REVIEW

THE JOURNAL OF
WILDLIFE
MANAGEMENT

THE WILDLIFE SOCIETY

Designing count-based studies in a world of hierarchical models

Quresh S. Latif[1] | Jonathon J. Valente[2] | Alison Johnston[3] | Kayla L. Davis[4] | Frank A. Fogarty[5] | Adam W. Green[6] | Gavin M. Jones[7] | Matthias Leu[8] | Nicole L. Michel[9] | David C. Pavlacky Jr.[10] | Elizabeth A. Rigby[11] | Clark S. Rushing[12] | Jamie S. Sanderlin[13] | Morgan W. Tingley[14] | Qing Zhao[1]

[1]Bird Conservancy of the Rockies, 230 Cherry Street, Suite 150, Fort Collins, CO 80521, USA

[2]U.S. Geological Survey, Alabama Cooperative Fish and Wildlife Research Unit, College of Forestry, Wildlife and Environment, Auburn University, Auburn, AL 36849, USA

[3]Centre for Research into Ecological and Environmental Modelling, University of St Andrews, United Kingdom

[4]Ecology, Evolution, and Behavior Program, Department of Integrative Biology, Michigan State University, East Lansing, MI 48906, USA

[5]Department of Wildlife, California State Polytechnic University, Humboldt, CA 95521, USA

[6]National Operations Center, Bureau of Land Management, Lakewood, CO 80226, USA

[7]USDA Forest Service, Rocky Mountain Research Station, Albuquerque, NM 87102, USA

[8]Biology Department, William & Mary, 540 Landrum Drive, Williamsburg, VA 23187, USA

[9]National Audubon Society, 225 Varick Street, New York, NY 10014, USA

[10]Bird Conservancy of the Rockies, 14500 Lark Bunting Lane, Brighton, CO 80603, USA

[11]U.S. Fish and Wildlife Service, Midwest Region, 5600 American Boulevard. West, Suite 990, Bloomington, MN 55437, USA

[12]Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA 30602, USA

[13]USDA Forest Service, Rocky Mountain Research Station, 2500 S. Pine Knoll Drive, Flagstaff, AZ 86001, USA

[14]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, Los Angeles, CA 90095, USA

Correspondence
Quresh S. Latif, Bird Conservancy of the Rockies, 230 Cherry Street, Suite 150, Fort Collins, CO 80521, USA.
Email: quresh.latif@birdconservancy.org

Abstract

Advances in hierarchical modeling have improved estimation of ecological parameters from count data, especially those quantifying population abundance, distribution, and dynamics by explicitly accounting for observation processes, particularly incomplete detection. Even hierarchical models that account for incomplete detection, however, cannot compensate for

data limitations stemming from poorly planned sampling. Ecologists therefore need guidance for planning count-based studies that follow established sampling theory, collect appropriate data, and apply current modeling approaches to answer their research questions. We synthesize available literature relevant to guiding count-based studies. Considering the central historical and ongoing contributions of avian studies to ecological knowledge, we focus on birds as a case study for this review, but the basic principles apply to all populations whose members are sufficiently observable to be counted. The sequence of our review represents the thought process in which we encourage ecologists to engage 1) the research question(s) and population parameters to measure, 2) sampling design, 3) analytical framework, 4) temporal design, and 5) survey protocol. We also provide 2 hypothetical demonstrations of these study plan components representing different research questions and study systems. Mirroring the structure of hierarchical models, we suggest researchers primarily focus on the ecological processes of interest when designing their approach to sampling, and wait to consider logistical constraints of data collection and observation processes when developing the survey protocol. We offer a broad framework for researchers planning count-based studies, while pointing to relevant literature elaborating on particular tools and concepts.

In the past 2 decades, hierarchical models (key terms defined in Table 1 and indicated with references to Table 1 upon first use) have improved our capacity for advancing ecological knowledge (Nichols et al. 2000, 2009; Royle and Dorazio 2008; Kéry and Royle 2016, 2020). In particular, models for analyzing data generated from counting readily observable animals (e.g., birds) have advanced our capacity to investigate wildlife populations across broad spatial and temporal scales (Chandler and King 2011; Pavlacky et al. 2012, 2017). By explicitly modeling the observation process (Table 1) as a distinct sub-model (Table 1) separate from an underlying ecological process (Table 1) model, state-space (Table 1) hierarchical models, particularly those that account for incomplete detection, can more accurately estimate population parameters (Nichols et al. 2009, Bailey et al. 2014). Moreover, researchers can use hierarchical models (i.e., both state-space and multi-level models [Table 1]) to explicitly distinguish and quantify various data-generating processes (Table 1) for more nuanced ecological inference, including distinguishing coarse-scale from fine-scale distributional patterns (Pavlacky et al. 2012, Latif et al. 2020), spatial variability from temporal population change (MacKenzie et al. 2003, Dail and Madsen 2011, Hostetler and Chandler 2015), spatial components of animal observability from temporal components (Nichols et al. 2009, Chandler et al. 2011, Amundson et al. 2014), and various ecological relationships among species forming a community (Kéry et al. 2009, Iknayan et al. 2014, Riecke et al. 2021, Zipkin et al. 2023).

TABLE 1 Glossary of terms used to describe a framework for planning studies based on counting unmarked populations of animals.

Term	Definition
Abundance	The total number of individual members of a species occurring within a particular geographic or spatial region (e.g., a country, state, province, or management unit), including density (see definition below)
Analytical framework	The paradigm for data analysis, which for count-based studies consists of expected data-generating processes and an analytical model capable of representing those processes sufficiently to answer a research question
Analytical model	A mathematical construct representing processes thought to give rise to observed (e.g., count) data and thus used for data analysis, the basic building blocks of which are probability functions and covariates. Basic analytical models describe variation in observed data using a probability function (typically binomial or Poisson for counts) modulated by at least one covariate.
Assumption	A condition of the data-generating or sampling process that must be met for reliable inference from an analytical model
Auxiliary data	Data collected alongside counts to help distinguish underlying processes. Usually, auxiliary data allow models to distinguish observation from ecological processes. Common examples include distances to or time of detections.
Closure assumption	The assumption required for observation sub-models that depend on replicate counts, which is that the ecological state (i.e., species presence or number of individuals present) is the same (i.e., closed to change) across replicates for a given survey unit within a given survey period
Covariate	Variables measuring attributes of the study system, sampling design, or counting process with which dependent variables are related (usually linearly) in analytical models. For this review, dependent variables are usually ecological or observation parameters (e.g., abundance, occupancy, or detectability), or raw counts.
Data-generating process	Any process that gives rise or contributes to data, including both ecological and observation processes
Density	The number of individual members of a species per unit area (e.g., km^2) occurring within a specified temporal period (i.e., abundance scaled by area)
Ecological process	All patterns and underlying mechanisms of interest relevant to a particular research question in ecology
Fundamental parameters	The basic population parameters that count-based studies aim to measure are abundance or occupancy. All other ecological parameters of possible interest (e.g., environmental relationships, trend, or dynamics) describe variation in these.
Hierarchical model	An analytical model containing >1 linked sub-models. Considering their purpose in ecology, hierarchical models typically contain ≥1 ecological sub-model and ≥1 observational sub-model (i.e., state-space models), allowing researchers to explicitly distinguish ecological from observation processes. Additionally, multi-level models can employ probability functions to link multiple sub-models representing multiple groups (i.e., random effects).
Multi-level models	Hierarchical models that employ a probability function to describe variability among groups while also linking those groups to allow partial information sharing across groups (i.e., random effects). Fundamentally, random effects in multi-level models link >1 parameter, each of which quantifies a group. Where linked parameters are subject to covariates, however, random effects effectively link sub-models.

TABLE 1 (Continued)

Term	Definition
Observation process	The process(es) by which we detect, identify, record, and enumerate population members when implementing count-based surveys (*contra* ecological process). For planned studies, we do not include the sampling process.
Occupancy	The probability of at least one member of a species occurring within a specified spatial unit and temporal period
Probabilistic sampling	Selection of sampling units from the sampling frame at random with known probabilities for inclusion in a sample
Probability function	A mathematical function describing the distribution of a random variable. The 2 types of probability functions are probability density functions, which describe variation in continuous variables (e.g., normal distributions), or probability mass functions, which describe variation in discrete variables (e.g., binomial or Poisson distributions).
Sample size	With probabilistic sampling, the number of sampling units selected from the sampling frame
Sampling design	The process by which we select a set of locations to count population members in a manner that we expect to represent a study population. Formally, sampling design consists of 3 components: 1) the sampling unit, 2) the sampling frame, and 3) selection of sampling units from the sampling frame.
Sampling frame	The area containing all possible sampling units that could represent the study population and that the researcher could survey. As such, the sampling frame defines the spatial extent of the population represented by sampled units and the data. For our purposes, we consider a sampling frame inclusive of both defined and undefined frames (Thompson et al. 1998). A sampling frame is defined if all possible sampling units are clearly delineated and sampling units do not overlap. A defined sampling frame is generally preferable where possible because it allows both design-based and model-based inference (Williams and Brown 2019), whereas undefined sampling frames (i.e., where sampling units can overlap and point coordinates for unit centers are selected from continuous space) require model-based inference to draw inference to un-sampled portions of the study population.
Sampling frequency	The number of sampling periods that occur at a given sampling unit, and thus how often the portion of the study population at that unit is represented in the data, over the course of the study period
Sampling period	A discrete unit of time in which birds will be counted at a designated sampling unit, and thus the time period over which any given count represents the portion of the study population at that unit. The sampling period determines interpretation and thus inference from population estimates (i.e., occupancy or abundance).
Sampling unit	A unique unit of space potentially selected from the sampling frame so that counts at units selected for sampling can represent the study population. Abundance and occupancy are estimated at the level of the sampling unit, so the dimensions of a sampling unit determine interpretation of and inference from these fundamental parameters.
Spatially balanced sampling	Probabilistic sampling wherein the sample selection probability is adaptively adjusted upon selection of each sampling unit to ensure the selected sample represents conditions within the sampling frame in proportion to their prevalence on the landscape. Generalized random tessellation stratified (GRTS) sampling is a common algorithm used to achieve spatially balanced sampling (Stevens and Olsen 2004).
State-space model	A hierarchical model that links sub-models whereby one sub-model is conditioned on another. Such structuring is most typically used to distinguish ecological from

TABLE 1 (Continued)

Term	Definition
	observation processes, and are foundational to hierarchical models in ecology as defined by Kéry and Royle (2016).
Stratified sampling	One type of probabilistic sampling wherein the researcher divides the sampling frame into non-overlapping strata and sets selection probabilities at levels that deviate from the relative area of the strata
Study period	The temporal extent of the study, usually number of years for species with seasonal phenology
Study population	The group of individuals relevant to the research question, which typically consists of all members of a population or species that occur within the sampling frame during study period and that are available for sampling (i.e., detectability > 0)
Sub-model	An analytical model that forms a component of a hierarchical model (defined above)
Survey frequency	The number survey periods implemented within the sampling period at a given sampling unit
Survey period	The unit of time within which a single count of population members is implemented at a designated survey unit
Survey protocol	The manner in which population members are counted at designated sampling units and within the designated sampling period
Survey unit	A unit of space where a single count-based survey is focused. Depending on logistical constraints for surveys, a survey unit may coincide spatially with a sampling unit or it may represent a subset of area within the sampling unit. In the latter case, multiple survey units may be nested within each sampling unit to better cover the sampling unit.
Temporal design	The temporal component of sampling, which consists of the study period, the sampling period, and sampling frequency

With the increased information they provide, hierarchical models place greater demands on data that are often not fully appreciated by ecologists. Explicitly modeling observation and ecological processes usually requires additional sampling or auxiliary data (Table 1) to distinguish between them. Moreover, without careful planning of where and when to implement surveys, researchers cannot assume that counts represent the populations they seek to measure (Sanderlin et al. 2019, Boyd et al. 2023). The emergence of complex models for mining large opportunistic datasets (Sauer and Link 2011, Kelling et al. 2019, Johnston et al. 2020) may convince researchers it is unnecessary to consider nuances of sampling and data collection for reliable inference; however, models cannot supplant the critical value of high-quality data from well-planned studies (Blanco et al. 2012, Sanderlin et al. 2019, Boyd et al. 2023). Moreover, the advent of data integration makes rigorous data even more valuable as a reference for identifying and adjusting for deficiencies in opportunistic data in the context of data mining (Miller et al. 2019; Zipkin et al. 2021, 2023). Advances in modeling therefore raise the need for concomitant updates to guidance for study planning (Popovic et al. 2024).

Birds provide a useful case study for understanding the need for study plan guidance. Avian responsiveness to environmental change (Furness and Greenwood 1993), the various ecosystem services they provide (Whelan et al. 2015), and the feasibility of bird surveys relative to other taxa (Bibby et al. 2000) have made counting birds attractive for investigating both population-level and macroecological questions (Wiens and Rotenberry 1985, Brown 1995, Hanski 1998). Historically, ornithologists primarily counted birds following standardized sampling designs (Table 1) and survey protocols (Table 1; Ralph et al. 1993, Rosenstock et al. 2002, Matsuoka et al. 2014) and analyzed resulting data with conventional analytical models (Table 1; Wiens and Rotenberry 1985, Link and Sauer 1998, Hutto and Patterson 2016). Accepting that incomplete detection (i.e., imperfect detection; Bailey et al. 2007) limits counts to a subset of detected population members, researchers related counts with covariates

representing spatial, temporal, or environmental variation, and interpreted model estimates as indices quantifying relative variability or population change (Link and Sauer 1998, Johnson 2008). This approach assumed that variability in the proportion of the population counted was either constant or at least not confounded with ecological patterns of interest. Ecologists have since demonstrated the tenuousness of this assumption (Table 1) and the consequent need to better account for the observation process (Table 1) for rigorous and clear inference (Ruiz-Gutiérrez and Zipkin 2011, Marques et al. 2017). Advancement of modeling methods that do so (Kéry and Royle 2016, 2020), however, has rendered most commonly referenced guides for counting birds incomplete (Hutto et al. 1986, Ralph et al. 1993, Hamel et al. 1996, Bibby et al. 2000, Scott 2002).

Here, we synthesize available literature to guide wildlife biologists in planning count-based studies that can effectively employ hierarchical models to answer their research questions. We use birds as a case study to focus our review, but the concepts we cover and approach we outline applies broadly to many taxa sufficiently observable to be counted (Otto et al. 2013, Dénes et al. 2015, Tobler et al. 2015). We start with a broad overview of potential questions that count data can help answer and population metrics that they can help measure. We then describe core components of effective study plans (Table 1) that incorporate hierarchical models to answer questions. We present the components of our planning framework in the order we suggest they be considered: 1) research question(s) and population parameter(s) to measure, 2) sampling design, 3) analytical framework, 4) temporal design, and 5) survey protocol (for definitions, see Table 1; Figure 1). Our framework mirrors the distinction of ecological from observation processes offered by hierarchical models; sampling and temporal design components concern ecological process(es), whereas the survey protocol concerns observation process(es). Moreover, we distinguish between sampling design and temporal design because the latter is more determined by the analysis model and therefore better follows development of the analytical framework. Finally, we provide examples that demonstrate application of this framework for research and monitoring. Throughout, we strive to explain key concepts and ideas in broad, accessible terms while providing references with further details.

RESEARCH QUESTIONS AND POPULATION METRICS

Count data can help researchers investigate various ecological questions (Table 2), which determine the ecological parameters researchers need to measure. Researchers focusing on macroecological questions (Brown 1995, Hanski 1998) may primarily seek to quantify occupancy (i.e., presence or absence of a species within a spatial and temporal unit; Table 1). For questions concerning population size, researchers can measure abundance (Table 1) for a given geography. For questions about factors governing a population's spatial distribution, researchers may instead need to estimate relationships of density (Table 1) with variability in environmental conditions. For count-based studies, researchers typically want at minimum to measure abundance or occupancy (i.e., fundamental parameters [Table 1]), whereby additional parameters describe variation in these (e.g., environmental relationships, trend, dynamics). Clearly articulating objectives and keeping them in focus is important when planning a study to ensure the sampling design, analytical framework (Table 1), temporal design, and survey protocol adequately address the primary question(s) of interest (Yoccoz et al. 2001).

SAMPLING DESIGN

After identifying their question and parameters to measure, researchers should consider their sampling design. Sampling is the process by which we select a set of locations to survey because we expect counts at these locations to accurately represent the study population (Table 1; Thompson et al. 1998). In most cases, researchers cannot feasibly survey all locations containing the study population, so they select a representative sample, from which they can confidently estimate parameters representing the population as a whole. Sampling design therefore
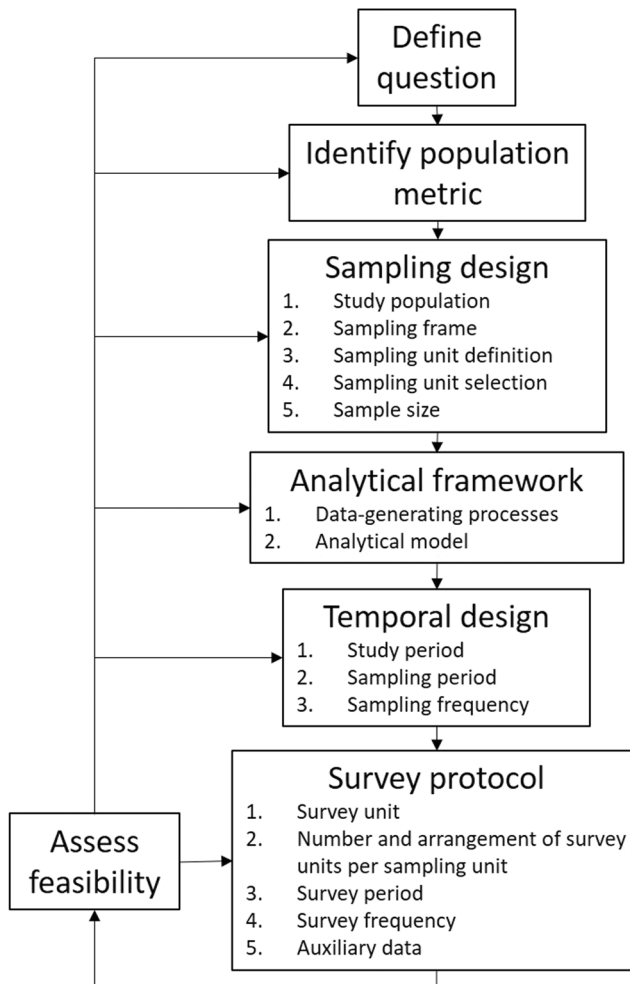
FIGURE 1 Schematic outlining suggested steps in planning studies based on counting unmarked populations of animals. Researchers may need to rethink initial components of the study plan after assessing its overall feasibility. Regardless, researchers will most likely succeed by planning all aspects of a study prior to implementation.

consists of identifying the study population, defining the sampling frame (i.e., the set of all sampling units that together contain the study population; Table 1), and selecting sampling units (Table 1) from the sampling frame (Thompson et al. 1998). Researchers can also apply models to predict beyond the study population, but the required assumptions for prediction make it less reliable than estimation within the sampling frame (Wenger and Olden 2012). The sampling frame is therefore best defined so as to contain the entire study population relevant to the research question.

Defining the sampling frame requires defining a sampling unit, the dimensions of which should reflect the ecological parameter(s) researchers want to measure. Sampling units smaller than a single territory or home range are better suited for measuring fine-scale space use. For territorial species, sampling units close to the size of a territory will tend to yield counts consisting primarily of 0 s and 1 s that are well suited for measuring occupancy that tracks variability in abundance (Linden et al. 2017, Latif et al. 2018, Steenweg et al. 2018). Sampling units large enough to contain multiple individuals will be better suited for measuring spatial distribution at coarser resolutions (Steenweg et al. 2018). For colonial or spatially aggregated populations, sampling units may correspond with

TABLE 2 Examples of research questions or monitoring objectives potentially addressed with count-based population surveys. Example studies and broad aspects of population ecology targeted in each example are listed with each question.

Question or objective	Example paper	Occupancy or abundance[a]	Distribution or dynamics
What is the species' population size and distribution within a specified area?	Hierarchical distance sampling models to estimate population size and habitat-specific abundance of an island endemic (Sillett et al. 2012)	Abundance	Distribution
How does natural disturbance affect bird abundance?	Changes in forest bird abundance, community structure and composition following a hurricane in Sweden (Chevalier et al. 2019)	Abundance	Distribution
Has population size increased or decreased?	The role of detectability on bird population trend estimates in an open farmland landscape (Sanz-Pérez et al. 2020)	Abundance	Dynamics
What is the population's geographic distribution?	Using Light Detection and Ranging (LiDAR)-derived vegetation metrics for high-resolution, species distribution models for conservation planning (Farrell et al. 2013)	Occupancy	Distribution
What climate niche does the species occupy, and how will this niche be affected by environmental change?	Birds track their Grinnellian niche through a century of climate change (Tingley et al. 2009)	Occupancy	Distribution
Produce predictive map to inform conservation or management planning	A multi-scale occupancy model for the grasshopper sparrow in the Mid-Atlantic (Irvin et al. 2013)	Occupancy	Distribution
Where are species reintroduction efforts most likely to succeed?	Surviving with a resident despot: do revegetated patches act as refuges from the effects of the noisy miner (*Manorina melanocephala*) in a highly fragmented landscape? (Mortelliti et al. 2016)	Occupancy	Distribution
Which environmental features determine habitat for a species?	An occupancy approach to monitoring regent honeyeaters (Crates et al. 2017)	Occupancy	Distribution
What are the implications of alternative management scenarios for population, demographic, or community parameters?	Utilization of a species occupancy model for management and conservation (McFarland et al. 2012)	Occupancy	Dynamics
How has migratory phenology shifted over time?	Estimating unbiased phenological trends by adapting site-occupancy models (Roth et al. 2014)	Occupancy	Dynamics
How functionally diverse is the community, how are traits structured, and what is the integrity of these structures to environmental perturbation?	The importance of accounting for imperfect detection when estimating functional and phylogenetic community structure (Si et al. 2018)	Occupancy	Dynamics

[a]Abundance here includes density (abundance per unit area).

locations or natural features capable of containing a colony or aggregation (e.g., wetlands for amphibians; Wright et al. 2020). For measuring populations at multiple spatial scales, researchers can define nested sampling units consisting of relatively large primary units that each contain multiple secondary units (Mordecai et al. 2011, Pavlacky et al. 2012). Researchers may unconsciously define a sampling unit based on the logistic feasibility of conducting surveys, to the detriment of their sampling design. To avoid this pitfall, we suggest distinguishing sampling units from survey units (Table 1), and reserving logistical considerations for the latter (Pescott et al. 2019).

Next, researchers need to decide how to select sampling units from the sampling frame. Probabilistic sampling (Table 1), whereby researchers select sampling units with known probabilities from the sampling frame, most readily and unambiguously provides data representing the study population (Morrison et al. 2008). A sample is representative of the population if inclusion of sampling units in the sample is not correlated with the ecological parameter of interest (Boyd et al. 2023). Simply selecting sampling units with equal probability from the sampling frame can effectively represent relatively small, homogenous populations. Sophisticated algorithms that ensure spatially balanced sample selection (i.e., spatially balanced sampling [Table 1]) more efficiently represent larger heterogeneous populations (Stevens and Olsen 2004, Theobald et al. 2007). Uneven selection probabilities may be necessary for various reasons, in which case estimation can correct for selection probabilities as long as they are known (i.e., design-based inference; Williams and Brown 2019). In particular, stratified sampling (Table 1) can provide relatively efficient sampling for research questions concerning population distribution along environmental gradients. Strata are mutually exclusive portions of the sampling frame. Elevating sample selection probabilities in strata delineating relatively rare environmental conditions can help ensure sufficient sampling to estimate relationships with those conditions. Moreover, estimating relationships with environmental conditions or hypothesized drivers of population distribution or dynamics may be necessary for hypothesis testing and prediction beyond the study population (Williams and Brown 2019). Even spacing of sampling units can also represent the study population but requires an additional assumption that the population is distributed independently of selected unit spacing, which is typically difficult to know for certain (Thompson et al. 1998). More complex selection strategies are less common but potentially worth considering to improve sampling efficiency (Pacifici et al. 2016). Because analyses cannot fully compensate for inappropriate sampling, the importance of understanding the strengths and weaknesses of different sampling designs cannot be overstated.

Finally, researchers must decide how many sampling units to select (i.e., sample size [Table 1]). Although information for estimating ecological parameters generally scales with sample size, increasing sample size cannot compensate for unrepresentativeness of the sample (Boyd et al. 2023), which is why we describe sample size after other components of sampling design. Ideally, researchers should conduct power analyses before initiating a study to ensure sufficient sample size. Alternatively, researchers can simply maximize their sample size given available funding, in which case consulting published power analyses or successful studies similar to theirs can help verify their sample size is reasonable. Within a broader research program, power analyses can also ensure sample size is not larger than necessary, potentially conserving funds for other questions (Bailey et al. 2007, Latif et al. 2018). Moreover, sample size may trade off with components of the survey protocol such as the number of replicate counts implemented to provide data for estimating incomplete detection (Bailey et al. 2007, Sanderlin et al. 2014).

ANALYTICAL FRAMEWORK

Identifying ecological and observation processes

We recommend researchers first consider likely data-generating processes relevant to their question, and then select a model capable of representing those processes. The number of individuals counted at any one place and time reflects both the number of individuals actually present and how readily the surveyor(s) can detect and correctly identify those individuals. We therefore need to account for the latter before we can correctly infer the

former, which is typically the ecological attribute of interest. As an initial step, we suggest researchers list likely data-generating processes (e.g., habitat, population dynamics, incomplete detection) that could affect count data within the context of the established sampling design.

While listing data-generating processes, we also suggest categorizing them as ecological, observational, or both. We might be inclined to consider any process concerning the study system (e.g., population dynamics, behavior, environmental change) as ecological and any process involved in counting (e.g., surveyor perception, survey protocol, sampling design) as observational. For selecting or developing an analytical model, however, researchers should view processes in light of their research question. Surveyor skill and attention represent elements of the observation process for virtually all studies (Ralph et al. 1993, Rosenstock et al. 2002, Matsuoka et al. 2014). Aspects of biology, such as singing rate or territorial movement for birds may at first seem ecological, but more typically represent factors modulating our ability to detect and count individuals (Betts et al. 2008, Robertson et al. 2010, Amundson et al. 2019, Berigan et al. 2019). Perhaps even less obvious are factors that influence both observation and ecological processes. Vegetation or environmental structure can affect the number of individuals present and how easily we can detect them (Ruiz-Gutiérrez and Zipkin 2011). Abundance can modulate species detectability (Royle and Nichols 2003), and may therefore be part of the observation process when measuring occupancy (Kéry et al. 2009). Within-territory movement may confound our ability to estimate site-level population dynamics among years (Berigan et al. 2019), but such movements may be of biological interest when investigating fine-scale habitat selection and space use patterns (Valente et al. 2017, Steenweg et al. 2018). We note that with a well-planned sampling design, researchers can avoid complicating their analysis by having to include the sampling process in their list of observational data-generating processes (Johnston et al. 2020, Boyd et al. 2023).

Following listing and categorization, we suggest researchers describe expected relationships between the data-generating processes and their count data. Graphs depicting expected relationships provide a useful approach to map data-generating processes in a manner that can also describe the structure of analytical models (Ethier and Nudds 2017). Lists and graphs of data-generating processes will best guide analytical model structure when rooted in established ecological theory. For example, researchers interested in estimating population trends over multiple years might remember that trends ultimately reflect dynamics of survival, reproduction, immigration, and emigration, which are typically measured annually (MacKenzie et al. 2003, Dail and Madsen 2011). For questions concerning spatial distribution at multiple scales, theory describes habitat selection at coarse scales as an antecedent to finer scale selection (Cody 1985). Comprehensive listing and graphing of all data-generating processes is generally impossible and would often include many trivial processes that are either unimportant or irrelevant to the primary research questions. We therefore recommend maintaining focus on the processes most relevant to the research questions.

Analytical model

Hierarchical models (defined here following Kéry and Royle 2016) expand the toolbox available for quantifying data-generating processes to meaningfully answer research questions. The fundamental building blocks for analytical models are probability functions (Table 1) and covariates (Table 1). Conventional analytical models describe variation in observed data using a probability function (e.g., binomial for occupancy or Poisson for abundance), the basic parameters of which may be modulated by ≥1 covariate. Hierarchical models link multiple analytical models (i.e., sub-models). Ecologists typically use state-space hierarchical models (Hostetler and Chandler 2015) to explicitly distinguish the ecological from observation process by representing each with ≥1 sub-model. Additionally, multi-level models incorporate probability functions that can link multiple sub-models representing different groups to allow partial information sharing across groups (i.e., random effects); for example, species within community models (Dorazio and Royle 2005, Chandler et al. 2013). While these approaches clarify and strengthen ecological inference, linking multiple sub-models comes with greater demands on data and often

requires auxiliary data collection and additional assumptions. As detailed in the remainder of this section, researchers must therefore decide which processes (or groups) to represent with distinct sub-models, or instead with simpler and less-demanding approaches (e.g., covariates).

Researchers must first decide whether to represent ecological and observation processes as distinct sub-models. Three conceptually foundational and commonly used hierarchical models are occupancy (MacKenzie et al. 2002, 2003; Bailey et al. 2014), N-mixture (Royle 2004), and distance sampling models (Royle et al. 2004, Buckland et al. 2015). Occupancy and N-mixture models require replicate counts, and distance sampling models require distance measurements from the surveyor to detected individuals. These foundational hierarchical models assume that all detected individuals are correctly identified such that observation errors only arise from incomplete detection. Researchers using these models must constrain their survey protocols to exclude false detections. Where necessary, researchers can add sub-models to account for false positive errors (e.g., from species misidentification) with additional demands on the data (Royle and Link 2006, Clare et al. 2021, Zhao et al. 2022). Additional assumptions are also required for each approach (Table 3; Matsuoka et al. 2014, Duarte et al. 2018, Louvrier et al. 2018, Marques et al. 2017, Fogarty and Fleishman 2021). Occupancy and N-mixture models also assume the ecological state (i.e., species presence or abundance, respectively) at a surveyed location remains the same between replicate surveys (i.e., closure assumption [Table 1]); distance sampling assumes 100% detectability at distance = 0 and that distances are measured accurately.

To avoid these additional data requirements and assumptions, some researchers continue to rely on non-hierarchical approaches, including covariates (Sauer and Link 2011), error partitioning (Dennis et al. 2006), or constrained sampling and surveys (Hutto and Patterson 2016), to address heterogeneity in the observation process. Inference with these approaches assumes either an invariant observation process or one that at least does not confound intended ecological inference (Figure 2; Guillera-Arroita et al. 2015, Hutto and Patterson 2016), assumptions that can be very difficult to verify without explicit testing. State-space hierarchical models can more readily account for confounding heterogeneity in the observation process but do not do so automatically. Both hierarchical and non-hierarchical models require appropriate structure (e.g., covariates) to address heterogeneity in the observation process, which is often what primarily confounds desired inference (Royle and Nichols 2003, Efford and Dawson 2012, Veech et al. 2016, Latif et al. 2018).

Having opted for (state-space) hierarchical modeling, researchers must decide how many and which data-generating processes to represent with distinct ecological and observational sub-models. In addition to fundamental parameters, ecological sub-models can distinguish various processes, including population dynamics, range dynamics, spatial distribution at multiple scales, and population parameters for multiple species, allowing derivation of community structure and composition. Such sub-models invariably impose additional demands or constraints on the data (Table 3). For example, models explicitly quantifying annual dynamics require implementing counts at selected sampling units repeatedly across successive years (MacKenzie et al. 2003, Dail and Madsen 2011). In contrast, repeated and successive sampling of individual units is less important for models that simply include year as a categorical covariate, although repeated sampling across years can still help distinguish temporal from spatial variation (Rhodes and Jonzén 2011).

Multiple sub-models can also help account and correct for multiple observation processes (Nichols et al. 2009, Schmidt et al. 2022). Incomplete detection sub-models fitted with data from replicate counts (e.g., repeat counts or multiple observers), such as in N-mixture models, can account for observer skill and temporal processes, such as singing rate for birds and movement, but ambiguity in the effective area surveyed can bias abundance estimates from N-mixture models or obscure their interpretation (Nichols et al. 2009). In contrast, the detection sub-model in distance sampling models accounts for effective area surveyed but ignores temporal processes (Bächler and Liechti 2007). Integrating both sub-models may be necessary for accurate abundance estimation if counts both decline with perceptibility of individuals farther from the surveyor and with intermittent availability of individuals for detection (e.g., most point-based counts of territorial songbirds; Chandler et al. 2011, Amundson et al. 2014). Models that account for false positive errors along with incomplete detection may also do so with a distinct sub-model for each (Clare et al. 2021).

TABLE 3 Example hierarchical models for count data of unmarked animals and required data and assumptions.

Model	Sub-models		Data		Additional assumptions[e]	Reference(s)
	Ecological[a]	Observation[b]	Type[c]	Requirement[d]		
Occupancy	ψ	p	O1	R	FN, CL	MacKenzie et al. (2002)
N-mixture	N	r	c	R	FN, CL	Royle (2004)
Distance sampling	N	$r_{[p]}$	c	DST	FN, PDg(0), DistMA	Buckland et al. (2015), Royle et al. (2004)
Time-to-detection	N	r	c	TD	FN, CL	Farnsworth et al. (2002)
	ψ	p	c	TD	FN, CL	Rota et al. (2009)
False positive occupancy	ψ	p, p_{O1}	O1	R or TD, ConfObs(+)	CL	Clare et al. (2021)
Dynamic occupancy	$\psi 0$, γ, ε	p	O1	R, RT	FN, CL	MacKenzie et al. (2003)
Dynamic N-mixture	$N0$, S, R or $N0$, λ	r	c	$R_{(+)}$, RT	FN, CL	Dail and Madsen (2011)
Multi-scale occupancy	ψ, θ	p	c	R/TD, NS	FN, CL	Mordecai et al. (2011), Pavlacky et al. (2012)
Community occupancy	ψ_i	p_i	O1	R/TD	FN, CL	Dorazio and Royle (2005)
Community abundance	N_i	p_i	c	R/TD/DST	FN, CL	Chandler et al. (2013)
Multi-scale community occupancy	ψ_i, θ_i	p_i	O1	R/TD, NS	FN, CL	Latif et al. (2020)
Hierarchical distance sampling and time-to-detection	N	$r_{[a]}$, $r_{[p]}$	O1	R/TD, DST	FN, CL, DistMA	Chandler et al. (2011), Amundson et al. (2014)

TABLE 3 (Continued)

Model	Sub-models		Data	Requirement[d]	Additional assumptions[e]	Reference(s)
	Ecological[a]	Observation[b]	Type[c]			
Spatial dynamic N-mixture with interspecific interactions	$N_i, \lambda, \delta, \phi_i$	r, r_{01}	c	R/TD, RT	CL	Zhao et al. (2022)
Nψ-mixture reaction-diffusion	ψ, N, δ, λ	p, r	c	R	FN, CL	Williams et al. (2017)

[a]Ecological sub-models are occupancy probability (ψ), occupancy of secondary sampling units, wherein ψ is then occupancy of primary sampling units (θ), abundance or density (N), initial occupancy (ψ0), colonization (γ), extirpation (ε), initial abundance (N0), survival (S), recruitment (R), population growth (λ), movement or dispersal (δ), occupancy (ψ, or θ_i) for each species represented in a community model, wherein species-level sub-models are linked with community-level hyper-parameters, or species interactions whereby the abundance of one species positively or negatively influences the abundance of another (ϕ_i).

[b]Observation sub-models are detectability of the species (p), false positive error in detection or non-detection data (e.g., species misidentification; p_{01}), false positive error in counts (e.g., double counting of individuals; r_{01}), detectability of an individual (r), availability an individual for detection (e.g., birds singing for aural detection; $r_{[a]}$), surveyor perceptibility of an individual ($r_{[p]}$), or detectability (p_i) for each species represented in a community model, wherein species-level sub-models are linked with community-level hyper-parameters.

[c]Data types are binary detection data (01) or count data (c).

[d]Data requirements are repeat or replicate surveys (R), distances to detections (DST), time to detection (TD), which is often discretized and formulated as replicate surveys with replicates after the first detection discarded (i.e., removal sampling), repeat sampling of sampling units across multiple sampling periods (RT), or nested sampling of secondary within primary sampling units (NS). Additionally, data not necessarily required for meaningfully precise estimates and parameter identifiability are repeat or replicate surveys ($R_{(+)}$) or a subsample of confirmed observations (ConfObs(+)).

[e]Assumptions required are false negative errors (i.e., incomplete detection) only (FN), closure assumption, i.e., no change in the ecological state between replicate counts at a survey unit within a sampling period (CL), perfect detectability at distance = 0.

(PDg(0)), or distance measurements are accurate and not influenced by the surveyor (DistMA).
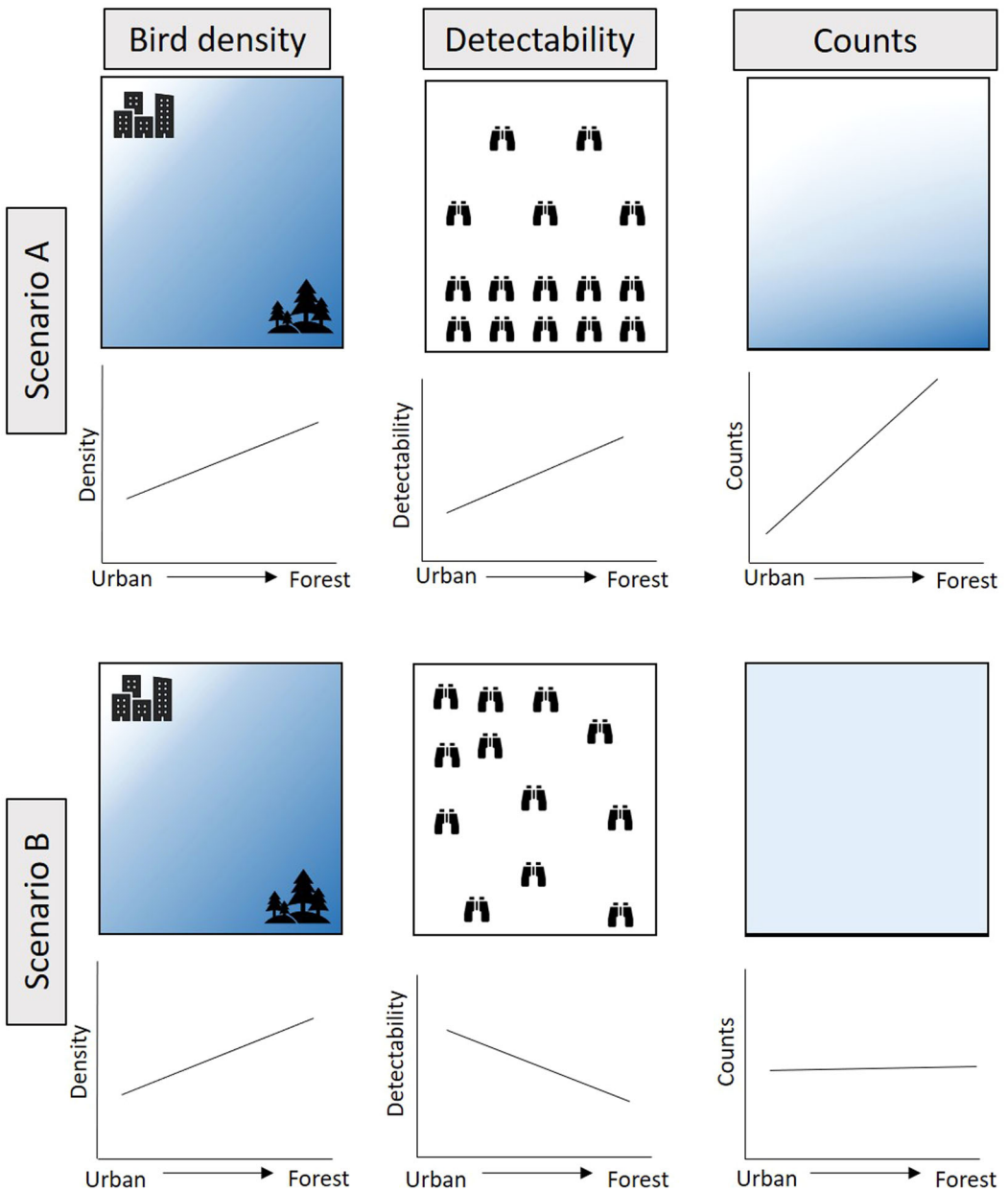
FIGURE 2 Hypothetical scenarios demonstrating how detectability can confound estimation of environmental relationships with data from count surveys of unmarked animal populations. Scenario A shows a hypothetical scenario where detectability is positively correlated with abundance, whereas scenario B shows the opposite. Both represent situations where explicit observation models would be critical for accurately estimating the depicted environmental relationship with population density.

Using analytical models to derive ecological inference inevitably requires making assumptions, so researchers must consider whether their data likely conform sufficiently to key model assumptions to provide information useful for answering their question. Providing a comprehensive list of all assumptions is beyond our scope, but basic assumptions for virtually all analytical models include that data accurately represent the study population and that

residual variation not accounted for by the model is independent among selected sampling units (i.e., independence). Hierarchical models require additional assumptions (Table 3). Probabilistic sampling designs provide clear advantages over opportunistic sampling for meeting assumptions of independence and representativeness of the study population (Williams and Brown 2019). Adding structure to models (e.g., covariates, random effects, autocorrelation parameters; Royle and Kéry 2007, Guélat and Kéry 2018) can help meet independence assumptions while also focusing models on the research questions. Additionally, we may be able to adjust our interpretation of model parameters to relax the closure assumption while maintaining or even improving inference (Latif et al. 2016, Valente et al. 2017, Steenweg et al. 2018, Fogarty and Fleishman 2021). Simulations before analysis and model checking afterwards are important tools for evaluating sufficient adherence to assumptions to support valid inference for addressing the study question (Conn et al. 2018, DiRenzo et al. 2023).

TEMPORAL DESIGN

We refer to the frequency, distribution, and extent of sampling over time as the temporal design. The temporal extent of the study—or the study period (usually number of years for species with seasonal phenology; Table 1)—determines the time period across which the study population is represented, and thus how broadly we can infer population change (e.g., dynamics and trend) from the data. The sampling period (Table 1) represents the unit of time within which the sample is collected (i.e., individuals are counted) at each selected sampling unit, and as such intersects with species ecology to determine interpretation of fundamental parameters estimated with count data. With a short sampling period that excludes the potential for individuals moving in or out of the sampling unit during sample collection, fundamental parameter estimates represent populations for a snapshot in time. Such estimates are especially relevant for inferring population density, population size, or the physical distribution of individuals at any given time (Buckland 2006, Latif et al. 2016). With longer sampling periods that allow movement during sample collection, fundamental parameter estimates instead represent the population within range of units and are more relevant to questions concerning species geographic range, habitat use, and metapopulation dynamics (Valente et al. 2017). Finally, the sampling frequency (Table 1) describes how many and how often sampling periods occur at a sampling unit within the study period. Sampling frequency determines the resolution at which we can estimate population change. For example, annual sampling is needed to quantify inter-annual population dynamics (MacKenzie et al. 2003, Dail and Madsen 2011), whereas less-frequent sampling may be sufficient to measure longer term population trends (Gitzen et al. 2012). The optimal ratio of spatial to temporal intensity of sampling increases with increasing spatial heterogeneity of the study population (Rhodes and Jonzén 2011). Simulations can help researchers navigate tradeoffs between different aspects of sampling and temporal design (e.g., sample size vs. study period; Ross and Weegman 2022). For most ecological systems, seasonal timing of sampling will additionally constrain inference (e.g., breeding season sampling will represent breeding populations). Analogous to our distinction between sampling and survey units, we recommend focusing on ecological processes of interest to the research question when defining sampling period and sampling frequency, and waiting to consider logistical constraints on surveys and observation processes when defining the survey period (Table 1) and survey frequency (Table 1).

SURVEY PROTOCOL

The survey protocol encompasses the process of conducting counts within selected sampling units and sampling periods. Researchers need a survey protocol capable of generating data necessary to answer their questions, but unlike sampling design and temporal design, logistical constraints and observation processes are more relevant to development of the survey protocol.

We suggest first defining the survey period and survey unit, which (as noted in the sampling design section) may reflect logistical constraints imposed on the surveyor and how we want model parameters to reflect observation processes. For small territorial breeding landbirds, surveyors typically implement counts while standing at points representing the center of circular survey units (i.e., point counts; Ralph et al. 1993, Rosenstock et al. 2002, Matsuoka et al. 2014). For other systems (e.g., non-territorial birds, waterbirds, large mammals), surveyors may travel along linear transects or meander across areal units (Dunn et al. 2005, Certain and Bretagnolle 2008, Brown et al. 2017, Harris et al. 2021, Schmidt et al. 2022). Researchers often limit the size and shape of survey units based in part on *a priori* knowledge of the perceptual limitations of surveyors or survey equipment (e.g., automated recording units, cameras). Logistical considerations may also shape the survey period, such as the need to distribute limited surveyor time across multiple survey units. Re-considering the independence assumption required by most analytical models can benefit researchers as they define survey units. Conversely, misunderstanding these assumptions can mislead study planning (Appendix A). If survey units are much smaller than a sampling unit, researchers may need multiple survey units per selected sampling unit for adequate sampling. Moreover, survey units may also serve as secondary sampling units for examining spatial distribution at multiple spatial scales (Pavlacky et al. 2012, 2017).

The survey period entails tradeoffs between detectability and assumptions about the distribution of individuals in space, and thus affects how counts and count-based model estimates reflect surveyor perception and behavior. Relatively long surveys allow observers more time to look and listen for individuals present within the survey unit, potentially increasing detectability within the survey. Long surveys, however, may also allow individuals to move in or out of the survey unit, expanding the subset of the population potentially detected and resulting in ambiguity regarding the spatial area sampled (Schmidt et al. 2013). Additionally, the likelihood of false positives in count data may increase with survey period length as it becomes increasingly difficult to keep track of individuals. Considering the typical variability in the observation process with seasonal and daily fluctuations in behavior and activity, researchers will likely benefit by constraining the survey period to when individuals are most readily detectable (e.g., morning hours for territorial breeding songbirds).

Survey protocols for state-space hierarchical models usually need to include auxiliary data collection or additional structure to inform observation sub-models. Distance sampling models require distance measurements from the survey unit center (typically a point or line transect) to each detected individual (Buckland et al. 2015). Covariates can also provide auxiliary information required by some models (Lele et al. 2012) with additional assumptions required (Knape and Korner-Nievergelt 2015). Replicated counts required for observation sub-models in N-mixture and occupancy models can arise from various protocols, including repeat surveys (Valente et al. 2017), double-observer surveys (Nichols et al. 2000), recording the time to first detection (Farnsworth et al. 2002, Latif et al. 2020), or (with caution) spatial replication (Kendall and White 2009, Guillera-Arroita 2011). We emphasize that researchers may analyze data of similar dimensions with a model that is mathematically identical across these protocols, but they may nevertheless generate estimates with different interpretations largely because of differences in survey period, which alters underlying observation processes (see previous paragraph). The survey frequency (i.e., number of repeat surveys or replicate counts within the sampling period; Table 1) determines the overall probability of detecting an individual or species where present at least once during the sampling period (i.e., p^* in MacKenzie and Royle 2005). Estimates of fundamental parameters are biased when overall detection probability is too low, so literature guiding replicate counts to provide data for detectability estimates focuses heavily on optimizing survey frequency (MacKenzie and Royle 2005, Bailey et al. 2007, Guillera-Arroita et al. 2010, Sanderlin et al. 2014, Reich 2020).

ASSESSING FEASIBILITY WITH SIMULATIONS

Simulations can be invaluable for assessing potential study plans. By comparing model estimates fitted with simulated data to the true parameter values used to generate those data, we can quantify various estimator properties, including precision, accuracy, and statistical power. Simulations can represent either ideal scenarios for

evaluating basic estimator properties or non-ideal scenarios for evaluating how estimators hold up when model assumptions are violated (Bailey et al. 2007, Efford and Dawson 2012, Brown et al. 2017, Latif et al. 2018, Fogarty and Fleishman 2021). Moreover, simulations representing even a limited range of plausible scenarios can provide researchers the opportunity to familiarize themselves with their intended analytical model while evaluating feasibility of drawing meaningful inference (DiRenzo et al. 2023). Despite the broad availability of computational tools, implementing extensive custom simulations for every study is probably infeasible for most researchers and unnecessary for well-established sampling designs. At the very least, researchers should be familiar with published simulation work relevant to sampling and analysis approaches under consideration. Much of the foundational literature for hierarchical models includes simulation studies that compare alternative sampling designs and survey protocols (Kéry and Royle 2016, 2020; MacKenzie et al. 2018). Simulation studies for observation sub-models informed with replicate counts describe minimum levels of replication for sufficient detectability to reliably estimate fundamental parameters (e.g., MacKenzie et al. 2002, Royle and Nichols 2003). Similarly, distance sampling literature describes minimum sample sizes needed to reliably estimate density (10 survey units and 60 independent detections; Buckland et al. 2001). Once acquainted with foundational literature, researchers can then gauge the value of additional tailored simulations to guide their study plan (e.g., Ross and Weegman 2022) given budget constraints (e.g., Sanderlin et al. 2014).

EXAMPLE STUDY PLANS

Red-faced warbler

Research question

In our first hypothetical example, we considered researchers working in the southwestern United States with red-faced warbler (*Cardellina rubrifrons*) as their study species. Red-faced warblers are a species of concern owing to expected habitat loss and degradation with interacting effects of climate change and increased risk of high-severity wildfire (Kirkpatrick and Conway 2010, Ganey et al. 2015). Red-faced warblers nest on the ground in higher elevation forests dominated by coniferous trees intermixed with oaks (Ganey et al. 2015), where they occupy uniformly distributed breeding territories (Flesch 2019).

The researchers were interested in studying the distribution of red-faced warbler in relation to fire history, particularly burn severity and time since wildfire. Additionally, in light of theory describing habitat selection as occurring first at coarser scales and subsequently at finer scales (Cody 1985), the researchers wanted to distinguish spatial scales at which red-faced warbler distributions exhibit the strongest relationships with fire history. Although negatively affected by high-severity fire, red-faced warblers may also respond negatively to small-diameter thinning treatments (Kalies et al. 2010) intended to reduce burn severity. The researchers therefore expected results from this study to help guide fuels management aimed at restoration (Villarreal et al. 2020) to reduce overall fire severity risk.

Population metric

Reflecting their primary interest in quantifying the spatial distribution of red-faced warblers, the researchers chose density as their target ecological metric. Given the small size of home ranges, researchers expected to routinely count >1 individual within relatively small areas. The researchers therefore expected density to be appropriate for fully quantifying spatial and temporal patterns anticipated in expected data. The researchers considered but ultimately decided against other metrics, including occupancy (better suited to rare or sparsely distributed species), relative abundance (requires assuming detectability of red-faced warbler is unrelated to burn severity and time since wildfire), trend (requires more

years than possible with available funding), and population vital rates of persistence and growth (requires sampling the same units successively over multiple years, potentially limiting spatial representation of the population, and not necessarily required for evaluating habitat quality; Bock and Jones 2004).

Sampling design

The researchers identified all red-faced warblers occurring within one Sky Island mountain in a high elevation conifer forest as their study population. An analysis of remotely sensed land cover verified sufficient range of burn severity and time since fire within one mountain sufficient to meet study objectives. Sampling units consisted of 1 × 3-km rectangular cells large enough to include multiple breeding territories yet small enough for units to represent distinct levels of burn severity and time since fire. Each sampling unit contained a set of 12 evenly spaced points (500 m apart; Figure 3). In addition to facilitating efficient allocation of survey effort, points served as secondary sampling units for finer scale relationships with wildfire. The researchers' sampling frame consisted of all possible sampling units within the Sky Island mountain containing their study population. Burned forest stands of varying severity and unburned stands were not evenly distributed in space, so researchers stratified sampling to select an equal number of sampling units representing 4 levels of burn severity: unburned, low, medium, and high. Researchers selected sampling unit center points within each of these strata using a spatially balanced approach (Stevens and Olsen 2004, Theobald et al. 2007), and oriented units randomly while constraining them to not overlap neighboring units. Thus, sampling units could include multiple levels of burn severity but would span the full severity gradient present within the sampling frame. The researchers conducted a simulation-based power analysis with parameters reflecting pilot data to identify a sufficient sample size to estimate abundance relationships with burn severity and time since wildfire.
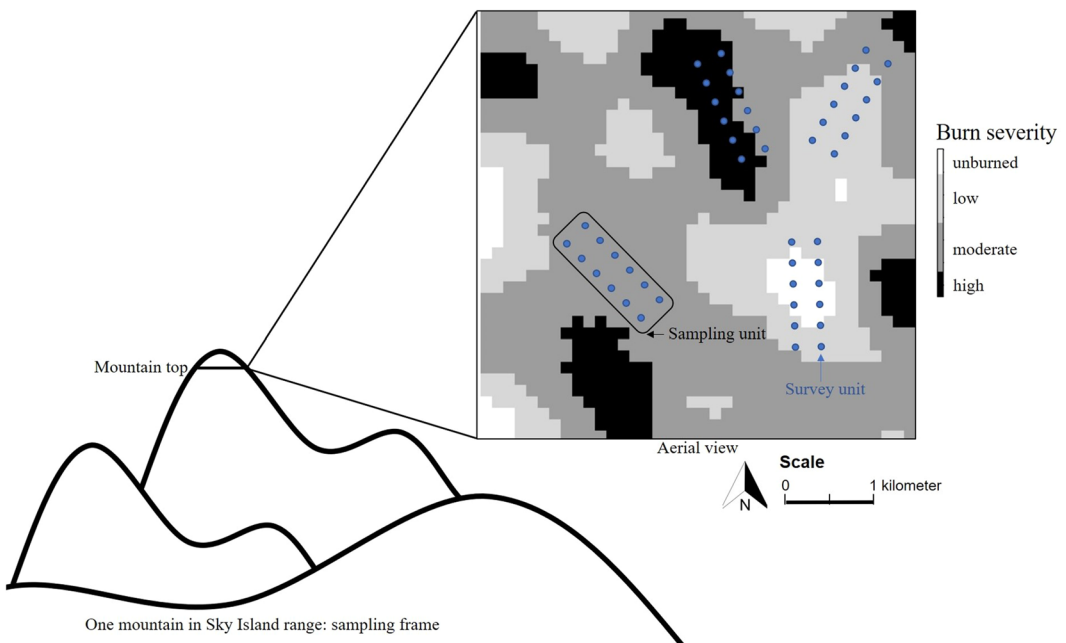

FIGURE 3   A portion of the sampling frame in one Sky Island mountain range (Arizona, USA) with the spatial arrangement of sampling units across different burn severities for a hypothetical study of red-faced warblers. Sampling units are rectangular cells containing evenly distributed points where surveys were centered.

Analytical framework

The researchers created a diagram of the relevant ecological and observation processes expected to influence count data (Figure 4). The primary ecological processes of interest are wildfire altering habitat conditions, and consequently affecting the spatial distribution of the population at 2 different spatial scales. Red-faced warbler breeding territories are generally distributed uniformly within suitable areas, so the researchers did not anticipate any spatial variation in density beyond what could be explained by variable habitat conditions.

During breeding, male red-faced warblers sing regularly, making them easily heard and visible. The researchers expected training to further boost surveyor ability to detect and distinguish red-faced warblers, particularly males, although detectability of red-faced warblers was expected to decline with distance from the surveyor. Moreover, they expected the steepness of this decline to depend on vegetation density (more vegetation obstructs vision and potentially reduces how far sound travels), burn severity and time since fire (via effects on vegetation density), wind (reduces how far sound travels), and observer (skill may vary even after training). The possibility of burn severity and time since fire affecting detectability was especially problematic for this study. With reduced vegetation density in recently burned areas, the researchers expected detectability to increase with increasing burn severity and decreasing time since fire, potentially obscuring relationships between population density and fire. Despite being readily detectable during breeding, the researchers expected singing rate to vary with time of day and progression of the breeding season, potentially adding variation to detectability. The researchers expected restricting sampling
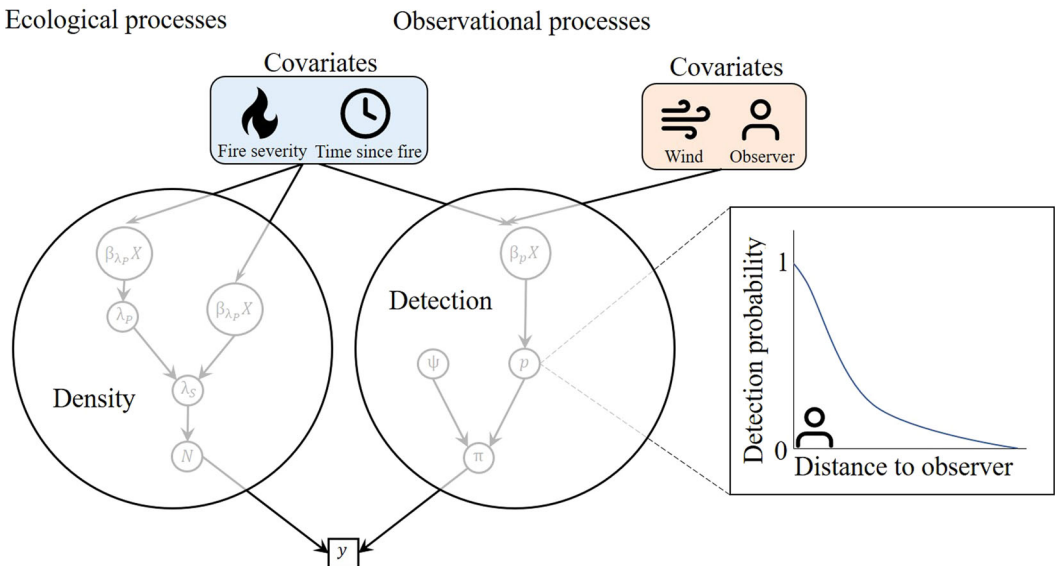

FIGURE 4 Directed acyclic graph of the distance sampling model intended for analyzing hypothetical red-faced warbler count data (Sky Island mountain range, Arizona, USA) with relevant ecological and observational processes highlighted. Data include number of individuals (y) detected at survey units in different distance classes, and covariates. Ecological parameters represented in the model are expected abundance at primary and secondary sampling units (i.e., 2 spatial scales: λ_P and λ_S), realized abundance red-faced warblers present within a survey unit, which coincides with the survey unit (N), and covariate relationships with expected abundance ($\beta_{\lambda_P} X$ and $\beta_{\lambda_S} X$). Observation parameters are the detection probability in each distance class (p), the distribution of individuals present among distance classes (ψ), the expected proportion of detected individuals among distance classes (π), and covariate relationships with detectability ($\beta_p X$). Arrows indicate dependencies among parameters (circle nodes) and data (square nodes). The inset depicts the decline of detection probability with distance from the observer.

to the height of the breeding season would limit this source of variation, but daily variation in singing rate could remain a factor of concern.

The researchers selected hierarchical distance sampling (Royle et al. 2004) as their modeling approach for estimating population density while accounting for observation processes (Figure 4). This model assumes perfect detectability for individuals located at distance = 0 from the surveyor, and then estimates the decline in detectability with increasing distance using distances to detected individuals recorded by the surveyor. The model allows covariates to modulate how steeply detectability declines with distance while also allowing researchers to estimate covariate relationships with population density. The researchers included burn severity and time since fire as covariates of density at primary and secondary spatial scales, and detectability to account for potential confounding effects of fire on ecological and observation processes. Although sampling was stratified among burn severity classes, the researchers identified a continuous measure of burn severity for use as a modeling covariate. Field measurements of wind speed and observer identity represented additional covariates for detectability (Figure 4). Assuming a relatively uniform distribution of red-faced warblers within suitable areas and assuming fire is the primary driver of habitat condition, the researchers did not anticipate density to vary spatially beyond what could be explained by variation in burn severity and time since fire. Nevertheless, they planned to evaluate model goodness of fit (Buckland et al. 2015) and include additional parameters to account for additional sources of variation (e.g., Pacifici et al. 2016) as needed for adequate fit.

Considering the particularly high detectability of singing males, the researchers planned to focus density estimation on breeding males and assume a balanced sex ratio to infer population density. Infrequent or variable singing could limit detectability and more importantly compromise the assumption of perfect detectability at distance = 0 (Amundson et al. 2014). By restricting sampling to the height of the breeding season and carefully planning surveys (see Survey protocol below), the researchers hoped to sufficiently adhere to the model assumption of perfect detectability of males at distance = 0. Nevertheless, the researchers planned to collect auxiliary data needed to check this assumption (i.e., time to detection) and potentially add a sub-model for incomplete temporal availability for detection if needed.

Temporal design

The researchers deemed a study period of 3 years as sufficient to meet their primary objective of quantifying spatial distribution along an environmental gradient. In particular, 3 years ensured estimated habitat relationships represented more than just 1 potentially aberrant year. Given their interest in estimating population density for a snapshot in time, the researchers wanted to limit the potential for individuals moving into range of the surveyor during sampling, so they restricted the sampling period to a short time window corresponding with a single survey on a single day of the breeding season. Red-faced warblers are migratory and found within the sampling frame only during the breeding season, when breeding behavior (e.g., singing) makes them particularly detectable. The researchers therefore restricted sampling periods to occur within a 5-week period at the center of the 9-week breeding season (i.e., excluding the first 2 and last 2 weeks of the breeding season). In addition to ensuring representation of the breeding season population, this sample timing maximized the intensity of breeding behavior during sampling. The researchers planned to sample each unit once during the study period (i.e., sampling frequency = 1), allowing sampling of different units in each year and thus maximizing spatial representation of the population.

Survey protocol

The researchers defined a survey unit as a circular plot centered on 1 of the 12 evenly distributed points within a selected sampling unit. Rugged terrain makes travel difficult, so clustering multiple survey units within

larger (primary) sampling units limited logistical demands associated with travel between surveys, along with benefits for ecological inference of scale-specific patterns (see above). Considering logistical constraints (e.g., travel time to and among sampling units, number of observers), the researchers expected that <12 survey units may be completed in some sampling units, and specified a minimum of 5 surveys for inclusion of a sampling unit.

The researchers selected 5 minutes as their survey period because this length of time was sufficient for all males within aural range of the surveyor to sing while being short enough to exclude substantial movement of singing males towards the surveyor during the survey. To meet requirements of their distance sampling model, surveyors recorded distances to each red-faced warbler they detected using a laser range finder, while planning to confirm the location of aurally detected individuals as needed after each survey. Surveyors recorded all detected individuals along with their sex (distinguished by plumage and behavior; Martin and Barber 2020) to allow the analysis to focus on males. Surveys were constrained to occur in the morning (i.e., 30 minutes before sunrise to 3 hours after sunrise) when singing rate is highest to limit variability in detectability and meet the required assumption of perfect detectability at distance = 0. The surveyors also recorded the timing of detections within the 5-minute survey period to allow researchers to verify the model assumption of perfect detectability at distance = 0 (Amundson et al. 2014).

Assessing feasibility

As a final step, the researchers assessed the feasibility of their study. They verified selected sampling units represented an adequate range of burn severity and time since fire values to address their research question. Conducting surveys in a mountainous setting can be logistically demanding in unexpected ways, so they conducted a pilot season (Sanderlin et al. 2019). They developed contingency plans for possible unexpected events, such as observers dropping out of the study or a new wildfire burning selected units. They also simulated data with parameters reflecting a pilot study to verify sufficient power to answer their research question with intended sampling design, analysis approach, and survey protocol. The researchers defined sufficient power as an 80% chance of an 80% credible interval for wildfire covariate relationships to exclude the null hypothesis given 30% more abundance with an increase in any wildfire covariate of 1 standard deviation.

Red-shouldered hawk

Research question

In our second hypothetical example, we consider researchers examining temporal population dynamics of red-shouldered hawk (*Buteo lineatus*) populations breeding in the eastern United States. Studies have reported decreasing range-wide population trends (Dykstra et al. 2020, Sauer et al. 2020), raising conservation interest. Eastern populations breed in a variety of forest ecosystems, particularly in mature deciduous or mixed coniferous-deciduous forests in flooded areas or along riparian corridors but also in forests embedded in urban landscapes (Dykstra et al. 2020).

The researchers were interested in examining the population dynamics and spatial distribution of this species along a rural–urban gradient. They hypothesized that populations in urban areas are increasing compared to rural populations that are either stable or declining. They expected information on spatial variation in population dynamics would elucidate source-sink dynamics needed to infer spatial distribution and prioritization of conservation action. Additionally, the researchers wanted to quantify the overall population trajectory to assess conservation status and guide prioritization of their study population for conservation action relative to other populations described in the literature.

Population metric

The researchers identified species occupancy dynamics as their focus for quantifying population dynamics. Specifically, they decided to measure spatial variation in colonization (probability of unoccupied sites becoming occupied) and extirpation (probability of occupied sites becoming unoccupied; MacKenzie et al. 2003). The researchers considered occupancy dynamics an appropriate focus because red-shouldered hawks are wide-ranging and sparsely distributed, so data were expected to be binary (detection or non-detection of a breeding pair). In contrast, collecting sufficient data to quantify dynamics in abundance (Dail and Madsen 2011) appeared less feasible for this species. To allow inference of population status and trajectory from occupancy, the researchers planned to measure occupancy at a spatial grain approximating the size of a breeding territory (Linden et al. 2017).

Sampling design

The researchers defined their study population as all red-shouldered hawks within the area encompassed by the Virginia peninsula in southeastern Virginia, USA (Figure 5), which is bounded by the York River to the north, the James River to the south, and by major metropolitan areas of Hampton and Richmond, to the east and west, respectively. This area consisted of a coastal plain ecosystem with patches of eastern deciduous forest interspersed by agriculture, wetlands, and low- to high-density urban areas. The study population consisted of all red-shouldered hawks within this area, which are a resident species in this system.
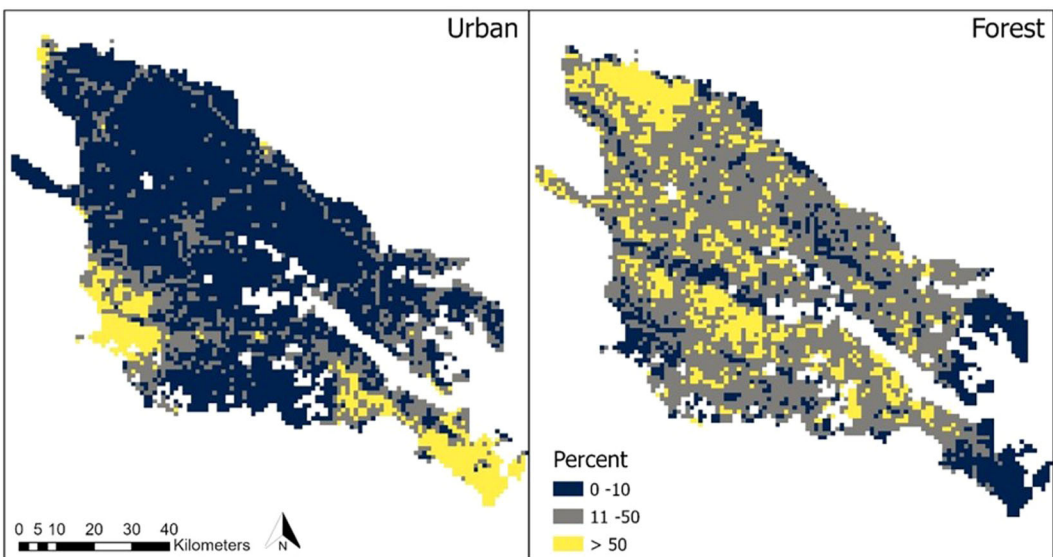

FIGURE 5 Sampling frame for a hypothetical study of red-shouldered hawk occupancy dynamics in the eastern United States. Sampling units represent 1.3-km^2 cells, approximating the size of a typical breeding territory, superimposed on rural–urban and non-forest–forest gradients. Forest is defined as National Land Cover Data (NLCD) values 41 and 43 and urban as NLCD values 21–24 (NLCD 2019). Panels represent cell-specific percent urban (left) and percent forest (right) discretized into bins representing low (0–10%), medium (11–50%), and high values (>50%). Crossing these bins generated the 8 strata across which sampling was stratified (ninth bin of high urban and high forest did not exist on the landscape).

The researchers identified their sampling units as 1.3-km^2 grid cells, approximating the size of a typical breeding territory (Dykstra et al. 2020). This spatial grain minimizes the number of breeding territories that could intersect a sampling unit, and thus allows stronger inference of population distribution and dynamics from occupancy (Linden et al. 2017). The researchers defined their sampling frame as all non-overlapping grid cells contained in their study area (Figure 5). By defining their sampling frame as a discrete set of non-overlapping sampling units (i.e., a well-defined sampling frame; Thompson et al. 1998), the researchers aimed to facilitate estimation of total area occupied to more clearly infer conservation status and trajectory of their study population. Additionally, the researchers ensured representation of environmental conditions of interest by stratifying sampling across 8 strata representing urban–rural and forest cover gradients. They mapped percent forest and urban land covers within 1.3-km^2 cells discretized into low (0–10%), medium (11–50%), and high percent (>50%; Figure 5). Cells representing both high proportion urban and high proportion forest did not exist on the landscape, so crossing the 3 bins for the 2 gradients generated 8 strata. Spatially balanced sampling (Stevens and Olsen 2004) within each stratum generated a sample of 240 units for the study. By consulting foundational literature describing dynamic occupancy models and published studies estimating occupancy dynamics for raptors, the researchers verified this sample size was reasonable for estimating distribution and dynamics along urban–rural and forest land cover gradients.

Analytical framework

The researchers visualized ecological and observation processes expected to influence red-shouldered hawk detection data in a directed acyclic graph (Figure 6). The primary ecological processes of interest were
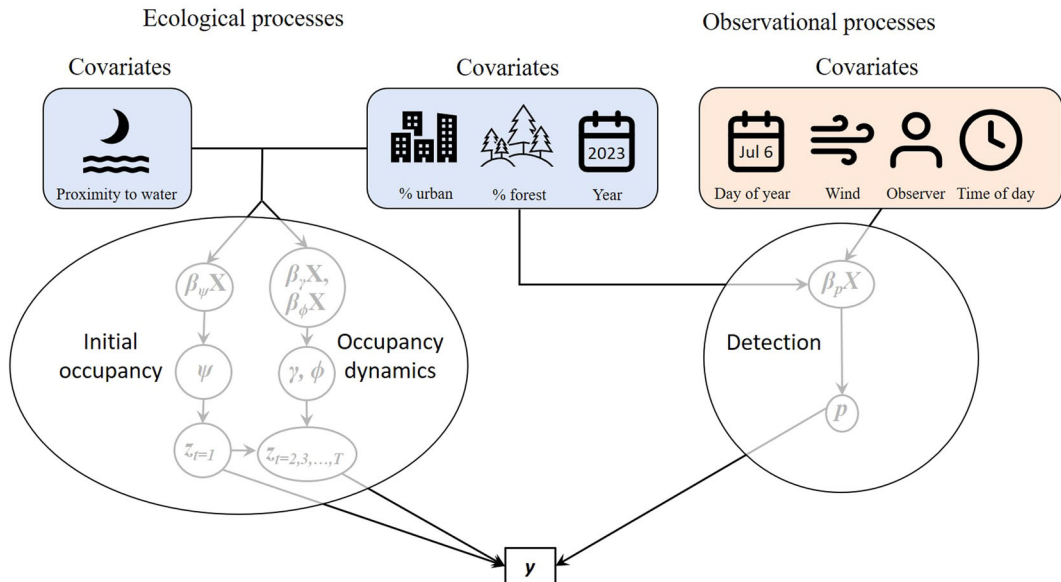

FIGURE 6 Directed acyclic graph representing ecological and observation processes for a hypothetical study of occupancy dynamics for red-shouldered hawks in the eastern United States. Expected data are year-specific detection histories (y). Occupancy model parameters representing ecological processes are occupancy probability in year 1 (ψ), colonization (γ) and extirpation (ϕ) probabilities for years 2–T, covariate relationships with these ($\beta_\psi X$, $\beta_\gamma X$, and $\beta_\phi X$), and occupancy states for sampling units in each year (z_t). Parameters representing observation processes are detection probability (p) and covariate relationships with detection ($\beta_p X$). Arrows indicate dependencies among parameters (circle nodes) and data (square nodes).

occurrence of breeding territories, space use of red-shouldered hawks within breeding territories (i.e., components of spatial distribution), annual changes in territory and space use distributions, and environmental effects on these processes. The researchers expected spatial distribution to relate positively with forest, particularly those associated with aquatic land cover types (e.g., flooded areas, riparian corridors, storm ponds), reflecting a diet that includes aquatic prey (Dykstra et al. 2020). The researchers also expected distributional shifts towards urban and away from rural areas combined with an overall population decline resulting in a decline in total area occupied.

The researchers expected red-shouldered hawk detectability to be lower in forested areas because of decreased visibility, and also potentially correlate with urbanization because urbanization affects vegetation structure (affecting visibility) and resource availability (affecting red-shouldered hawk behavior). In addition, the researchers anticipated detectability to vary among years, reflecting turnover of surveyors and consequent fluctuations in skill and experience, and annual variation in environmental conditions and behavior associated with resource availability. The researchers planned to incorporate call broadcasts in their surveys to boost detectability (see Survey protocol), but red-shouldered hawks are more responsive to call broadcasts before trees are fully leafed (McLeod and Andersen 1998), potentially causing a negative detectability relationship with survey date. The researchers also expected activity and therefore detectability to be lower on windy days.

The researchers planned to use a dynamic occupancy model (MacKenzie et al. 2003) to analyze their data and quantify ecological parameters of interest while accounting for relevant observation processes (Figure 6). This model assumes that the true occupancy status of a site remains constant within the sampling period and that changes among replicate surveys within a sampling period are therefore due to incomplete detectability. Replicate surveys therefore need to be spaced sufficiently in time to avoid dependencies from red-shouldered hawk movement speed limitations and behavior.

The model estimates occupancy in the initial year as a function of spatially varying covariates, with occupancy in subsequent years estimated from colonization and extinction rates, which can vary with spatial or temporal covariates. Considering their objectives, the researchers related both initial occupancy and occupancy dynamics (colonization and extirpation) with spatial covariates (percent urbanization, percent forest, and proximity to water), and occupancy dynamic parameters were additionally allowed to vary annually.

The researchers modeled detection probability as a function of covariates representing expected observation processes. They related detectability as a linear function of forest cover, percent urbanization, day of year, wind speed, and time of day. The researchers also planned to include categorical effects of year and observer as covariates of detectability (either as fixed or random effects; Zuur et al. 2009).

Temporal design

To provide time series data needed to quantify occupancy dynamics, the researchers planned a study period of 5 years. Their sample selection algorithm generated an ordered list of units, whereby any ordered subset of units would itself represent a spatially balanced sample (Stevens and Olsen 2004). Thus, the researchers could reduce the number of units surveyed in each stratum in any given year if needed to deal with unexpected logistical constraints while maintaining an annual sampling frequency for a core set of spatially balanced units representing the gradients of interest. The researchers expected an annual sampling frequency for a core set of sampling units to allow estimation of yearly rates of colonization and extirpation while also including data from units sampled less frequently (MacKenzie et al. 2003). The researchers bounded their sampling period within each year by the breeding season to ensure relatively predictable movements centered on a nest (Dykstra et al. 2020), thereby better meeting the closure assumption (MacKenzie et al. 2003, 2018).

Survey protocol

The researchers planned to implement call broadcast surveys from the centers of 1.3-km^2 grid cells representing selected sampling units. Thus, each grid cell contained a single survey unit at its center. Because the researchers lacked a straightforward way to measure the reach of call broadcasts, the size of the survey unit was ambiguous. Nevertheless, the researchers expected that the minimum spacing of 1.14 km between neighboring survey unit centers (same as distance between neighboring sampling unit centers) would limit the potential for a single breeding territory to intersect >1 survey unit, so as to sufficiently meet assumptions required to infer population change from occupancy dynamics (Linden et al. 2017).

Following established methods, the researchers planned a survey period of 10 minutes, during which the surveyor would broadcast conspecific calls at approximately 100 dB 6 times for 20 seconds every minute, followed by a 4-minute and 40-second listening period, and restricted surveys to morning hours (McLeod and Andersen 1998). The researchers defined a survey detection (visual or aural) as an individual flying overhead or approaching the speaker, confirming that a breeding territory included the survey (and sampling) unit center. If necessary to stay on schedule, surveyors could discontinue surveys once they recorded a survey detection, but they planned to complete the full survey period where time allowed to confirm their assumption of territorial exclusivity.

The researchers planned to repeat surveys at a survey unit within each sampling period to provide data needed to estimate detection probability and allow interpretation of occupancy probabilities as the probability of a breeding territory including a sampling unit center. The researchers used a pilot study to evaluate how many surveys were needed for unbiased occupancy estimation. Results from the pilot study indicated 4 surveys were required for cumulative detection probabilities of $p^* \geq 0.85$ (i.e., sufficient probability of recording at least one detection where the species is present for unbiased estimation of occupancy; MacKenzie and Royle 2005). To maximize independence of detection outcomes among surveys, the researchers spaced surveys evenly across the sampling period by completing surveys of all units before returning for subsequent surveys.

Assessing feasibility

The researchers used established software inputted with detection and occupancy probability estimates from pilot data to verify sufficient power to discern occupancy dynamic relationships with environmental gradients of interest (Bailey et al. 2007). Considering that a majority of land in the sampling frame was privately owned, the researchers sought access to each survey unit from private landowners prior to the first field season. In cases where access was denied to a selected sampling unit, the researchers replaced the unit with the next unit on the list generated by their selection algorithm to maintain a spatially balanced sample (Stevens and Olsen 2004).

RESEARCH IMPLICATIONS

Contemporary hierarchical models expand researchers' capacity to ask increasingly complex and nuanced ecological questions from count surveys. We provide general guidance for collecting count data of necessary quality and structure to support desired inference with these models. Our framework outlines the major components of study planning that, considered in the order presented, can help ensure appropriate data are collected to answer research questions of interest. This framework and all of its components are demonstrated in 2 detailed examples.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ETHICS STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

Data sharing not applicable—no new data generated.

ORCID

Quresh S. Latif http://orcid.org/0000-0003-2925-5042
Gavin M. Jones http://orcid.org/0000-0002-5102-1229

REFERENCES

Amundson, C. L., J. A. Royle, and C. M. Handel. 2014. A hierarchical model combining distance sampling and time removal to estimate detection probability during avian point counts. Auk 131:476–494.

Bächler, E., and F. Liechti. 2007. On the importance of g(0) for estimating bird population densities with standard distance-sampling: implications from a telemetry study and a literature review. Ibis 149:693–700.

Bailey, L. L., J. E. Hines, J. D. Nichols, and D. I. MacKenzie. 2007. Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. Ecological Applications 17:281–290.

Bailey, L. L., D. I. MacKenzie, and J. D. Nichols. 2014. Advances and applications of occupancy models. Methods in Ecology and Evolution 5:1269–1279.

Berigan, W. J., G. M. Jones, S. A. Whitmore, R. J. Gutiérrez, and M. Z. Peery. 2019. Cryptic wide-ranging movements lead to upwardly biased occupancy in a territorial species. Journal of Applied Ecology 56:470–480.

Betts, M. G., N. L. Rodenhouse, T. S. Sillett, P. J. Doran, and R. T. Holmes. 2008. Dynamic occupancy models reveal within-breeding season movement up a habitat quality gradient by a migratory songbird. Ecography 31:592–600.

Bibby, C. J., N. D. Burgess, D. A. Hill, and S. Mustoe. 2000. Bird census techniques. Second edition. Academic Press, London, United Kingdom.

Blanco, G., F. Sergio, J. A. Sanchéz-Zapata, J. M. Pérez-García, F. Botella, F. Martínez, I. Zuberogoitia, O. Frías, F. Roviralta, J. E. Martínez, and F. Hiraldo. 2012. Safety in numbers? Supplanting data quality with fanciful models in wildlife monitoring and conservation. Biodiversity and Conservation 21:3269–3276.

Bock, C. E., and Z. F. Jones. 2004. Avian habitat evaluation: should counting birds count? Frontiers in Ecology and the Environment 2:403–410.

Brown, D. J., M. M. Cochrane, and R. A. Moen. 2017. Survey and analysis design for wood turtle population monitoring. Journal of Wildlife Management 81:868–877.

Brown, J. H. 1995. Macroecology. University of Chicago Press, Chicago, Illinois, USA.

Boyd, R. J., G. D. Powney, and O. L. Pescott. 2023. We need to talk about nonprobability samples. Trends in Ecology & Evolution 38:521–531.

Buckland, S. T. 2006. Point-transect surveys for songbirds: robust methodologies. Auk 123:345–357.

Buckland, S. T., E. A. Rexstad, T. A. Marques, and C. S. Oedekoven. 2015. Distance sampling: methods and applications. Springer International Publishing, Dordrecht, Switzerland.

Certain, G., and V. Bretagnolle. 2008. Monitoring seabirds population in marine ecosystem: the use of strip-transect aerial surveys. Remote Sensing of Environment 112:3314–3322.

Chandler, R. B., and D. I. King. 2011. Habitat quality and habitat selection of golden-winged warblers in Costa Rica: an application of hierarchical models for open populations. Journal of Applied Ecology 48:1038–1047.

Chandler, R. B., D. I. King, R. Raudales, R. Trubey, C. Chandler, and V. J. Arce Chávez. 2013. A small-scale land-sparing approach to conserving biological diversity in tropical agricultural landscapes. Conservation Biology 27:785–795.

Chandler, R. B., J. A. Royle, and D. I. King. 2011. Inference about density and temporary emigration in unmarked populations. Ecology 92:1429–1435.

Chevalier, M., Å. Lindström, T. Pärt, and J. Knape. 2019. Changes in forest bird abundance, community structure and composition following a hurricane in Sweden. Ecography 42:1862–1873.

Clare, J. D. J., P. A. Townsend, and B. Zuckerberg. 2021. Generalized model-based solutions to false-positive error in species detection/nondetection data. Ecology 102:e03241.

Cody, M. L. 1985. Habitat selection in birds. Academic Press, San Diego, California, USA.

Conn, P. B., D. S. Johnson, P. J. Williams, S. R. Melin, and M. B. Hooten. 2018. A guide to Bayesian model checking for ecologists. Ecological Monographs 88:526–542.

Crates, R., A. Terauds, L. Rayner, D. Stojanovic, R. Heinsohn, D. Ingwersen, and M. Webb. 2017. An occupancy approach to monitoring regent honeyeaters. Journal of Wildlife Management 81:669–677.

Dail, D., and L. Madsen. 2011. Models for estimating abundance from repeated counts of an open metapopulation. Biometrics 67:577–587.

Dénes, F. V., L. F. Silveira, and S. R. Beissinger. 2015. Estimating abundance of unmarked animal populations: accounting for imperfect detection and other sources of zero inflation. Methods in Ecology and Evolution 6:543–556.

Dennis, B., J. M. Ponciano, S. R. Lele, M. L. Taper, and D. F. Staples. 2006. Estimating density dependence, process noise, and observation error. Ecological Monographs 76:323–341.

DiRenzo, G. V., E. Hanks, and D. A. W. Miller. 2023. A practical guide to understanding and validating complex models using data simulations. Methods in Ecology and Evolution 14:203–217.

Dorazio, R. M., and J. A. Royle. 2005. Estimating size and composition of biological communities by modeling the occurrence of species. Journal of the American Statistical Association 100:389–398.

Duarte, A., M. J. Adams, and J. T. Peterson. 2018. Fitting N-mixture models to count data with unmodeled heterogeneity: bias, diagnostics, and alternative approaches. Ecological Modelling 374:51–59.

Dunn, E. H., C. M. Francis, P. J. Blancher, S. R. Drennan, M. A. Howe, D. Lepage, C. S. Robbins, K. V. Rosenberg, J. R. Sauer, and K. G. Smith. 2005. Enhancing the scientific value of the Christmas bird count. Auk 122:338–346.

Dykstra, C. R., J. L. Hays, and S. T. Crocoll. 2020. Red-shouldered hawk (*Buteo lineatus*), version 1.0. Account *in* A. F. Poole, editor. The birds of the world. Cornell Lab of Ornithology, Ithaca, New York, USA. https://doi.org/10.2173/bow.reshaw.01

Efford, M. G., and D. K. Dawson. 2012. Occupancy in continuous habitat. Ecosphere 3:article 32.

Ethier, D. M., and T. D. Nudds. 2017. Complexity of factors affecting bobolink population dynamics communicated with directed acyclic graphs. Wildlife Society Bulletin 41:4–16.

Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119:414–425.

Farrell, S. L., B. A. Collier, K. L. Skow, A. M. Long, A. J. Campomizzi, M. L. Morrison, K. B. Hays, and R. N. Wilkins. 2013. Using LiDAR-derived vegetation metrics for high-resolution, species distribution models for conservation planning. Ecosphere 4:art42.

Flesch, A. D. 2019. Patterns and drivers of long-term changes in breeding bird communities in a global biodiversity hotspot in Mexico. Diversity and Distributions 25:499–513.

Fogarty, F. A., and E. Fleishman. 2021. Bias in estimated breeding-bird abundance from closure-assumption violations. Ecological Indicators 131:108170.

Furness, R. W., and J. J. D. Greenwood, editors. 1993. Birds as monitors of environmental change. Springer, Dordrecht, Netherlands.

Ganey, J. L., W. M. Block, J. S. Sanderlin, and J. M. Iniguez. 2015. Comparative nest-site habitat of painted redstarts and red-faced warblers in the Madrean Sky Islands of southeastern Arizona. Western North American Naturalist 75:291-300, 210.

Gitzen, R. A., A. B. Cooper, J. J. Millspaugh, and D. S. Licht. 2012. Design and analysis of long-term ecological monitoring studies. Cambridge University Press, Cambridge, United Kingdom.

Guélat, J., and M. Kéry. 2018. Effects of spatial autocorrelation and imperfect detection on species distribution models. Methods in Ecology and Evolution 9:1614–1625.

Guillera-Arroita, G. 2011. Impact of sampling with replacement in occupancy studies with spatial replication. Methods in Ecology and Evolution 2:401–406.

Guillera-Arroita, G., J. J. Lahoz-Monfort, J. Elith, A. Gordon, H. Kujala, P. E. Lentini, M. A. McCarthy, R. Tingley, and B. A. Wintle. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecology and Biogeography 24:276–292.

Guillera-Arroita, G., M. S. Ridout, and B. J. T. Morgan. 2010. Design of occupancy studies with imperfect detection. Methods in Ecology and Evolution 1:131–139.

Hamel, P. B., W. P. Smith, D. J. Twedt, J. R. Woehr, E. Morris, R. B. Hamilton, and R. J. Cooper. 1996. A land manager's guide to point counts of birds in the Southeast. 120. U.S. Department of Agriculture, Forest Service, Southern Research Station, New Orleans, Louisiana, USA.

Hanski, I. 1998. Metapopulation dynamics. Nature 396:41–49.

Harris, S. J., D. Massimino, D. E. Balmer, M. A. Eaton, D. G. Noble, J. W. Pearce-Higgins, P. Woodcock, and S. Gillings. 2021. The breeding bird survey 2020. BTO Research Report 736. British Trust for Ornithology, Thetford, United Kingdom.

Hostetler, J. A., and R. B. Chandler. 2015. Improved state-space models for inference about spatial and temporal variation in abundance from count data. Ecology 96:1713–1723.

Hutto, R. L., and D. A. Patterson. 2016. Positive effects of fire on birds may appear only under narrow combinations of fire severity and time-since-fire. International Journal of Wildland Fire 25:1074–1085.

Irvin, E., K. R. Duren, J. J. Buler, W. Jones, A. T. Gonzon, and C. K. Williams. 2013. A multi-scale occupancy model for the grasshopper sparrow in the Mid-Atlantic. Journal of Wildlife Management 77:1564–1571.

Iknayan, K. J., M. W. Tingley, B. J. Furnas, and S. R. Beissinger. 2014. Detecting diversity: emerging methods to estimate species diversity. Trends in Ecology & Evolution 29:97–106.

Johnson, D. H. 2008. In defense of indices: the case of bird surveys. Journal of Wildlife Management 72:857–868.

Johnston, A., N. Moran, A. Musgrove, D. Fink, and S. R. Baillie. 2020. Estimating species distributions from spatially biased citizen science data. Ecological Modelling 422:108927.

Kalies, E. L., C. L. Chambers, and W. W. Covington. 2010. Wildlife responses to thinning and burning treatments in southwestern conifer forests: a meta-analysis. Forest Ecology and Management 259:333–342.

Kelling, S., A. Johnston, A. Bonn, D. Fink, V. Ruiz-Gutierrez, R. Bonney, M. Fernandez, W. M. Hochachka, R. Julliard, R. Kraemer, and R. Guralnick. 2019. Using semistructured surveys to improve citizen science data for monitoring biodiversity. Bioscience 69:170–179.

Kendall, W. L., and G. C. White. 2009. A cautionary note on substituting spatial subunits for repeated temporal sampling in studies of site occupancy. Journal of Applied Ecology 46:1182–1188.

Kéry, M., and J. A. Royle. 2016. Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS. Volume 1, prelude and static models. Elsevier, London, United Kingdom.

Kéry, M., and J. A. Royle. 2020. Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS. Volume 2, dynamic and advanced models. Elsevier, London, United Kingdom.

Kéry, M., J. A. Royle, M. Plattner, and R. M. Dorazio. 2009. Species richness and occupancy estimation in communities subject to temporary emigration. Ecology 90:1279–1290.

Kirkpatrick, C., and C. J. Conway. 2010. Importance of montane riparian forest and influence of wildfire on nest-site selection of ground-nesting birds. Journal of Wildlife Management 74:729–738.

Knape, J., and F. Korner-Nievergelt. 2015. Estimates from non-replicated population surveys rely on critical assumptions. Methods in Ecology and Evolution 6:298–306.

Latif, Q. S., M. M. Ellis, and C. L. Amundson. 2016. A broader definition of occupancy: comment on Hayes and Monfils. Journal of Wildlife Management 80:192–194.

Latif, Q. S., M. M. Ellis, V. A. Saab, and K. Mellen-McLean. 2018. Simulations inform design of regional occupancy-based monitoring for a sparsely distributed, territorial species. Ecology and Evolution 8:1171–1185.

Latif, Q. S., R. L. Truex, R. A. Sparks, and D. C. Pavlacky Jr. 2020. Dry conifer forest restoration benefits Colorado Front Range avian communities. Ecological Applications 30:e02142.

Lele, S. R., M. Moreno, and E. Bayne. 2012. Dealing with detection error in site occupancy surveys: what can we do with a single survey? Journal of Plant Ecology 5:22–31.

Linden, D. W., A. K. Fuller, J. A. Royle, and M. P. Hare. 2017. Examining the occupancy–density relationship for a low-density carnivore. Journal of Applied Ecology 54:2043–2052.

Link, W. A., and J. R. Sauer. 1998. Esimating population change from count data: application to the North American breeding bird survey. Ecological Applications 8:258–268.

Louvrier, J., T. Chambert, E. Marboutin, and O. Gimenez. 2018. Accounting for misidentification and heterogeneity in occupancy studies using hidden Markov models. Ecological Modelling 387:61–69.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84:2200–2207.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Baily, and J. E. Hines. 2018. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Second edition. Academic Press, London, United Kingdom.

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

Marques, T. A., L. Thomas, M. Kéry, S. T. Buckland, D. L. Borchers, E. Rexstad, R. M. Fewster, D. I. MacKenzie, J. A. Royle, G. Guillera-Arroita, et al. 2017. Model-based approaches to deal with detectability: a comment on Hutto (2016a). Ecological Applications 27:1694–1698.

Martin, T. E., and P. M. Barber. 2020. Red-faced warbler (*Cardellina rubrifrons*), version 1.0. Account *in* A. F. Poole and F. B. Gill, editors. The birds of the world. Cornell Lab of Ornithology, Ithaca, New York, USA. https://doi.org/10.2173/bow.refwar.01

Matsuoka, S. M., C. L. Mahon, C. M. Handel, P. Sólymos, E. M. Bayne, P. C. Fontaine, and C. J. Ralph. 2014. Reviving common standards in point-count surveys for broad inference across studies. Condor 116:599–608.

McFarland, T. M., H. A. Mathewson, J. E. Groce, M. L. Morrison, J. C. Newnam, R. T. Snelgrove, K. L. Skow, B. A. Collier, and R. N. Wilkins. 2012. Utilization of a species occupancy model for management and conservation. Wildlife Society Bulletin 36:432–439.

McLeod, M. A., and D. E. Andersen. 1998. Red-shouldered hawk broadcast surveys: factors affecting detection of responses and population trends. Journal of Wildlife Management 62:1385–1397.

Miller, D. A. W., K. Pacifici, J. S. Sanderlin, B. J. Reich. 2019. The recent past and promising future for data integration methods to estimate species' distributions. Methods in Ecology and Evolution 10:22–37.

Mordecai, R. S., B. J. Mattsson, C. J. Tzilkowski, and R. J. Cooper. 2011. Addressing challenges when studying mobile or episodic species: hierarchical Bayes estimation of occupancy and use. Journal of Applied Ecology 48:56–66.

Morrison, M. L., W. M. Block, M. D. Strickland, B. A. Collier, and M. J. Peterson. 2008. Wildlife study design. Springer, New York, New York, USA.

Mortelliti, A., K. Ikin, A. I. T. Tulloch, R. Cunningham, J. Stein, D. Michael, and D. B. Lindenmayer. 2016. Surviving with a resident despot: do revegetated patches act as refuges from the effects of the noisy miner (*Manorina melanocephala*) in a highly fragmented landscape? Diversity and Distributions 22:770–782.

Nichols, J. D., J. E. Hines, J. R. Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393–408.

Nichols, J. D., L. Thomas, and P. B. Conn. 2009. Inferences about landbird abundance from count data: recent advances and future directions. Pages 201–235 *in* D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Springer US, Boston, Massachusetts, USA.

Otto, C. R. V., L. L. Bailey, and G. J. Roloff. 2013. Improving species occupancy estimation when sampling violates the closure assumption. Ecography 36:1299–1309.

Pacifici, K., B. J. Reich, R. M. Dorazio, and M. J. Conroy. 2016. Occupancy estimation for rare species using a spatially-adaptive sampling design. Methods in Ecology and Evolution 7:285–293.

Pavlacky, D. C., J. A. Blakesley, G. C. White, D. J. Hanni, and P. M. Lukacs. 2012. Hierarchical multi-scale occupancy estimation for monitoring wildlife populations. Journal of Wildlife Management 76:154–162.

Pavlacky, Jr., D. C., P. M. Lukacs, J. A. Blakesley, R. C. Skorkowsky, D. S. Klute, B. A. Hahn, V. J. Dreitz, T. L. George, and D. J. Hanni. 2017. A statistically rigorous sampling design to integrate avian monitoring and management within Bird Conservation Regions. PLoS ONE 12:e0185924.

Pescott, O. L., K. J. Walker, F. Harris, H. New, C. M. Cheffings, N. Newton, M. Jitlal, J. Redhead, S. M. Smart, and D. B. Roy. 2019. The design, launch and assessment of a new volunteer-based plant monitoring scheme for the United Kingdom. PLoS ONE 14:e0215891.

Popovic, G., T. J. Mason, S. M. Drobniak, T. A. Marques, J. Potts, R. Joo, R. Altwegg, C. C. I. Burns, M. A. McCarthy, A. Johnston, et al. 2024. Four principles for improved statistical ecology. Methods in Ecology and Evolution 15:266–281.

Ralph, C. J., G. R. Geupel, P. Pyle, T. E. Martin, and D. F. DeSante. 1993. Handbook of field methods for monitoring landbirds. Report PSW-GTR-144. Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, Albany, California, USA.

Reich, H. T. 2020. Optimal sampling design and the accuracy of occupancy models. Biometrics 76:1017–1027.

Rhodes, J. R., and N. Jonzén. 2011. Monitoring temporal trends in spatially structured populations: how should sampling effort be allocated between space and time? Ecography 34:1040–1048.

Riecke, T. V., D. Gibson, M. Kéry, and M. Schaub. 2021. Sharing detection heterogeneity information among species in community models of occupancy and abundance can strengthen inference. Ecology and Evolution 11:18125–18135.

Robertson, B. A., R. L. Hutto, and J. J. Fontaine. 2010. Evaluating food availability and nest predation risk as sources of bias in aural bird surveys. Journal of Field Ornithology 81:420–429.

Rosenstock, S. S., D. R. Anderson, K. M. Giesen, T. Leukering, and M. F. Carter. 2002. Landbird counting techniques: current practices and an alternative. Auk 119:46–53.

Ross, B. E., and M. D. Weegman. 2022. Relative effects of sample size, detection probability, and study duration on estimation in integrated population models. Ecological Applications 32:e2686.

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 46:1173–1181.

Roth, T., N. Strebel, and V. Amrhein. 2014. Estimating unbiased phenological trends by adapting site-occupancy models. Ecology 95:2144–2154.

Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. Biometrics 60:108–115.

Royle, J. A., D. K. Dawson, and S. Bates. 2004. Modeling abundance effects in distance sampling. Ecology 85:1591–1597.

Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology. Elsevier Academic Press, Oxford, United Kingdom.

Royle, J. A., and M. Kéry. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 88:1813–1823.

Royle, J. A., and W. A. Link. 2006. Generalized site occupancy models allowing for false positive and false negative errors. Ecology 87:835–841.

Royle, J. A., and J. D. Nichols. 2003. Estimating abundance from repeated presence-absence data or point counts. Ecology 84:777–790.

Ruiz-Gutiérrez, V., and E. F. Zipkin. 2011. Detection biases yield misleading patterns of species persistence and colonization in fragmented landscapes. Ecosphere 2:art61.

Sanderlin, J. S., W. M. Block, and J. L. Ganey. 2014. Optimizing study design for multi-species avian monitoring programmes. Journal of Applied Ecology 51:860–870.

Sanderlin, J. S., M. L. Morrison, and W. M. Block. 2019. Analysis of population monitoring data. Pages 131–148 in L. A. Brennan, A. N. Tri, and B. G. Marcot, editors. Quantitative analyses in wildlife science. Johns Hopkins University Press, Baltimore, Maryland, USA.

Sanz-Pérez, A., R. Sollmann, F. Sardà-Palomera, G. Bota, and D. Giralt. 2020. The role of detectability on bird population trend estimates in an open farmland landscape. Biodiversity and Conservation 29:1747–1765.

Sauer, J. R., and W. A. Link. 2011. Analysis of the North American Breeding Bird Survey using hierarchical models. Auk 128: 87–98.

Sauer, J. R., W. A. Link, and J. E. Hines. 2020. The North American Breeding Bird Survey, analysis results 1966 – 2019. U.S. Geological Survey data release, Reston, Virginia, USA. https://doi.org/10.5066/P96A7675

Schmidt, J. H., C. L. McIntyre, and M. C. MacCluskie. 2013. Accounting for incomplete detection: What are we estimating and how might it affect long-term passerine monitoring programs? Biological Conservation 160:130–139.

Schmidt, J. H., W. L. Thompson, T. L. Wilson, and J. H. Reynolds. 2022. Distance sampling surveys: using components of detection and total error to select among approaches. Wildlife Monographs 210:e1070.

Scott, J. M. 2002. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

Si, X., M. W. Cadotte, Y. Zhao, H. Zhou, D. Zeng, J. Li, T. Jin, P. Ren, Y. Wang, P. Ding, and M. W. Tingley. 2018. The importance of accounting for imperfect detection when estimating functional and phylogenetic community structure. Ecology 99:2103–2112.

Sillett, T. S., R. B. Chandler, J. A. Royle, M. Kéry, and S. A. Morrison. 2012. Hierarchical distance-sampling models to estimate population size and habitat-specific abundance of an island endemic. Ecological Applications 22:1997–2006.

Steenweg, R., M. Hebblewhite, J. Whittington, P. Lukacs, and K. McKelvey. 2018. Sampling scales define occupancy and underlying occupancy-abundance relationships in animals. Ecology 99:172–183.

Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278.

Theobald, D. M., D. L. Stevens, D. White, N. S. Urquhart, A. R. Olsen, and J. B. Norman. 2007. Using GIS to generate spatially balanced random survey designs for natural resource applications. Environmental Management 40:134–146.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, San Diego, California, USA.

Tingley, M. W., W. B. Monahan, S. R. Beissinger, and C. Moritz. 2009. Birds track their Grinnellian niche through a century of climate change. Proceedings of the National Academy of Sciences 106:19637–19643.

Tobler, M. W., A. Zúñiga Hartley, S. E. Carrillo-Percastegui, and G. V. N. Powell. 2015. Spatiotemporal hierarchical modelling of species richness and occupancy using camera trap data. Journal of Applied Ecology 52:413–421.

Valente, J. J., R. A. Hutchinson, and M. G. Betts. 2017. Distinguishing distribution dynamics from temporary emigration using dynamic occupancy models. Methods in Ecology and Evolution 8:1707–1716.

Veech, J. A., J. R. Ott, and J. R. Troy. 2016. Intrinsic heterogeneity in detection probability and its effect on N-mixture models. Methods in Ecology and Evolution 7:1019–1028.

Villarreal, M. L., J. M. Iniguez, A. D. Flesch, J. S. Sanderlin, C. Cortés Montaño, C. R. Conrad, and S. L. Haire. 2020. Contemporary fire regimes provide a critical perspective on restoration needs in the Mexico-United States borderlands. Air, Soil and Water Research 13:1178622120969191.

Wenger, S. J., and J. D. Olden. 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. Methods in Ecology and Evolution 3:260–267.

Whelan, C. J., Ç. H. Şekercioğlu, and D. G. Wenny. 2015. Why birds matter: from economic ornithology to ecosystem services. Journal of Ornithology 156:227–238.

Wiens, J. A., and J. T. Rotenberry. 1985. Response of breeding passerine birds to rangeland alteration in a North American shrubsteppe locality. Journal of Applied Ecology 22:655–668.

Williams, B. K., and E. D. Brown. 2019. Sampling and analysis frameworks for inference in ecology. Methods in Ecology and Evolution 10:1832–1842.

Williams, P. J., M. B. Hooten, J. N. Womble, G. G. Esslinger, M. R. Bower, and T. J. Hefley. 2017. An integrated data model to estimate spatiotemporal occupancy, abundance, and colonization dynamics. Ecology 98:328–336.

Wright, A. D., E. H. C. Grant, and E. F. Zipkin. 2020. A hierarchical analysis of habitat area, connectivity, and quality on amphibian diversity across spatial scales. Landscape Ecology 35:529–544.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology & Evolution 16:446–453.

Zhao, Q., A. K. Fuller, and J. A. Royle. 2022. Spatial dynamic N-mixture models with interspecific interactions. Methods in Ecology and Evolution 13:2209–2221.

Zipkin, E. F., J. W. Doser, C. L. Davis, W. Leuenberger, S. Ayebare, and K. L. Davis. 2023. Integrated community models: a framework combining multispecies data sources to estimate the status, trends and dynamics of biodiversity. Journal of Animal Ecology 92:2248–2262.

Zipkin, E. F., E. R. Zylstra, A. D. Wright, S. P. Saunders, A. O. Finley, M. C. Dietze, M. S. Itter, and M. W. Tingley. 2021. Addressing data integration challenges to link ecological processes across scales. Frontiers in Ecology and the Environment 19:30–38.

Zuckerberg, B., A. Desrochers, W. M. Hochachka, D. Fink, W. D. Koenig, and J. L. Dickinson. 2012. Overlapping landscapes: a persistent, but misdirected concern when collecting and analyzing ecological data. Journal of Wildlife Management 76:1072–1080.

Zuur, A. F., E. N. Ieno, N. J. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed effects models and extensions in ecology with R. Springer Science+Business Media, New York, New York, USA.

Associate Editor: Emily Sinnott.

APPENDIX A: IMPLICATIONS OF MOVEMENT FOR STATISTICAL INDEPENDENCE

In general, analytical models require that any deviations of counts from the model-predicted value for the sampling unit within the sampling period can be reasonably approximated with the specified error distribution (e.g., typically Poisson for count data and binomial for binary detection data). To help meet this assumption, researchers often pay special attention to spacing between sampling units and limiting the distance at which surveyors can record detections to avoid counting the same individual at multiple units. The potential for movement between adjacent units, however, should not be confused with statistical independence. Such movement need not violate independence assumptions as long as it is random with respect to the sampling process and ecological parameters of interest (Zuckerberg et al. 2012). Conversely, spacing that avoids detection of the same individual at neighboring units does not necessarily guarantee statistical independence. Non-independence can arise, for example, by birds following or avoiding the surveyor, spatial association or avoidance among neighbors, or unmodeled environmental covariate relationships.