

ARTICLE

A hierarchical N-mixture model to estimate behavioral variation and a case study of Neotropical birds

Alison Ke¹  | Rahel Sollmann^{1,2}  | Luke O. Frishkoff³  | Daniel S. Karp¹

¹Department of Wildlife, Fish, and Conservation Biology, University of California, Davis, California, USA

²Department of Ecological Dynamics, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

³Department of Biology, University of Texas at Arlington, Arlington, Texas, USA

Correspondence

Alison Ke

Email: ake@ucdavis.edu

Funding information

National Geographic Society, Grant/Award Number: NGS-67543R-20; National Science Foundation, Grant/Award Number: GRFP

Handling Editor: John M. Marzluff

Abstract

Understanding how and why animals use the environments where they occur is both foundational to behavioral ecology and essential to identify critical habitats for species conservation. However, some behaviors are more difficult to observe than others, which can bias analyses of raw observational data. To our knowledge, no method currently exists to model how animals use different environments while accounting for imperfect behavior-specific detection probability. We developed an extension of a binomial N-mixture model (hereafter the behavior N-mixture model) to estimate the probability of a given behavior occurring in a particular environment while accounting for imperfect detection. We then conducted a simulation to validate the model's ability to estimate the effects of environmental covariates on the probabilities of individuals performing different behaviors. We compared our model to a naïve model that does not account for imperfect detection, as well as a traditional N-mixture model. Finally, we applied the model to a bird observation data set in northwest Costa Rica to quantify how three species behave in forests and farms. Simulations and sensitivity analyses demonstrated that the behavior N-mixture model produced unbiased estimates of behaviors and their relationships with predictor variables (e.g., forest cover, habitat type). Importantly, the behavior N-mixture model accurately characterized uncertainty, unlike the naïve model, which often suggested erroneous effects of covariates on behaviors. When applied to field data, the behavior N-mixture model suggested that Hoffmann's woodpecker (*Melanerpes hoffmannii*) and Inca dove (*Columbina inca*) behaved differently in forested versus agricultural habitats, while turquoise-browed motmot (*Eumomota superciliosa*) did not. Thus, the behavior N-mixture model can help identify habitats that are essential to a species' life cycle (e.g., where individuals nest, forage) that nonbehavioral models would miss. Our model can greatly improve the appropriate use of behavioral survey data and conclusions drawn from them. In doing so, it provides a valuable path forward for assessing the conservation value of alternative habitat types.

KEYWORDS

agricultural systems, Bayesian hierarchical model, Bayesian methods, behavior change, behavior survey, N-mixture model, tropical birds

INTRODUCTION

Species responses to environmental conditions are most often assessed by measuring effects on their occurrences or abundances (e.g., Hatfield et al., 2018; Newbold et al., 2013). Yet changes in incidence provide little information about how species use the environments where they occur (Gilroy & Edwards, 2017; Ortega-Álvarez et al., 2021). For example, a species regularly detected in agricultural habitats could simply be passing through, as opposed to actively using agricultural habitats to forage, reproduce, and complete its life cycle (Vickery et al., 2001). Variation in the behavioral use of different environments can have cascading implications for individual fitness and population persistence (e.g., Luck, 2002; Lyons, 2005). For example, a dense habitat may not be necessary for an individual to survive, but it is still important for population persistence because it contains the necessary substrates for reproduction (e.g., nesting, mating). Therefore, changes in behavior can serve as a warning signal that a population is in trouble (Berger-Tal et al., 2016), and understanding the fitness effects of animals behaving differently in different habitats is critical to conserve species effectively. In contrast, if we base costly, lengthy management actions solely on incidence, without understanding how a species uses that habitat, we may fail to protect habitats that are essential to critical phases of the species' life cycle.

Unfortunately, many behaviors are difficult to observe (Durso et al., 2011), and the detectability of individuals often depends on the behaviors they perform (Crowe & Longshore, 2010). Moreover, the detectability of different behaviors may change depending on the surrounding environment. This may be particularly problematic when changes in behaviors are confounded by changes in their detection probability, such as when the same covariate (e.g., vegetation density) influences behaviors and the probability of detecting both individuals and their behaviors. For example, the probability of detecting bird vocalizations may be similar between open areas (e.g., farms) and dense environments (e.g., forests), whereas the probability of detecting behaviors that are observed visually may decline. Meanwhile, it is possible that a species performs behaviors detected visually, such as foraging, more frequently in denser environments. Without accounting for behavior-specific detection, we would risk falsely concluding that open environments are more beneficial to this species because we observe it foraging there more often. To our knowledge, no method currently exists to model how different habitats support different behaviors while accounting for behavior-specific detection probability.

Ecologists have long grappled with the problem of imperfect detection when modeling the abundance of wildlife (Hilborn et al., 1976; Kéry & Royle, 2008; Seber, 1982).

Failing to account for variation in detection (e.g., among species, observers, habitats) can bias estimates of abundances (Kellner & Swihart, 2014). One method frequently used to account for imperfect detection in abundance estimation is the N-mixture model (Kéry, 2018; Royle, 2004). N-mixture models estimate abundance and detection using spatially and temporally replicated surveys where the number of individuals is counted. They assume that populations are closed, such that the same number of individuals is present during each visit to a site, and that all individuals have the same detection probability. If there is unmodeled heterogeneity in detection (such as variable detection probabilities between behaviors), N-mixture models are known to underestimate abundance (Kéry & Royle, 2015). Therefore, modeling behavior-specific detection probabilities could potentially account for behavior-driven heterogeneity (e.g., between individuals, age-sex classes, populations) in detection.

Given a data set of species observations with spatial and temporal replication, where behaviors are noted for all detected individuals, a conceptually straightforward way to account for behavior-specific detection probabilities is to extend the N-mixture model to estimate the abundance of each behavior separately, using repeated counts of behaviors. A species' total abundance is then the sum of the abundances of each behavior. This approach allows each behavior to have its own detection probability that can be modeled as a function of covariates. It also allows for modeling of covariate effects on the prevalence of a given behavior. However, a major challenge with this method is that the N-mixture model assumes populations are closed, which may be less probable with behavior-specific abundances than overall abundances. Specifically, it is quite possible that, across multiple visits to a single site, the same number of individuals is present but they perform different behaviors; we refer to this as a violation of the behavioral closure assumption.

Here we develop a behavior N-mixture model that estimates the effects of environmental covariates on behavior probabilities (i.e., the probability that an individual will perform a certain behavior) while accounting for imperfect detection of individuals and behaviors. This method adds another conditional layer to the traditional N-mixture model. Specifically, it estimates the probabilities that individual animals will use one type of habitat for a behavior (e.g., foraging, vocalizing) while accounting for the imperfect and variable detection of those behaviors. We conduct a simulation study to understand how variation in the detection, abundance, and violation of the behavioral closure assumption affects our ability to estimate relationships between environmental covariates and behavior probabilities. To assess performance, we compare our model to a naïve model that estimates behavior probabilities without accounting for imperfect

detection, as well as a traditional N-mixture model that accounts for overall imperfect detection but that estimates neither behavior probabilities nor behavior-specific detection probabilities. Finally, we apply our model to a 4-year data set of bird surveys in Costa Rica, in which behaviors (e.g., vocalizing, eating, passive behaviors) were noted during repeated point counts in agricultural and forested habitats.

MATERIALS AND METHODS

Formulation of the behavior N-Mixture model

We extended the N-mixture model framework to quantify how environmental covariates influence the relative frequency of behaviors performed (behavior probability) while accounting for variation in detection probability among different behaviors in different environments (Kéry, 2018; Royle, 2004). Under certain assumptions, our new model can also be used as a typical N-mixture model to estimate abundance, though in other cases doing so can be problematic (see subsequent discussion). In typical N-mixture models, the number of individuals observed (Y) at a site (j) and a visit (k) is modeled based on abundance and detection processes such that

$$Y_{j,k} \sim \text{Binomial}(N_j, P_{j,k}),$$

where N is the true number of individuals and P is the per-individual detection probability. P can be modeled as a function of site- and visit-specific covariates. Site-specific abundances, N_j , are modeled as a random count variable (e.g., with a Poisson probability distribution), and expected abundance can be modeled as functions of site-specific covariates. For the behavior N-mixture model, to model the counts (Y) and proportion (π) of individuals performing the b th behavior, we modified the typical N-mixture model such that

$$Y_{j,b,k} \sim \text{Binomial}(N.b_{j,b}, P_{j,b,k})$$

$$N.b_{j,b} \sim \text{Poisson}(\lambda_j \times \pi_{j,b})$$

$$N_j = \sum_{b=1}^B N.b_b.$$

where $Y_{j,b,k}$ is the number of individuals observed performing a certain behavior (b) at a site (j) and a visit (k), $P_{j,b,k}$ is the detection probability that an individual will

perform a certain behavior at a particular site and visit, $N.b_{j,b}$ is the number of individuals performing a certain behavior at a site, $\pi_{j,b}$ is the probability that an individual will perform a certain behavior at a site, and, λ_j is the expected value of the total abundance (N_j) at a site. Note that $\pi_{j,b}$ and $N.b_{j,b}$ are indexed by site and behavior, which implies closure of relative behavior frequencies between visits (i.e., assuming the same number of individuals performs each behavior during each visit). Expected abundance λ_j can be modeled as

$$\log(\lambda_j) = \alpha_0 + \boldsymbol{\alpha}'\mathbf{X}_j,$$

where α_0 is the intercept and $\boldsymbol{\alpha}'$ is a vector of coefficients that are multiplied by the covariates \mathbf{X} .

$B.p_{j,b}$ is modeled as a function of covariates using multinomial logistic regression. For the first behavior,

$$\pi_{j,1} = \frac{1}{1 + \sum_{b=2}^B e^{\beta_{0b} + \boldsymbol{\beta}'_b \mathbf{Z}_j}},$$

and for the remaining behaviors,

$$\pi_{j,b} = \frac{e^{\beta_{0b} + \boldsymbol{\beta}'_b \mathbf{Z}_j}}{1 + \sum_{b=2}^B e^{\beta_{0b} + \boldsymbol{\beta}'_b \mathbf{Z}_j}}.$$

The first behavior serves as a reference category that the other behaviors are compared to; any category can be used as the reference. B is the number of behavior categories. In the multinomial logistic regression, β_{0b} is a behavior-specific intercept and $\boldsymbol{\beta}'$ is the vector of coefficients that is multiplied by the covariates \mathbf{Z} .

Finally, the detection probability that an individual will perform a certain behavior at a given site and visit ($P_{j,b,k}$) can be modeled as a function of site- and visit-specific covariates:

$$\text{logit}(P_{j,b,k}) = \gamma_0 + \boldsymbol{\gamma}'_b \mathbf{V}_j,$$

where γ_0 is the intercept and $\boldsymbol{\gamma}'$ is the vector of behavior-specific coefficients that are multiplied by the covariates \mathbf{V} .

Simulation study

We evaluated the performance of the behavior N-mixture model under five combinations of mean expected abundance (λ) and detection probability (P). For each

scenario, we applied the behavior N-mixture model to 100 sets of simulated observations of behaviors and abundances for one species across 50 sites (indexed by j), with five visits to each site (indexed by k). In all scenarios, individuals could perform one of three behaviors (indexed by b), with intercepts of the behavior probabilities fixed at 0.6, 0.15, and 0.25 ($\beta_0 = [0, -1.38, -0.87]$). The intercepts of the behavior probabilities must sum to 1, and the first behavior served as the reference level. First, we tested model performance with a detection probability intercept fixed at 0.2 (a realistic value for bird communities) (Karp et al., 2018) and three levels of abundance intercepts representing 1, 2.7, and 7.4 individuals per site. Second, we tested model performance with an abundance intercept fixed at 2.7 individuals per site ($\alpha_0 = 1$) and two additional levels of detection probability intercepts: 0.5 and 0.8 ($\gamma_0 = 0$ and 1.38).

In all cases, the abundance intercept α_0 was added to the product of one site-specific covariate \mathbf{X} and an abundance coefficient α' . The behavior probability intercept β_0 was added to the product of a behavior-specific coefficient β'_b and one site-specific covariate \mathbf{Z} for two behaviors, while the first behavior served as the reference level. The number of individuals performing each behavior at each site and visit was determined using a multinomial random distribution with size $n = N_j$ and cell probabilities $\boldsymbol{\pi}_{1,3,j}$. This case reflects behavioral closure, where the numbers of individuals performing each behavior at each site is constant across visits. Finally, we generated detection probabilities by adding the intercept γ_0 and the product of a behavior-specific detection coefficient and the covariate \mathbf{Z} . For each replicate, the abundance coefficient and detection intercept were drawn from a normal distribution with mean 0 and variance 0.25, while the behavior-specific coefficients were drawn from a normal distribution with mean 0 and variance 1. We wanted to explore more of the parameter space of the behavior-specific coefficients in contrast to abundance and detection parameters, which we aimed to control for in each simulation scenario. Site-specific covariates were also drawn from a normal distribution with mean 0 and variance 1, so that we did not have to standardize environmental covariates prior to analysis.

Assuming behavioral closure could be problematic in practice because animals can change behaviors frequently. To explore how the behavior N-mixture model performs under complete violation of behavioral closure, we applied it to an additional data set paired with each simulated data set described earlier, where the number of individuals performing each behavior was resampled from the same multinomial distribution for each visit. This way, the number of individuals was constant across visits, but the number of them engaging in each behavior

could change. Next, to understand model performance under intermediate levels of violation of the behavioral closure assumption, we applied the behavior N-mixture model to more paired data sets where the numbers of individuals performing each behavior each visit were determined by resampling 0%, 25%, 50%, 75%, and 100% of the total abundance while holding the detection intercept at 0.2 and abundance intercept at 2.7 (see earlier discussion). When the behavioral closure assumption was violated, we expected abundance to be overestimated and detection probability to be underestimated. This is because when an individual is observed performing different behaviors between visits, it contributes to overall abundance estimates multiple times.

Alternative models

To compare how the behavior N-mixture model estimated total abundance to an alternative model, we applied a traditional N-mixture model to every simulated data set where we retained covariates for estimating abundances and detection probabilities. Under behavioral closure, we expected the traditional N-mixture model to underestimate abundance due to unmodeled variation in detection probabilities. When behavioral closure was violated, we expected the behavior N-mixture model to produce more biased abundance estimates than the traditional N-mixture model.

We also sought to compare the behavior N-mixture model's estimates of behavior parameters to a naïve model that did not account for imperfect detection. To do so, we applied a multinomial regression model to every simulated data set, where we aggregated behavioral observations across visits for each simulation replicate. This was the most straightforward way to model behavior frequencies without accounting for imperfect detection. We expected the naïve model to produce biased relationships with covariates, especially when the same covariate affected detection and behavior.

We assessed model estimates for abundance (N , α_0 , α'), detection ($P_{j,b,k}$, γ_0 , γ'_b), and behavior ($\pi_{j,b}$, β_0 , β'_b) parameters. For each replicate, we calculated the absolute bias of the posterior means and 95% Bayesian credible interval (BCI) coverage, which represents how often the range between the 2.5th and 97.5th percentiles of the posterior distribution contains the true value. We compared bias and BCI coverage across levels of mean detection, abundance, and violation of behavioral closure. We also compared bias and BCI coverage of parameters between the behavior N-mixture model, traditional N-mixture model, and naïve multinomial regression model.

Implementation

We implemented all models in R version 4.0.0 using the *nimble* package, which runs Markov chain Monte Carlo (MCMC) algorithms (de Valpine et al., 2017; RC Team, 2013). For each simulation replicate, we ran three chains starting at random initial values and 5000 burn-in iterations. We included 150,000, 100,000, and 50,000 burn-in iterations and thinning rates of 75, 50, and 25 for the behavior N-mixture, traditional N-mixture, and naïve multinomial regression models, respectively. We treated models as having converged if all chains for abundance, detection probabilities, behavior probabilities, and intercept and slope parameters had Gelman–Rubin statistics ≤ 1.1 (Gelman et al., 2004). If any chains did not converge, we excluded the entire replicate (i.e., all models fit to that data set) from further analysis.

Case study: Effects of agriculture on bird behavior in northwest Costa Rica

We applied the behavior N-mixture model to behavioral observations of Hoffmann's woodpecker (*Melanerpes hoffmannii*), Inca dove (*Columbina inca*), and turquoise-browed motmot (*Eumomota superciliosa*) in adjacent agricultural and forested sites in northwest Costa Rica. All species are regularly observed both in forest and agricultural habitats; our objective was to determine whether species changed behavior frequencies between habitats.

At 20 forest-adjointing farms and five protected areas, birds were surveyed at six sites each, half in agricultural sites and half in forest ($N = 150$ sites total). At each site, the same expert observer (J. Zook) surveyed all birds seen or heard in 20-min, 50-m fixed-radius point counts in May–July from 2016 to 2019. Half of the point counts were sampled three times within a 1-week period, and the other half were surveyed once to increase spatial replication while still being able to estimate detection probabilities. One farm or protected area (six sites) was surveyed each day, beginning at sunrise and continuing for ~ 5 h. The observer recorded species identity, number of individuals observed, time of day, and ambient noise, which we considered to be noise levels that were above typical background noises and were thought by the observer to interfere with his ability to detect bird vocalizations. Each observation of a species was associated with one of 32 behaviors (Appendix S1: Table S2) that we classified into three categories: vocalizing, eating/foraging, and other. When an individual performed more than one behavior during a point count (10.5% of observations), we randomly selected one of the behaviors from those recorded by the observer.

We modeled the annual abundance of each species at each site as a function of an intercept, a coefficient multiplied by the fraction of forest cover within 50 m, and random effects for farm, point, and year. We modeled behavior probability as a function of a behavior-specific intercept and a behavior-specific coefficient multiplied by a binary habitat covariate (agriculture vs. forest). We modeled detection probability as a function of an intercept, a behavior-specific coefficient multiplied by a binary noise variable (no noise vs. noise, e.g., machinery, cicadas), a behavior-specific coefficient multiplied by the binary habitat covariate, and a coefficient multiplied by the time of day.

We considered the composition of behaviors to differ between habitats when the BCIs for any behavior coefficients did not include zero. Because the values of behavior coefficients are relative to the reference behavior, their values do not necessarily reflect how a behavior probability differs between habitats. Thus, we considered a behavior probability to be significantly different between forested and agricultural habitats when the BCI for the predicted difference in the behavior probability between habitats did not include zero. We compared estimates of behavior probabilities and coefficients to the naïve model and compared estimates of abundance to the traditional N-mixture model (Appendix S1). For all models, we ran three chains of 5000 burn-in iterations and 150,000 post-burn-in iterations and thinned chains by 75. We implemented models in R version 4.0.0 using the *nimble* package and checked for convergence as described earlier (de Valpine et al., 2017; RC Team, 2013).

RESULTS

Model convergence

Chains for every parameter within the behavior N-mixture, traditional N-mixture, and naïve models converged for 60%–85% of the simulation replicates when the detection intercept was 0.2 and 0.5 across all levels of abundance. When the detection intercept was 0.8, all chains converged in only 14% of replicates (Appendix S1: Table S1). When we adjusted the levels of violation of the behavioral closure assumption, all chains from 15 models (three models applied to five levels of closure violation) converged in 23% of replicates.

Behavior parameters

Overall, we found that the behavior N-mixture model better estimated behavior probabilities and effects of

covariates (e.g., land-use type) on behavior compared to the naïve model. Specifically, the mean estimate of behavior probabilities was 0.96–3.34 times more likely to be less biased under the behavior N-mixture model than under the naïve model across scenarios (Appendix S1: Table S3). When data met the behavioral closure assumption, BCIs from the behavior N-mixture model captured the true value of the behavior probabilities 90%–95% of the time across all scenarios of abundance and detection (Table 1). The naïve model, however, only captured the true value 42%–74% of the time.

The mean absolute bias of estimates of behavior probability and behavior probability coefficients (i.e., the effect of a covariate on each behavior, β'_b) from the behavior N-mixture model were small (−0.001 to 0.03 and −0.01 to 0.21, respectively; true β'_b values ranged from −4.11 to 3.06) (Table 1). When the detection intercept was 0.2 or 0.5, the behavior N-mixture model's BCI coverage of the behavior probability coefficients was 91%–95%. However, when mean detection was 0.8, the behavior N-mixture model's BCI coverage was only 82% (Appendix S1: Table S3). Again, coverage of behavior probability coefficients was always higher for the behavior N-mixture model than the naïve model (Figure 1). Further, the behavior N-mixture model's mean estimate of behavior probability coefficients was 1.33–6.14 times more likely to be less biased than the naïve model (Appendix S1: Table S3). For behavior probabilities and coefficients, BCIs from the behavior N-mixture model were about 50% wider than BCIs from the naïve model, which clearly underestimated uncertainty. Therefore, the better coverage of the behavior N-mixture model can be attributed to a combination of both lower bias and wider BCIs.

Abundance parameters

Under behavioral closure, the mean absolute bias of the behavior N-mixture model's estimate of abundance (N) ranged from 0.03 to 0.22 (true values ranged from 0 to 47, with three outlier values of 88) across all scenarios (Appendix S1: Table S4). The mean absolute bias of the traditional N-mixture model's abundance estimate ranged from −1.23 to −0.02 across scenarios (Appendix S1: Table S6). The behavior N-mixture model's BCI coverage of abundance was 98%–99% (S1: Table S4), and across scenarios, the behavior N-mixture model's estimate of abundance was 1.17–2.13 times more likely to be less biased than the traditional N-mixture model (Appendix S1: Table S3).

The mean absolute bias of the behavior N-mixture model's estimate of the abundance coefficient (α') ranged from −0.03 to 0.02 across scenarios, which was similar to that from the traditional N-mixture model (−0.02 to 0.02) (Appendix S1: Tables S3 and S6). The behavior N-mixture model's BCI coverage of the true value of α' was 84%–100%, and the highest coverage occurred when the detection intercept was 0.2 (Appendix S1: Table S4). The behavior N-mixture model's mean estimate of the abundance coefficient was 0.59–3.35 times more likely to be less biased than the traditional N-mixture model across trials (Appendix S1: Table S3), so for some simulation scenarios, the behavior N-mixture model resulted in more bias than the traditional N-mixture model.

Detection parameters

Under behavioral closure, mean absolute bias of the behavior N-mixture model's estimate of detection probability ($P_{j,b,k}$) ranged from −0.01 to 0.00, and the BCI coverage

TABLE 1 Absolute bias for behavior probability coefficients

Parameter	Measure	Detection 0.2, abundance 2.7	Detection 0.5, abundance 2.7	Detection 0.8, abundance 2.7	Detection 0.2, abundance 1	Detection 0.2, abundance 7.4
Behavior probability coefficient β'_b	Mean absolute bias	0.21 (−0.04)	0.06 (−0.03)	0.03 (−0.10)	0.07 (−0.04)	−0.01 (0.25)
	2.5th and 97.5th percentiles, absolute bias	−1.28, 1.56 (−3.34, 1.88)	−0.63, 0.82 (−2.03, 1.89)	−0.74, 0.59 (−1.65, 2.00)	−1.14, −0.94 (−2.16, 1.58)	−0.70, 0.62 (−1.38, 1.60)
Behavior probability $B_{j,b}$	Mean absolute bias	−0.001 (0.01)	0.008 (0.02)	0.003 (0.02)	0.03 (0.03)	0.00 (−0.01)
	2.5th and 97.5th percentiles, absolute bias	−0.17, 0.19 (−0.29, 0.30)	−0.12, 0.12 (−0.33, 0.33)	−0.15, 0.13 (−0.36, 0.32)	−0.28, 0.27 (−0.36, 0.35)	−0.11, 0.12 (−0.27, 0.24)

Notes: Mean and 2.5th and 97.5th percentiles of absolute bias, and 95% Bayesian credible interval coverage across replicates where all chains converged (Appendix S1: Table S1) for each scenario (defined by input values of mean detection and abundance) for the behavior N-mixture model from the simulation study. The first values in each cell correspond to when the behavioral closure assumption was met, and the values in parentheses correspond to the model under complete violation of behavioral closure, such that 100% of the behaviors were resampled on each visit. The first column contains the baseline levels of the simulation, where we fixed mean detection at 0.2 and mean abundance at 2.7.

of the true value was 88%–97% (Appendix S1: Table S4). The mean absolute bias of the detection probability coefficient (γ'_b) ranged from -0.10 to 0.07 (true values ranged from -3.90 to 2.67) (Appendix S1: Table S4), and the BCI coverage of the true value was 90%–96%. The worst-performing scenario for detection parameters was one in which the mean detection probability was 0.8 (Appendix S1: Table S4). There was no alternative model that estimated behavior-specific detection probabilities for comparison because the naïve model did not estimate detection probabilities and the traditional N-mixture model did not estimate behavior-specific parameters.

Violation of behavioral closure assumption

When the behavioral closure assumption was violated, the behavior N-mixture model still outperformed the naïve model in estimating behavioral frequencies (and effects of covariates on behaviors) but performed worse than the traditional N-mixture model at estimating abundances. Specifically, when the behavioral closure

assumption was fully violated (i.e., 100% of the behaviors resampled on each visit), the mean absolute bias of behavior probability coefficients (β'_b) for the behavior N-mixture model ranged from -0.10 to 0.25 (Appendix S1: Table S4). Although the range of absolute bias increased when the behavioral closure assumption was violated, BCI coverage for the behavior N-mixture model (74%–91%) was still much greater than the naïve model (28%–67%). The only scenario in which the BCI coverage for the naïve model was greater was one in which the detection intercept was 0.8 (Figure 1). With violation of the behavioral closure assumption, the variance of the bias of behavior and detection parameters increased, but the mean absolute bias was close to zero (Appendix S1: Table S4). Across intermediate levels of behavioral closure violation, there was little variation in the BCI coverage of behavior probability coefficients (Figure 2).

With increasing violation of the behavioral closure assumption, bias in total abundance (N) became increasingly positive, and bias of detection probability (P) became increasingly negative (Figure 3; Appendix S1: Table S4). At low levels of behavioral closure violation,

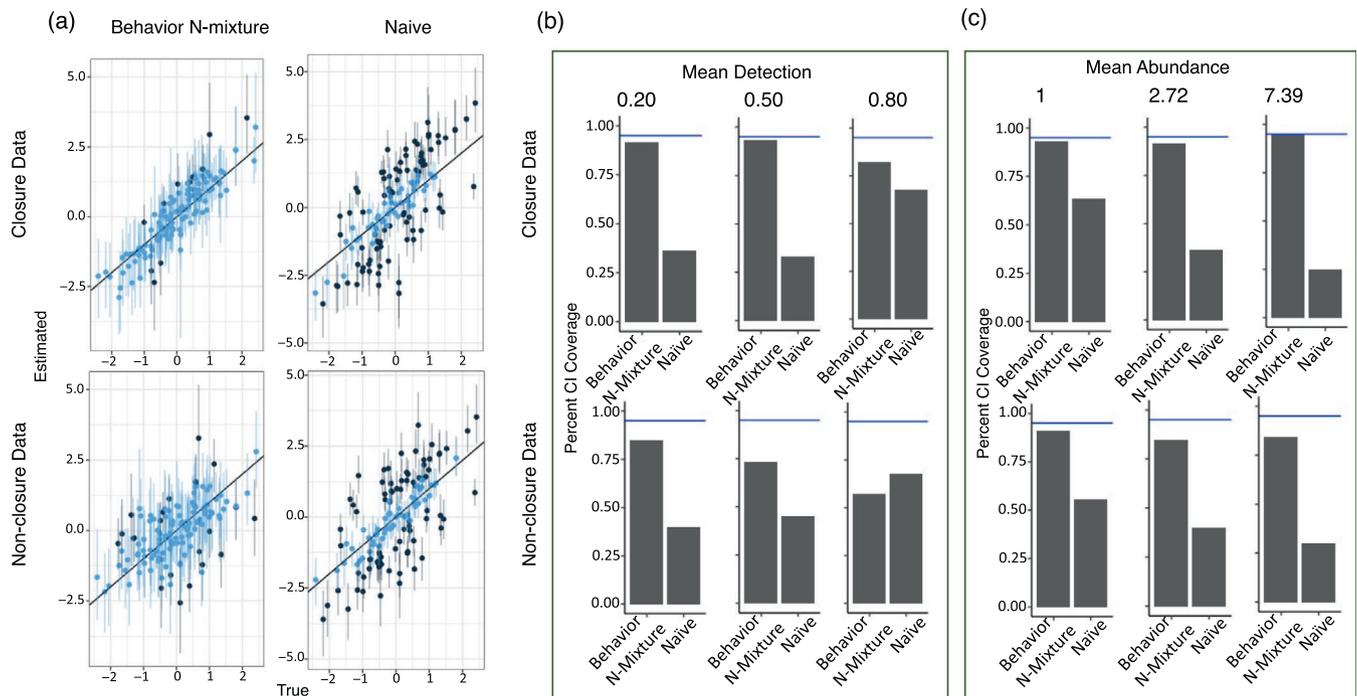


FIGURE 1 (a) Estimated versus true input values and 95% Bayesian credible intervals (BCIs) for β'_b (effect of a covariate on the probability that an individual will perform a behavior) based on a simulation study exploring the effectiveness of the behavior N-mixture model across varying levels of detection and abundance. Dark blue points and lines indicate BCIs that did not capture the true simulated values; light blue points and lines indicate BCI coverage of true values. (b) Percentage of 95% BCIs that captured true input values for β'_b across all simulation replicates at varying levels of mean detection. The mean abundance is fixed at 2.7 and the blue lines mark 0.95. (c) Percentage of 95% BCIs that capture simulated values for β'_b across all iterations at varying levels of mean abundance. The mean detection is fixed at 0.2 and the blue lines mark 0.95. Only replicates where all chains converged were used for analysis (Appendix S1: Table S1)

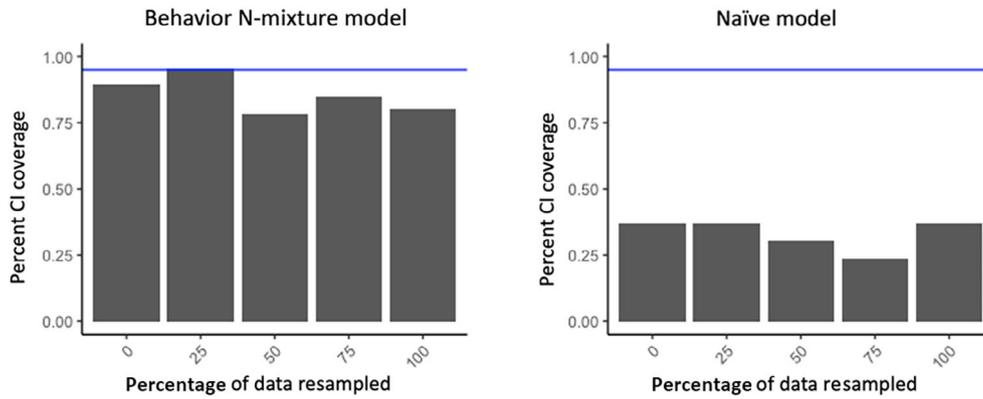


FIGURE 2 Percentage of 95% Bayesian credible intervals that captured the simulated values for β'_b (i.e., effect of a covariate on behaviors) across all replicates where all chains converged (Appendix S1: Table S1) of the simulation study where the level of the violation of the behavioral closure assumption (i.e., percentage of data resampled) varied from 0% to 100%. The blue lines mark 0.95

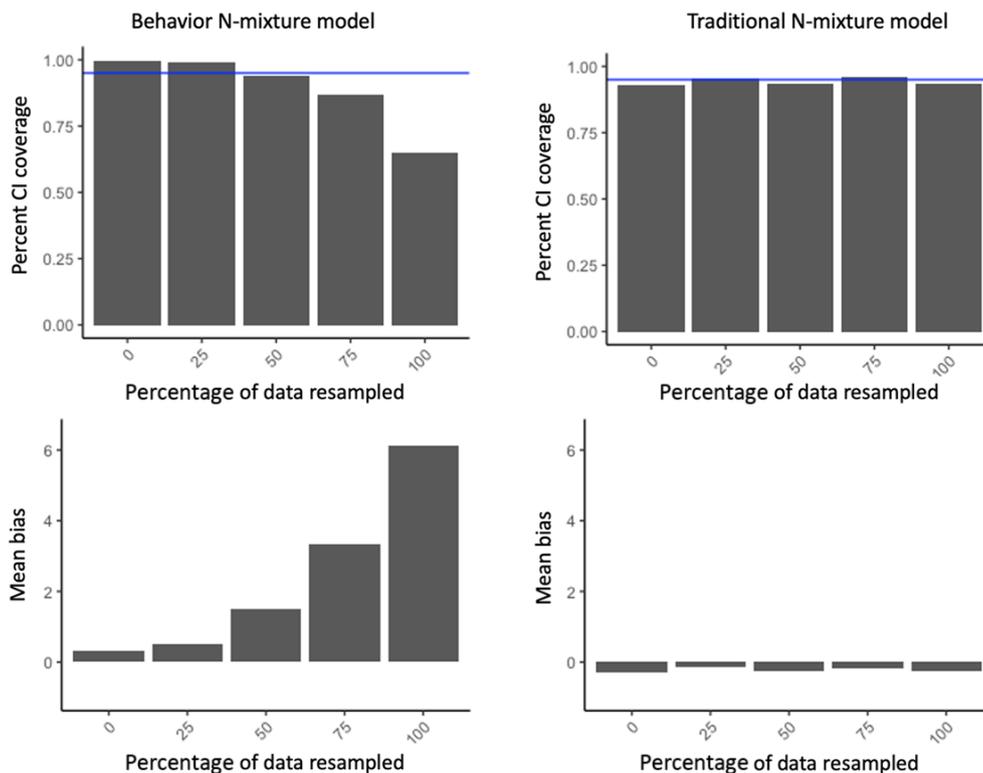


FIGURE 3 Percentage of 95% Bayesian credible intervals that captured the simulated values for N and the mean absolute bias of N across all replicates where all chains converged (Appendix S1: Table S1) of the simulation study where the level of the violation of the behavioral closure assumption (i.e., percentage of data resampled) varied from 0% to 100%. The blue lines mark 0.95

such as when 25% of the behaviors were resampled, the mean absolute bias of N was still rather low (i.e., 0.51; Figure 3; true values ranged from 0 to 13). However, at higher levels of mean abundance, mean absolute bias increased (Appendix S1: Table S4). Meanwhile, the traditional N-mixture model underestimated N on average—mean absolute bias ranged from -0.27 to -0.15 across all levels of behavioral closure violation. When the behavioral closure assumption was violated, the mean absolute

bias of the behavior N-mixture model's estimates of the abundance and detection coefficients (α' and γ') were not consistently positive or negative, ranging from -0.01 to 0.05 and -0.10 to 0.07 , respectively (Appendix S1: Table S4). In addition, when mean detection was 0.2 or 0.5, the behavior N-mixture model's estimates of α' were 1.04–2.44 times more likely to be less biased than the naïve and traditional N-mixture models across replicates (Appendix S1: Table S3).

Application: Bird behavior variation between forested and agricultural habitats

We estimated the effects of land use on the probabilities of eating and vocalizing for three focal species using the behavior N-mixture model and a naïve model. In some cases, predictions from the behavior N-mixture and naïve model largely aligned. The behavior N-mixture and naïve models both estimated higher probabilities of eating in agricultural habitats and vocalizing in forest for Inca dove, a species often found foraging in agricultural fields (Mueller, 2020), though they differed in which behavior coefficients were statistically significant (Figure 4). In other cases, the models produced very distinct results. For example, for Hoffmann’s woodpecker, a species known to occupy both forested and open areas (Stiles & Skutch, 1989), the behavior N-mixture model only found that the probability of the *other* behavior was significantly higher in agricultural land than in forest, whereas the naïve model suggested that the woodpecker was less likely to vocalize and more likely to perform the *other* behavior category in agricultural land than in forest (Figure 4). Similarly, for turquoise-browed motmot, a species that has been observed foraging and nesting in

both forested and anthropogenic landscapes (Snow & Kirwan, 2020), the behavior N-mixture model did not find significant effects of land use on behavior; however, the naïve model indicated a marginally significant higher probability of vocalizing and lower probability of *other* in forest compared to agricultural habitats (Figure 4), where the BCI bordered 0.00. Additionally, the behavior N-mixture model suggested that loud noises reduced the likelihood of detecting vocalizations of all three species, Hoffmann’s woodpecker was harder to detect later in the day, and forest cover increased the likelihood of observing turquoise-browed motmots eating (Appendix S1: Figure S1). The behavior N-mixture model also suggested that forest cover increased the abundance of Hoffmann’s woodpecker and turquoise-browed motmot but not Inca dove.

DISCUSSION

We developed and applied a novel hierarchical model to assess shifts in animal behavior across environments. Data on how animals behave can be used to understand how local management activities and global changes

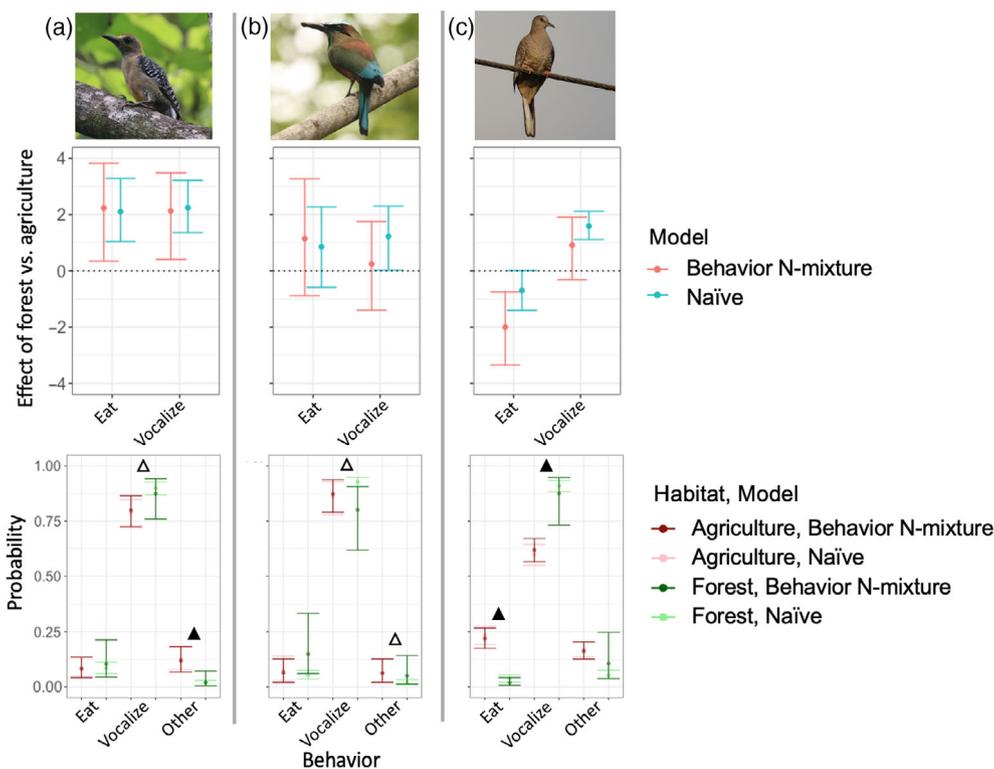


FIGURE 4 Top row: Behavior N-mixture model and naïve model estimates and 95% Bayesian credible intervals for the effect of land use (where forest = 1 and agriculture = 0) on the probability of eating and vocalizing for (a) Hoffmann’s woodpecker, (b) turquoise-browed motmot, and (c) Inca dove. Bottom row: Behavior N-mixture model and naïve model predictions of probabilities that an individual will perform each behavior in each habitat type. Significant differences between habitats according to only the naïve model are marked with an empty triangle; significant differences according to both models are marked with a solid triangle

affect species, quantify habitat suitability (e.g., by identifying potential source/sink habitats), and identify conservation interventions (Ortega-Álvarez et al., 2021). Indeed, behavior data can reveal trends that census data cannot. For example, urban parks may appear to be high-quality habitat for wildlife based on the presence and abundance of individuals, even as human activity reduces the availability of foraging resources, thereby decreasing fitness and increasing the probability of future extirpations (Jokimäki et al., 2011). Most studies of animal behavior, however, rely on intensively monitoring animals (e.g., Luck, 2002; Tremblay et al., 2005) and are thus limited to few individuals because they are labor-intensive and restricted to certain species (e.g., those that can carry tracking equipment).

Here we present a novel method that allows researchers to leverage survey-type behavioral data while accounting for imperfect detection, thereby enabling scientists or practitioners to analyze behaviors for many individuals of multiple species and better assess the conservation value of alternative habitat types. When the behavioral closure assumption was met, our behavior N-mixture model produced unbiased estimates of behavior probabilities, their relationships with predictor variables, and abundance, even when the same covariate affected both behavior and detection probability. When the behavioral closure assumption was violated, the behavior N-mixture model still produced unbiased estimates of behavior probabilities and the effects of covariates on behaviors, but it overestimated total abundance. Further, the behavior N-mixture model had better BCI coverage than the naïve model when estimating behavior probabilities in all scenarios, except when the behavioral closure assumption was completely violated and the mean detection probability was high. (Note: Figure 5 provides a decision tree of which models to choose in different circumstances.)

Behavior parameters

The behavior N-mixture model produced unbiased estimates of covariates affecting behavior probability and accurate estimates of uncertainty. As mentioned, the only scenario in which the behavior N-mixture model performed worse than the naïve model was when mean detection probability was 0.8 and the behavioral closure assumption was violated. In reality, a mean detection probability of 0.8 or greater is rare (Kéry, 2018; Kellner & Swihart, 2014), and if detection probability is near-perfect, a naïve model may be sufficient (Figure 5). While the naïve model is a simpler alternative that does not make assumptions about closure, the behavior N-mixture model produced similar or less biased estimates of behavior parameters in all cases where there was behavioral closure and at mean detection probabilities of 0.5 or lower when the behavioral closure assumption was violated (Appendix S1: Table S3).

A critical disadvantage of the naïve model is that it underestimated uncertainty, causing it to make incorrect inferences about how covariates affected behaviors. In the most egregious cases, the naïve model returns significant results, suggesting that a behavior increases in probability in an environment, when in fact it decreases. For example, under behavioral closure, when the naïve model was applied to the scenario with a mean detection probability of 0.2 and mean abundance of 2.7 (Figure 1a), 14% of the significant coefficients affecting behavior (i.e., where BCI did not include 0) were in the opposite direction of the true value. This type of error could be problematic if naïve model results were applied to management decisions because we might fail to protect essential habitats for a species' life cycle and potentially protect nonimportant habitats instead. In contrast, the behavior N-mixture model did not estimate any statistically significant coefficients affecting behavior in the wrong direction in the same scenario.

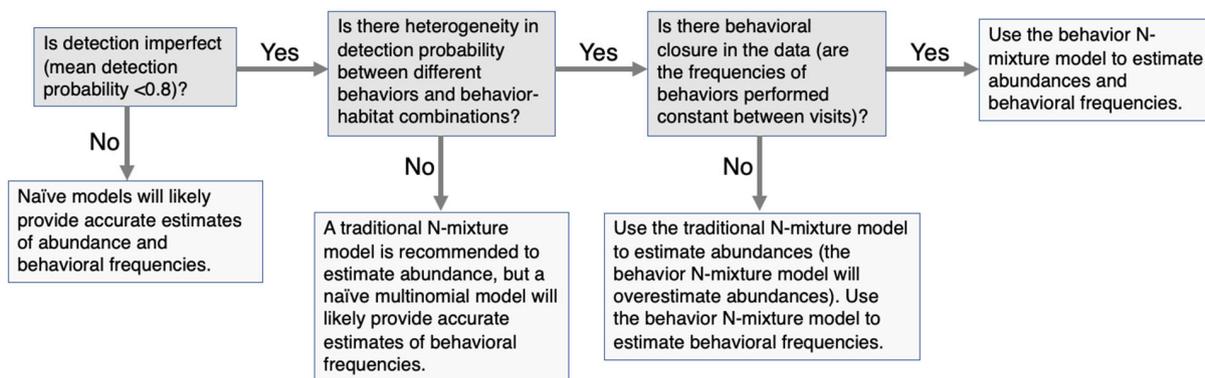


FIGURE 5 Decision tree to help determine when behavior N-mixture, naïve, and traditional N-mixture models should be used to estimate abundance and behavioral frequencies from wildlife survey data across various scenarios of imperfect detection and population closure

Abundance parameters

The behavior N-mixture model produced unbiased estimates of abundance parameters, as well as accurate estimates of uncertainty, when the data conformed to the behavioral closure assumption. In this case, it outperformed the traditional N-mixture model, which tended to underestimate abundance. When the behavioral closure assumption was violated, however, the behavior N-mixture model performed poorly (Figure 3), which we expected based on what is known about N-mixture models (Rota et al., 2009). Therefore, when the behavioral closure assumption is violated, the behavior N-mixture model should not be used to estimate abundances or detection probabilities. However, when mean detection probability is around 0.5 or lower, the behavior N-mixture model can still be used to estimate coefficients affecting abundance and detection (Table 1).

When the behavioral closure assumption is violated, the traditional N-mixture model can still be used to estimate abundance and detection in most cases. Although there was a negative bias in abundance estimates, BCI coverage was at least 93% across all scenarios. Bias was greatest when mean abundance was high (average of 7.4 individuals per site), when the traditional N-mixture model consistently overestimated detection and underestimated abundance (Appendix S1: Table S6).

In practice, it may be hard to determine whether the behavioral closure assumption is violated, so caution is warranted before using the behavior N-mixture model to estimate abundances. To assess behavioral closure, users could potentially contrast abundance estimates under both models and, if similar, just use the behavior N-mixture model. Overall, we recommend using the traditional N-mixture model to estimate abundances but the behavior N-mixture model to measure changes in behavior (Figure 5).

Bird behavior variation in northwest Costa Rica

The behavior N-mixture model produces important insights into the effects of environmental conditions on behaviors. As illustrated in the simulation study, these effects can be incorrectly determined when not accounting for imperfect and varying detection.

Among our case studies, the turquoise-browed motmot and Hoffmann's woodpecker highlight scenarios where the behavior N-mixture model suggests that birds do not behave differently in forested and agricultural habitats, while the naïve model suggests that they do. This is likely because the naïve model underestimates

the uncertainty around coefficients (as seen in the simulation study), and the detectability of some of these behaviors differed between habitats (Appendix S1: Figure S1). In contrast, the Inca dove highlights that sometimes both models can produce consistent estimates of important effects on behaviors. This may occur when there are large differences in behaviors performed between habitats that are apparent even without correcting for detection.

The results of the behavior N-mixture and naïve models seemed to differ the most when there was more heterogeneity in the effects of habitat on the detection probability of each behavior (Figure 4; Appendix S1: Figure S1). That is likely because the naïve model confounds effects of habitat on the detection of behaviors with effects on the behaviors themselves. For example, for the turquoise-browed motmot, the naïve model likely estimated a larger effect of forest on the probability of vocalizing than the behavior N-mixture model because it did not account for the fact that vocalizing is relatively easier to detect in forest than foraging, which becomes more difficult to see when there is dense cover (Appendix S1: Figure S1). Thus, a potential consequence of using the naïve model is falsely concluding that turquoise-browed motmots vocalize more frequently in forest. Because vocalization is essential for reproduction, this could lead to incorrect conclusions about the importance of forest habitat for reproduction. Indeed, turquoise-browed motmots are known to breed in human-dominated areas and even exploit nesting opportunities created by human infrastructure (Snow & Kirwan, 2020). In contrast, for Hoffmann's woodpecker, the effects of habitat on detection probability of each behavior were near zero (Appendix S1: Figure S1), likely because the foraging behavior of Hoffmann's woodpecker is often audible. Because habitat did not have a large effect on detection probability, the behavior N-mixture and naïve models produced similar coefficient estimates (Figure 4). More generally, this suggests that our model is especially critical in scenarios where there is heterogeneity in the detection probabilities of different behaviors and their relationships with the environment. For example, when conducting visual surveys of amphibians, it is important to account for the variation in detection probability depending on the activity level of individuals (Hammond et al., 2021).

The behavior N-mixture model provided useful information on how birds utilize different habitats. For the turquoise-browed motmot and Hoffmann's woodpecker, the behavior N-mixture model estimated a positive effect of forest cover on abundance, but no significant differences in the probability of eating and vocalizing between habitats (Appendix S1: Figure S1 and Table S4). This

implies that there may be higher densities of individuals in more forested habitats, but individuals have similar probabilities of eating and vocalizing in forested and in agricultural habitats. This could mean that these species indeed perform important behaviors for their life history in agricultural habitats but that forested habitat can support more individuals performing these key behaviors. In contrast, the Inca dove was estimated to eat more and vocalize less in agricultural habitats than in forest. This makes biological sense because they primarily eat seeds from grains, weeds, and grasses, which are more abundant in agricultural habitats (Johnston, 1960). Interestingly, the behavior N-mixture model did not find an effect of forest cover on abundance. This implies that agricultural habitats provide better foraging resources for the Inca dove than forest and that certain types of agriculture may boost populations of Inca doves. Without accounting for behaviors, managers may assume forested and agricultural habitats are equally beneficial for Inca doves, while in reality, agricultural habitats support more foraging resources. Thus, the behavior N-mixture model can add insights to abundance information that could lead to novel conservation strategies and ecological understanding, for example, by identifying where animals obtain resources to survive and reproduce.

Limitations

There are several limitations to applying the behavior N-mixture model in practice. First, there is currently no method to determine whether data meet the behavioral closure assumption. A potential method to measure the degree of behavioral closure is to model the abundance of each behavior category separately using an N-mixture model that allows for temporary emigration (Chandler et al., 2011). An animal switching behaviors between visits could result in a higher estimate of the probability of temporary emigration that indicates greater violation of the behavioral closure assumption, but we have not explored this further. Based on our simulation, another sign of behavioral nonclosure could be when estimates of abundance from the behavior N-mixture model are consistently higher than estimates from a traditional N-mixture model. An issue we did not explore is that not only may the behaviors performed differ between visits, but the probabilities of behaviors an individual will perform could change across visits, which could worsen the violation of the closure assumption.

Finally, all model chains converged in over 50% of the simulation trials except for when mean detection was 0.8, when there were often convergence issues. The reasons for nonconvergence of chains were unclear. For

abundance estimates that did not converge, the chains were often at zero most of the time, with a few nonzero values. For other parameters that did not converge, the chains appeared to be in different locations of parameter space. Splitting observations into behavior classes, as is necessary for the behavior N-mixture model, leads to sparse data, and increasing the number of observations could improve model convergence. Because of this, we chose frequently observed species and behaviors for our case study, and this model would work best for animals and behaviors that are more easily detected so that there are sufficient data to estimate detection probabilities. For example, in a camera trap study of mammals, it would likely be difficult to detect and analyze reproductive behaviors but easier to observe foraging behaviors (Abu Baker et al., 2015). A useful extension of the behavior N-mixture model could be to develop it into a multispecies community model so that behaviors of many species could be analyzed when data are sparse (Ovaskainen & Soininen, 2011).

CONCLUSIONS

Measuring behavior is important for understanding the ecology and conservation of species because behavioral changes are often an animal's first response to human-induced environmental changes (Wong & Candolin, 2015). The behavior N-mixture model advances the study of behavior because it accurately estimates behavior probabilities and the effects of covariates on behavior probabilities using observational surveys. In doing so, the model facilitates a low-cost method to monitor many individuals' behavior changes simultaneously. Under the behavioral closure assumption, the model can also be used to estimate abundance by modeling heterogeneity in detection between behaviors, though when the behavioral closure assumption is violated, it should not be used for abundance (Figure 5). Behavior-driven heterogeneity in detection is likely present in many taxa, and information on individual behavior can be obtained from many different survey types beyond the point-count data analyzed in our case study (e.g., camera traps for mammals, transects/area searches for herpetofauna) (Burton et al., 2015). Thus, scientists should attempt to collect behavioral observations and incorporate them into analyses whenever possible.

ACKNOWLEDGMENTS

We would like to thank the many people involved in collecting the bird survey data we used from northwest Costa Rica, including J. Zook, A. Echeverri, K. Chan, P. Juárez, X. Campos, E. Rodrigues, L. Bogantes, the FuturAgua team, staff from the National System of

Conservation Areas (SINAC) (from the Spanish: *Sistema Nacional de Áreas de Conservación*), and the landowners who supported our work. Funding was provided by a National Science Foundation Graduate Research Fellowship to Alison Ke and grants from the National Geographic Society (Nos. 9977-16; 67543R-20). All work was conducted with approval from the Costa Rican government (permits: SINAC-SE-CUS-PI-R-036-2016, SINAC-SE-CUS-PI-R-030-2017, SINAC-PNI-ACAT-043-2019, and SINAC-PNA-ACAT-039-2021).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data (Ke et al., 2022a) are available in Dryad at <https://doi.org/10.25338/B8VP85>. Code (Ke et al., 2022b) is available in Zenodo at <https://doi.org/10.5281/zenodo.5612158>.

ORCID

Alison Ke  <https://orcid.org/0000-0001-9111-449X>

Rahel Sollmann  <https://orcid.org/0000-0002-1607-2039>

Luke O. Frishkoff  <https://orcid.org/0000-0001-5738-2140>

REFERENCES

- Abu Baker, M. A., S. E. Emerson, and J. S. Brown. 2015. "Foraging and Habitat Use of Eastern Cottontails (*Sylvilagus floridanus*) in an Urban Landscape." *Urban Ecosystem* 18(3): 977–87. <https://doi.org/10.1007/s11252-015-0463-7>.
- Berger-Tal, O., D. T. Blumstein, S. Carroll, R. N. Fisher, S. L. Mesnick, M. A. Owen, D. Saltz, C. C. S. Claire, and R. R. Swaisgood. 2016. "A Systematic Survey of the Integration of Animal Behavior into Conservation." *Conservation Biology* 30(4): 744–53.
- Burton, A. C., E. Neilson, D. Moreira, A. Ladle, R. Steenweg, J. T. Fisher, E. Bayne, and S. Boutin. 2015. "REVIEW: Wildlife Camera Trapping: A Review and Recommendations for Linking Surveys to Ecological Processes." *Journal of Applied Ecology* 52(3): 675–85. <https://doi.org/10.1111/1365-2664.12432>.
- Chandler, R. B., J. Andrew Royle, and D. I. King. 2011. "Inference about Density and Temporary Emigration in Unmarked Populations." *Ecology* 92(7): 1429–35. <https://doi.org/10.1890/10-2433.1>.
- Crowe, D. E., and K. M. Longshore. 2010. "Estimates of Density, Detection Probability, and Factors Influencing Detection of Burrowing Owls in the Mojave Desert." *Journal of Raptor Research* 44(1): 1–11. <https://doi.org/10.3356/JRR-08-40.1>.
- Durso, A. M., J. D. Willson, and C. T. Winne. 2011. "Needles in Haystacks: Estimating Detection Probability and Occupancy of Rare and Cryptic Snakes." *Biological Conservation* 144(5): 1508–15. <https://doi.org/10.1016/j.biocon.2011.01.020> Ecoregional-scale monitoring within conservation areas, in a rapidly changing climate.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. *Bayesian Data Analysis*. CRC Texts in Statistical Science. London: Chapman & Hall.
- Gilroy, J. J., and D. P. Edwards. 2017. "Source-Sink Dynamics: A Neglected Problem for Landscape-Scale Biodiversity Conservation in the Tropics." *Current Landscape Ecology Reports* 2(1): 51–60.
- Hammond, T. T., M. J. Curtis, L. E. Jacobs, M. W. Tobler, R. R. Swaisgood, and D. M. Shier. 2021. "Behavior and Detection Method Influence Detection Probability of a Translocated, Endangered Amphibian." *Animal Conservation* 24(3): 401–11. <https://doi.org/10.1111/acv.12645>.
- Hatfield, J. H., C. David, L. Orme, J. A. Tobias, and C. Banks-Leite. 2018. "Trait-Based Indicators of Bird Species Sensitivity to Habitat Loss Are Effective within but Not across Data Sets." *Ecological Applications* 28(1): 28–34. <https://doi.org/10.1002/eap.1646>.
- Hilborn, R., J. A. Redfield, and C. J. Krebs. 1976. "On the Reliability of Enumeration for Mark and Recapture Census of Voles." *Canadian Journal of Zoology* 54(6): 1019–24.
- Johnston, R. F. 1960. "Behavior of the Inca Dove." *The Condor* 62(1): 7–24.
- Jokimäki, J., M.-L. Kaisanlahti-Jokimäki, J. Suhonen, P. Clergeau, M. Pautasso, and E. Fernández-Juricic. 2011. "Merging Wildlife Community Ecology with Animal Behavioral Ecology for a Better Urban Landscape Planning." *Landscape and Urban Planning* 100(4): 383–5. <https://doi.org/10.1016/j.landurbplan.2011.02.001>.
- Karp, D. S., L. O. Frishkoff, A. Echeverri, J. Zook, P. Juárez, and K. M. A. Chan. 2018. "Agriculture Erases Climate-Driven β -Diversity in Neotropical Bird Communities." *Global Change Biology* 24(1): 338–49.
- Ke, A., R. Sollmann, L. Frishkoff, and D. Karp. 2022a, "A Hierarchical N-mixture Model to Estimate Behavioral Variation and a Case Study of Neotropical Birds." Dryad, Dataset. <https://doi.org/10.25338/B8VP85>.
- Ke, A., R. Sollmann, L. Frishkoff, and D. Karp. 2022b. "A Hierarchical N-Mixture Model to Estimate Behavioral Variation and a Case Study of Neotropical Birds." Zenodo, Software. <https://doi.org/10.5281/zenodo.5612158>.
- Kellner, K. F., and R. K. Swihart. 2014. "Accounting for Imperfect Detection in Ecology: A Quantitative Review." *PLoS One* 9(10): e111436.
- Kéry, M., and J. A. Royle. 2008. "Hierarchical Bayes Estimation of Species Richness and Occupancy in Spatially Replicated Surveys." *Journal of Applied Ecology* 45(2): 589–98.
- Kéry, M. 2018. "Identifiability in N-Mixture Models: A Large-Scale Screening Test with Bird Data." *Ecology* 99(2): 281–8.
- Kéry, M., and J. A. Royle. 2015. *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS: Volume 1: Prelude and Static Models*. Cambridge, MA: Academic Press.
- Luck, G. W. 2002. "Determining Habitat Quality for the Cooperatively Breeding Rufous Treecreeper, *Climacteris rufa*." *Austral Ecology* 27(2): 229–37.
- Lyons, J. E. 2005. "Habitat-Specific Foraging of Prothonotary Warblers: Deducing Habitat Quality." *The Condor* 107(1): 41–9.
- Mueller, A. J. (2020). "Inca Dove (*Columbina inca*)." *Birds of the World*. <https://doi.org/10.2173/bow.incdov.01>
- Newbold, T., J. P. W. Scharlemann, S. H. M. Butchart, Ç. H. Şekercioglu, R. Alkemade, H. Booth, and D. W. Purves. 2013. "Ecological Traits Affect the Response of Tropical Forest Bird

- Species to Land-Use Intensity.” *Proceedings of the Royal Society B: Biological Sciences* 280(1750): 20122131.
- Ortega-Álvarez, R., V. Ruiz-Gutiérrez, O. J. Robinson, E. B. Benítez, I. M. Mena, and J. Jaime Zúñiga-Vega. 2021. “Beyond Incidence Data: Assessing Bird Habitat Use in Indigenous Working Landscapes through the Analysis of Behavioral Variation among Land Uses.” *Landscape and Urban Planning* 211(7): 104100. <https://doi.org/10.1016/j.landurbplan.2021.104100>.
- Ovaskainen, O., and J. Soininen. 2011. “Making More out of Sparse Data: Hierarchical Modeling of Species Communities.” *Ecology* 92(2): 289–95. <https://doi.org/10.1890/10-1251.1>.
- Rota, C. T., R. J. Fletcher Jr, R. M. Dorazio, and M. G. Betts. 2009. “Occupancy Estimation and the Closure Assumption.” *Journal of Applied Ecology* 46(6): 1173–81. <https://doi.org/10.1111/j.1365-2664.2009.01734.x>.
- Royle, J. A. 2004. “N-Mixture Models for Estimating Population Size from Spatially Replicated Counts.” *Biometrics* 60(1): 108–15.
- Seber, G. A. F. 1982. *The Estimation of Animal Abundance and Related Parameters*. London, UK: Charles Griffin.
- Snow, D., and G. M. Kirwan. 2020. “Turquoise-Browed Motmot (*Eumomota superciliosa*), Version 1.0.” *Birds of the World*. <https://doi.org/10.2173/bow.tubmot1.01>.
- Stiles, F. G., and A. F. Skutch. 1989. “Guide to the Birds of Costa Rica.” Comistock <http://agris.fao.org/agris-search/search.do?recordID=US201300640055>.
- R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Tremblay, I., D. Thomas, J. Blondel, P. Perret, and M. M. Lambrechts. 2005. “The Effect of Habitat Quality on Foraging Patterns, Provisioning Rate and Nestling Growth in Corsican Blue Tits *Parus caeruleus*.” *Ibis* 147(1): 17–24. <https://doi.org/10.1111/j.1474-919x.2004.00312.x>.
- Valpine, P. d., D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik. 2017. “Programming with Models: Writing Statistical Algorithms for General Model Structures with NIMBLE.” *Journal of Computational and Graphical Statistics* 26(2): 403–13. <https://doi.org/10.1080/10618600.2016.1172487>.
- Vickery, J. A., J. R. Tallwin, R. E. Feber, E. J. Asteraki, P. W. Atkinson, R. J. Fuller, and V. K. Brown. 2001. “The Management of Lowland Neutral Grasslands in Britain: Effects of Agricultural Practices on Birds and their Food Resources.” *Journal of Applied Ecology* 38(3): 647–64.
- Wong, B., and U. Candolin. 2015. “Behavioral Responses to Changing Environments.” *Behavioral Ecology* 26(3): 665–73.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher’s website.

How to cite this article: Ke, Alison, Rahel Sollmann, Luke O. Frishkoff, and Daniel S. Karp. 2022. “A Hierarchical N-Mixture Model to Estimate Behavioral Variation and a Case Study of Neotropical Birds.” *Ecological Applications* e2632. <https://doi.org/10.1002/eap.2632>