/10.1002/bimj.200900273

FINCH, S. (2004). Ornstein–Uhlenbeck process. Available at https://oeis.org/A249417/a249417.pdf.

GELMAN, A. (2003). A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. *Int*. *Stat*. *Rev*. **71** 369–382.

GIMENEZ, O., CRAINICEANU, C., BARBRAUD, C., JENOUVRIER, S. and MORGAN, B. J. T. (2006). Semiparametric regression in capture–recapture modeling. *Biometrics* **62** 691–698. MR2247195 https://doi.org/10.1111/j.1541-0420.2005.00514.x

GUNNARSSON, G., LATORRE-MARGALEF, N., HOBSON, K. A., WILGENBURG, S. L. V., ELMBERG, J., OLSEN, B., FOUCHIER, R. A. M. and WALDENSTRÖM, J. (2012). Disease dynamics and bird migration–linking mallards *Anas platyrhynchos* and subtype diversity of the influenza A virus in time and space. *PLoS ONE* **7** e35679. https://doi.org/10.1371/journal.pone.0035679

JOLLY, G. M. (1965). Explicit estimates from capture–recapture data with both death and immigration-stochastic model. *Biometrika* **52** 225–247. MR0210227 https://doi.org/10.1093/biomet/52.1-2.225

KAISER, A. (1995). Estimating turnover, movements and capture parameters of resting passerines in standardized capture–recapture studies. *J*. *Appl*. *Stat*. **22** 1039–1048.

KING, R. and BROOKS, S. P. (2002). Bayesian model discrimination for multiple strata capture–recapture data. *Biometrika* **89** 785–806. MR1946510 https://doi.org/10.1093/biomet/89.4.785

KING, R., MORGAN, B., GIMENEZ, O. and BROOKS, S. (2010). *Bayesian Analysis for Population Ecology*. CRC Press/CRC, Boca Raton, FL.

KUENZI, A. J., MOORE, F. R. and SIMONS, T. R. (1991). Stopover of Neotropical landbird migrants on East Ship Island following trans-Gulf migration. *Condor* **93** 869–883.

LATORRE-MARGALEF, N., GUNNARSSON, G., MUNSTER, V. J., FOUCHIER, R. A., OSTERHAUS, A. D., ELMBERG, J., OLSEN, B., WALLENSTEN, A., HAEMIG, P. D. et al. (2009). Effects of influenza A virus infection on migrating mallard ducks. *Proc*. *R*. *Soc*. *Lond*., *B Biol*. *Sci*. **276** 1029–1036.

LUNN, D. J., THOMAS, A., BEST, N. and SPIEGELHALTER, D. (2000). WinBUGS—A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat*. *Comput*. **10** 325–337.

LYONS, J. E., KENDALL, W. L., ROYLE, J. A., CONVERSE, S. J., ANDRES, B. A. and BUCHANAN, J. B. (2016). Population size and stopover duration estimation using mark-resight data and Bayesian analysis of a superpopulation model. *Biometrics* **72** 262–271. MR3500595 https://doi.org/10.1111/biom.12393

MATECHOU, E. (2010). Applications and extensions of capture–recapture stop-over models. Ph.D. thesis, The University of Kent.

MATECHOU, E., DENNIS, E. B., FREEMAN, S. N. and BRERETON, T. (2014). Monitoring abundance and phenology in (multivoltine) butterfly species: A novel mixture model. *J. Appl. Ecol.* **51** 766–775.

MATECHOU, E., NICHOLLS, G. K., MORGAN, B. J. T., COLLAZO, J. A. and LYONS, J. E. (2016). Bayesian analysis of Jolly–Seber type models. *Environ. Ecol. Stat.* **23** 531–547. MR3568110 https://doi.org/10.1007/s10651-016-0352-0

NEWTON, I. (2010). *The Migration Ecology of Birds*. Academic Press, London.

NICHOLS, J. D., RUNGE, M. C., JOHNSON, F. A. and WILLIAMS, B. K. (2007). Adaptive harvest management of North American waterfowl populations: A brief history and future prospects. *J. Ornithol.* **148** 343–349.

PLEDGER, S., EFFORD, M., POLLOCK, K., COLLAZO, J. and LYONS, J. (2009). Stopover duration analysis with departure probability dependent on unknown time since arrival. In *Environmental and Ecological Statistics* (D. Thomson, E. Cooch and M. Conroy, eds.) 349–363. Springer, Berlin.

PLUMMER, M., BEST, N., COWLES, K. and VINES, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6** 7–11.

RICE, S. M., COLLAZO, J. A., ALLDREDGE, M. W., HARRINGTON, B. A., LEWIS, A. R. and SANDERCOCK, B. (2007). Local annual survival and seasonal residency rates of semipalmated sandpipers (*Calidris pusilla*) in Puerto Rico. *Auk* **124** 1397–1406.

ROYLE, J. A. and DORAZIO, R. M. (2008). *Hierarchical Modeling and Inference in Ecology*: *The Analysis of Data from Populations*, *Metapopulations and Communities*. Academic Press, San Diego, CA.

RUPPERT, D., WAND, M. P. and CARROLL, R. J. (2003). *Semiparametric Regression*. *Cambridge Series in Statistical and Probabilistic Mathematics* **12**. Cambridge Univ. Press, Cambridge. MR1998720 https://doi.org/10.1017/CBO9780511755453

SCHAUB, M., JENNI, L. and BAIRLEIN, F. (2008). Fuel stores, fuel accumulation, and the decision to depart from a migration stopover site. *Behav. Ecol.* **19** 657–666.

SCHOFIELD, M. R. and BARKER, R. J. (2011). Full open population capture–recapture models with individual covariates. *J. Agric. Biol. Environ. Stat.* **16** 253–268. MR2818548 https://doi.org/10.1007/s13253-010-0052-4

SCHWARZ, C. J. and ARNASON, A. N. (1996). A general methodology for the analysis of capture–recapture experiments in open populations. *Biometrics* **52** 860–873. MR1411736 https://doi.org/10.2307/2533048

SEBER, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika* **52** 249–259. MR0210228 https://doi.org/10.1093/biomet/52.1-2.249

TANNER, M. A. and WONG, W. H. (1987). The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* **82** 528–550. MR0898357

R CORE TEAM (2018). *R*: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

VAN TOOR, M. L., AVRIL, A., WU, G., HOLAN, S. H. and WALDENSTRÖM, J. (2018a). As the duck flies—Estimating the dispersal of low-pathogenic avian influenza viruses by migrating mallards. *Front. Ecol. Evol.* **6** 208.

VAN TOOR, M. L., KRANSTAUBER, B., NEWMAN, S. H., PROSSER, D. J., TAKEKAWA, J. Y., TECHNITIS, G., WEIBEL, R., WIKELSKI, M. and SAFI, K. (2018b). Integrating animal movement with habitat suitability for estimating dynamic migratory connectivity. *Landsc. Ecol.* **33** 879–893.

VON BERTALANFFY, L. K. (1938). A quantitative theory of organic growth (inquiries on growth laws. II). *Hum. Biol.* **10** 181–213.

WILLIAMS, B., NICHOLS, J. and CONROY, M. (2002). *Analysis and Management of Animal Populations*. Academic Press, San Diego, CA.

WU, G., HOLAN, S. H., AVRIL, A. and WALDENSTRÖM, J. (2021). Supplement to "A Bayesian semiparametric Jolly–Seber model with individual heterogeneity: An application to migratory mallards at stopover." https://doi.org/10.1214/20-AOAS1421SUPPA, https://doi.org/10.1214/20-AOAS1421SUPPB

YONG, W. and MOORE, F. R. (1993). Relation between migratory activity and energetic condition among thrushes (Turdinae) following passage across the Gulf of Mexico. *Condor* **95** 934–943.