

Detecting diversity: emerging methods to estimate species diversity

Kelly J. Iknayan¹, Morgan W. Tingley², Brett J. Furnas^{1,3}, and Steven R. Beissinger¹

¹Department of Environmental Science, Policy and Management, 130 Mulford Hall, University of California, Berkeley, CA 94720-3110, USA

²Woodrow Wilson School, Princeton University, Princeton, NJ 08544, USA

³California Department of Fish and Wildlife, 601 Locust Street, Redding, CA 96001, USA

Estimates of species richness and diversity are central to community and macroecology and are frequently used in conservation planning. Commonly used diversity metrics account for undetected species primarily by controlling for sampling effort. Yet the probability of detecting an individual can vary among species, observers, survey methods, and sites. We review emerging methods to estimate alpha, beta, gamma, and metacommunity diversity through hierarchical multispecies occupancy models (MSOMs) and multispecies abundance models (MSAMs) that explicitly incorporate observation error in the detection process for species or individuals. We examine advantages, limitations, and assumptions of these detection-based hierarchical models for estimating species diversity. Accounting for imperfect detection using these approaches has influenced conclusions of comparative community studies and creates new opportunities for testing theory.

Diversity and imperfect detection

Diversity estimates are central to community and macroecology [1–4] and are frequently used in conservation planning as a surrogate for biodiversity and to identify areas in need of protection [5,6]. Diversity is classically divided into alpha (site level), beta (turnover across multiple sites), and gamma (composite of all sites in a region) components. The fundamental unit of all diversity metrics is a count of species, individuals, or both. Yet rarely do circumstances occur when all species or all individuals are detected during a survey, regardless of whether the study organisms are birds [7], mammals [8], insects [9], or plants [10,11]. Imperfect detection has predictable consequences: when species are common, missed individuals result in underestimation of populations; when species are rare, missed individuals result in false absences. Uncorrected counts of observed species often used in measures of diversity ignore detection altogether and established methods used to account for missed species do not disentangle detection from occurrence.

Corresponding author: Beissinger, S.R. (beis@berkeley.edu).

0169-5347/\$ – see front matter

© 2013 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tree.2013.10.012>



Hierarchical occupancy models have recently become a standard method to account for false absences when modeling the occurrence of a single species [12]. These models can distinguish between and elucidate two key processes [13]: (i) the ecological or state process of interest (e.g., site occupancy); and (ii) the observation process that always accompanies field sampling (e.g., variability in detection of individuals or species) and that makes identification of the state process imperfect. Recent advances have applied these hierarchical ‘detection-based’ models (i.e., models that incorporate the underlying detection process separately from the ecological measure of interest) to surveys comprising multiple species to estimate diversity metrics. They not only account for multiple sources of detection but can estimate the number of species never encountered during surveys [14].

In this review, we examine the origins of imperfect detection and discuss how hierarchical detection-based

Glossary

Bayesian shrinkage: increased precision and accuracy in a parameter estimate that occurs when parameters are modeled with a common prior distribution, which results in the individual estimates being pulled (‘shrunk’) toward the mean; also called ‘borrowing strength’.

False absence: an individual or species that is present at a site but not detected during sampling due either to observation error or to the individual not being available for detection during the survey period.

Hierarchical model: a statistical model described by a nested sequence of observed and unobserved random variables, which is a particularly flexible and transparent way to model complex dependencies in observed data.

Hill numbers: a set of mathematically unified diversity indices whose outputs are in units of ‘effective number of species’ also known as ‘numbers equivalent’.

Hyperparameter: a parameter that governs the community-level distributions from which species-specific probabilities are drawn.

Latent variable: a parameter that is inferred from other parameters through statistical models and that is not directly observed by sampling.

Markov chain Monte Carlo (MCMC) methods: a class of algorithms used to generate dependent random samples from statistical distributions that may be intractable analytically (typically posterior distributions in Bayesian analysis of a model).

Nonparametric estimator: an estimator that does not assume an underlying distribution for the data on which the estimate is being made.

Rarefaction: a plot of the average number of species encountered as a function of sampling effort, which facilitates comparison of diversity metrics across samples derived from unequal sampling effort.

Species-accumulation curve: a plot that records the total number of species encountered with increasing sampling effort. Effort is usually measured by number of samples or number of individuals.

Uncorrected count: the count of individuals or species from survey data that has not been statistically adjusted to account for sampling bias or sources of imperfect detection.

models can be applied to multiple species to improve estimates of diversity. We highlight important advantages, drawbacks, and considerations of using detection-based models for estimating diversity and compare results from this approach with traditional methods. Multispecies hierarchical approaches are relatively new, so both their potential and their limitations have yet to be realized.

Origins of imperfect detection

Arriving at valid estimates of species diversity from multispecies surveys requires an appreciation of the myriad ways in which species are imperfectly detected. The detection process can be separated into three components: individuals emit a signal (either auditory or visual) that is transmitted through the environment and received by an observer. Various factors affect the rate and strength of transmission, the environmental filtering of the signal's clarity, and the observer's ability to receive it. These factors can be decomposed into four groups (Figure 1) – two that characterize the organism being surveyed (species and individual) and two that relate to sampling logistics (site and survey). Whereas survey- and site-specific factors affecting detectability can be partly addressed through study design, species- and individual-specific differences in detectability can be accounted for only by using statistical methods.

Species vary in their inherent probabilities of detection. Differences in detectability arise from species' behavior (e.g., vocalization rate or volume, movement frequency) and distinctiveness (e.g., call, size, color). The rarity of a species also influences detectability and the probability of detecting a species is positively related to its abundance [15–17]. Less appreciated is that the local abundance of another species (e.g., a competitor or predator) can influence the behavior of the focal species, resulting in either positive or negative changes in detectability [18–20]. Substantial heterogeneity in detection among individuals can also occur within species. For example, in species detected by vocalizations, the frequency and volume of calls can differ by body size, sex, and age. Distance from observer affects the probability of detection of individual animals [21] and can also affect plant detection [22,23].

Site- and survey-specific detectability factors can influence the detectability of all species (Figure 1). The sampling site has intrinsic characteristics that affect detectability relative to comparable samples at other sites. Site-level heterogeneity in detection derives from factors that impede visual or auditory detection regardless of observer, such as habitat structure or noise. Many factors related to the individual survey or sampling event affect detectability. Weather, which generally changes across observation periods, has a strong influence on

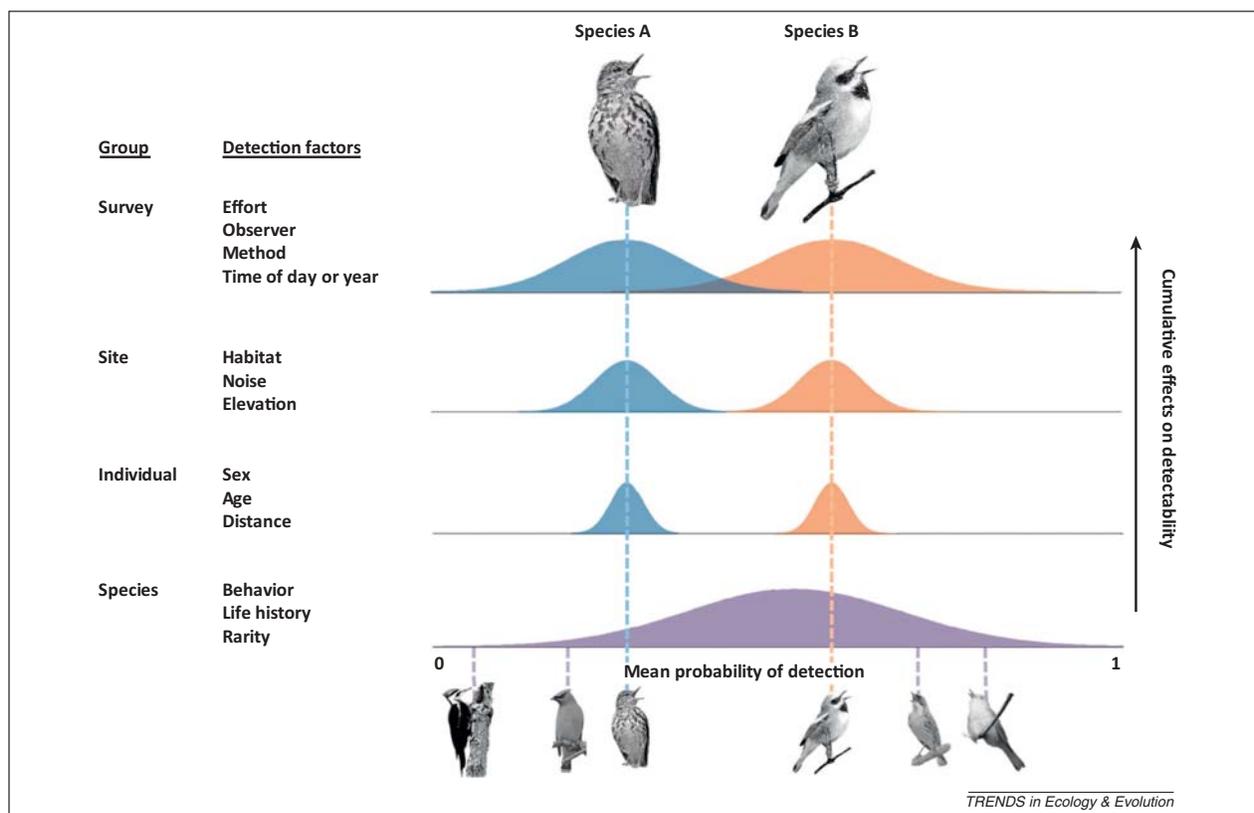


Figure 1. Factors affecting the probability of detection of a species. Mean detectability can be considered a species-level trait that is a function of behavior, life history, and rarity. In a hierarchical model, species-specific detectability is assumed to come from a common community-level distribution, shown here in purple. A hypothetical distribution of observed detection probabilities for two members of a bird community (indicated by dashed lines and color-coded probability curves) is shown for different aspects affecting the detection process (individual-, site-, and survey-specific factors). As more components of the observation process are considered, variation in the probability of detection increases.

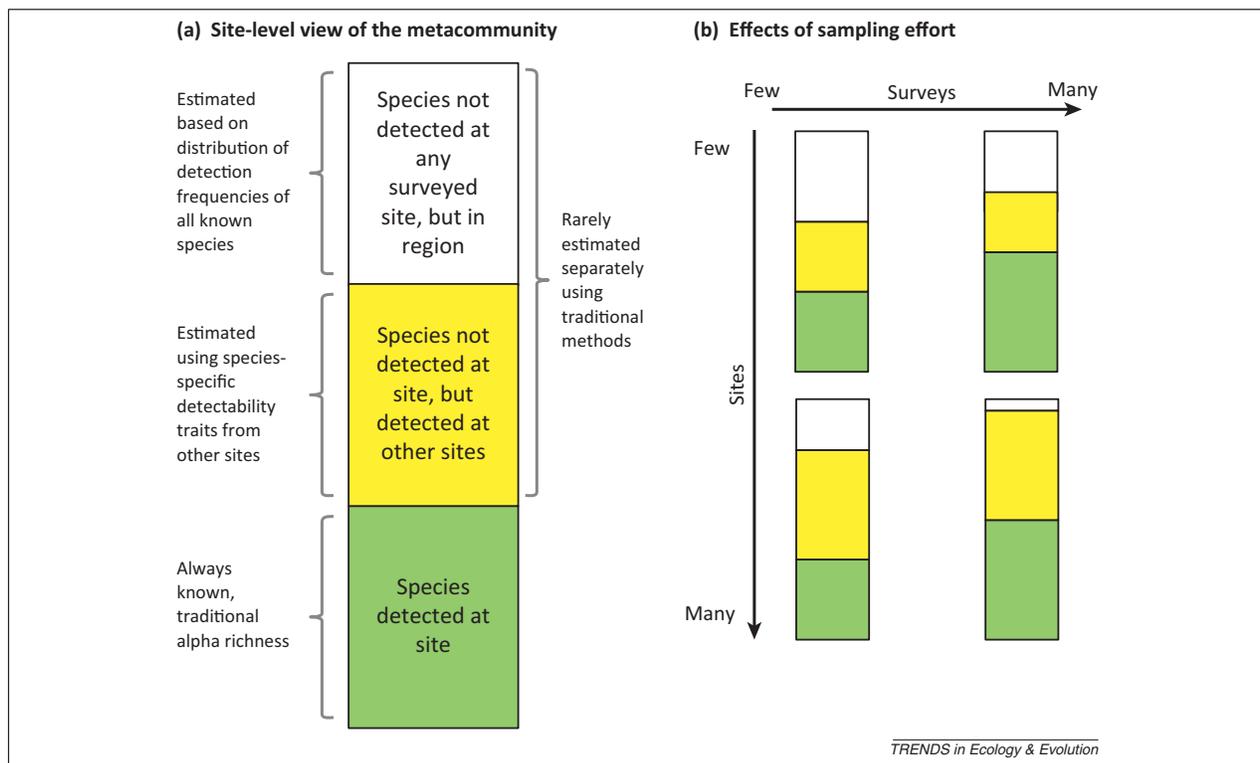


Figure 2. Categories of species at surveyed sites resulting from imperfect detection and how they change with different temporal and spatial sampling strategies. **(A)** The true (unknown) species pool of a metacommunity represented at a site comprises species that have been detected there (green bin), those that have not yet been detected at the site but have been detected at other surveyed sites (yellow bin), and those that have not yet been detected at this or any site but occur in the region (white bin). **(B)** As temporal and spatial replication (i.e., sampling effort) increases, knowledge of the species pool changes for both the site (green bins) and the metacommunity (green + yellow + white bins). When a site is surveyed few times, the relative size of each bin depends on the factors affecting detectability (Figure 1). If there are few sites and few surveys per site, a large portion of the metacommunity may not be detected (white bin of upper left rectangle), either at the site or at other sites. As the number of surveys per site increases (temporal replication) but not the number of sites surveyed (i.e., little spatial replication), the total number of species detected per site increases (green bin in upper right rectangle), mostly as a result of detecting species that are likely to occur at other sites (yellow bin). When the total number of sites surveyed increases (spatial replication) but not the number of surveys (i.e., little temporal replication), the number of species undetected in the region decreases (white bin in lower left rectangle), but the number of species detected per site remains the same (green bin). As both the number of surveys per site and the number of sites surveyed increase, a greater proportion of species in the metacommunity will be known (green + yellow bins in lower right rectangle), either from being directly detected at the site (green bin) or by being detected at other sites (yellow bin).

detectability, affecting both animal behavior and the detection process itself. For example, playback experiments found breezy conditions reduced auditory detections of songbirds by 28% [24]. Surveys are often conducted by different observers, and observer ability, age, and experience affect detectability [25,26]. These factors can also impact whether or not a species is identified correctly, which can result in false positives. Sampling design (e.g., survey methods, timing, effort) also affects detection rates. For example, the detectability of animals often differs by time of day or sampling date (e.g., calling frogs or birds). Although easily controlled, sampling design can differ across the course of a study or among studies.

As a result of imperfect detection, the species that occur at a site at any moment in a multisite survey comprise three categories (Figure 2A): (i) species detected at the site; (ii) species not detected at the site but detected at other sites; and (iii) species not detected at any surveyed site but which are known to, or could, occur in the metacommunity or regional pool of species, which may or may not be well described. When a site is surveyed only once (Figure 2B), the proportion of species undetected may be high. Temporal replication improves the resolution of diversity estimates for

a site. Spatial replication improves the resolution of diversity estimates for the metacommunity due to the patchiness of species' occurrences across a landscape. Because sampling is rarely sufficient to detect all individuals or species, each of the three categories of detected or undetected species should be modeled to estimate the true richness of a site.

Hierarchical detection-based, multispecies models to estimate diversity

Hierarchical detection-based modeling of occupancy or abundance offers an approach for estimating diversity metrics that can incorporate the processes causing imperfect detection (Figure 1). Hierarchical occupancy models are built on 'encounter histories' of species' detection or non-detection at a survey site (Box 1). These histories require repeated surveys during a period when the sampled population is assumed to be closed to changes in occupancy. The pattern of detections at occupied sites yields information on species' detectability that can be applied to sites without detections. 'Single-season' occupancy models provide a framework to estimate hierarchically a probability of detection (p) and a probability of occupancy (ψ) over a sampling period using either a

Box 1. Multispecies models for estimating occupancy and abundance

The MSOM is an extension of the single-species, single-season occupancy model [86] that combines and analyzes the history of detections and non-detections (denoted by 1s and 0s, respectively) of all species encountered during replicated surveys at a set of sites [14]. The hierarchical model includes three levels, one each for species (i), site (j), and replicate (k). The first level represents the true occurrence states (w) within the community of all partially observed and never observed species. The second level is the ecological process governing occurrences (z) of partially observed species. The third level explains the detection history (y) from the replicated surveys:

$$w_i \sim \text{Bernoulli}(\Omega) - \text{superpopulation process (data augmentation);} \quad \text{[I]}$$

$$z_{i,j} | w_i \sim \text{Bernoulli}(w_i * \psi_{i,j}) - \text{ecological process;} \quad \text{[II]}$$

$$y_{i,j,k} | z_{i,j} \sim \text{Bernoulli}(z_{i,j} * p_{i,j,k}) - \text{observation process.} \quad \text{[III]}$$

The second and third levels are always present in a single-species model. The MSOM expands these levels across all observed species within the same system of linked, hierarchical models. It also addresses the community of partially observed and never observed species through addition of the first level. The parameters ψ (proportion of sites occupied) and p (probability of detection) are the same occupancy and detection probability parameters, respectively, used in a single-species model. The parameter Ω governs the data augmentation variable w .

Data augmentation is used in MSOMs to estimate the number of species present in a community (or metacommunity) but not detected at any site. The detection histories of n observed species are augmented by m all-zero detection histories that represent species with unknown identities. The number of all-zero augmented detection histories, m , should be arbitrarily large without being computationally unwieldy (see [55] for guidance). An indicator variable, w , is modeled such that $w = 1$ for species that were either observed (total = n) or unobserved but available for sampling. The total number of unobserved species, x , for which $w = 1$, is equivalent to the asymptote of a species-accumulation curve [36]. Consequently, gamma diversity is represented by $n + x$ (where $0 \leq x \leq m$), which estimates the total number of species in a sample of sites. This can be directly calculated by multiplying the posterior estimate of Ω by the known quantities $n + m$.

The structure of the MSAM is similar to that of the MSOM, except that the detection history comprises observed counts ($y_{i,j,k}$) given the true abundances ($N_{i,j}$):

$$w_i \sim \text{Bernoulli}(\Omega) - \text{superpopulation process (data augmentation);} \quad \text{[IV]}$$

$$N_{i,j} | w_i \sim \text{Poisson}(w_i * \lambda_{i,j}) - \text{ecological process;} \quad \text{[V]}$$

$$y_{i,j,k} | N_{i,j} \sim \text{binomial}(N_{i,j}, p_{i,j,k}) - \text{observation process.} \quad \text{[VI]}$$

Instead of a Poisson distribution controlling the ecological process (Equation V), a zero-inflated Poisson, a negative binomial, or another distribution suitable for counts can be used.

frequentist (maximum likelihood estimation) or Bayesian framework [12,27]. One option is to use a ‘predict first, assemble later’ approach [28] to estimate species richness, where species are modeled individually and richness measures are calculated through aggregation with established metrics [29,30]. Although this approach allows occupancy to be modeled with species-specific covariates, modeling species individually restricts inferences about diversity to species that have been detected multiple times, which excludes species that are rare or are undetected across

all sites (Figure 2). Additionally, this approach makes it difficult to propagate uncertainty in occurrence estimates of a single species into uncertainty for community metrics.

As a result of the limitations of aggregating single-species occupancy models, estimation of diversity metrics in a detection-based modeling framework typically take an ‘assemble and predict together’ approach using a hierarchical community occupancy model. We will refer to these as MSOMs when used to estimate incidence-based measures of diversity [14,31,32]. Hierarchical diversity models have been developed without explicitly accounting for the detection process [33,34]. However, a great advantage of incorporating the detection process into community models is the ability to account explicitly for the effects of survey-, site-, species-, and individual-level factors affecting detectability (Figure 1) through the inclusion of one or multiple detection covariates [14,31]. MSOMs model undetected species in a biologically oriented, process-driven way. The mathematical framework used by MSOMs is discussed in Box 1. Although MSOMs can be analyzed using either frequentist or Bayesian methods, current implementations, including available code [13,35], favor the latter.

The hierarchical structure of MSOMs allows data from the entire sample to inform the estimation of diversity, despite encounter histories being stratified by species and site [14]. The detection and occupancy probabilities of each species are assumed to come from a community-level distribution (e.g., Figure 1, bottom panel). For the sake of mathematical convenience, a normal distribution on the logit scale is often used, with an associated mean and variance as hyperparameters. This construction is equivalent to treating the different species as random effects [13,27,36]. Community-level hyperparameters facilitate the modeling of all species, including rare ones, through a property often referred to as ‘borrowing strength’ [37] or Bayesian shrinkage [35,38]. Each species’ estimate informs the overall (community level) mean and variance and, as a result, estimates of individual species are pulled (i.e., shrunk) toward the community mean. Many of the benefits of MSOMs arise from this ‘information sharing’, because data are used more efficiently compared with single-species models and individual species estimates are improved [27,38]. Hierarchical community models have improved the precision of diversity descriptors [35], which can offset the costs of multispecies monitoring efforts [39]. Moreover, borrowing strength allows MSOMs to estimate site-level occupancy for rarely detected species, which is often not possible with single-species models due to the limited number of detections [40,41]. Although Bayesian shrinkage can improve parameter estimation, estimates for infrequently encountered species will be pulled more toward the metacommunity mean because they are estimated with less precision and inform the community-level mean less than species that are frequently detected [27,42].

Although most hierarchical detection-based occupancy models can be used to produce estimates of species richness and turnover based on incidence, the MSOM framework easily accommodates replacement of species incidence with counts of individuals to create a MSAM to estimate abundance-based diversity metrics (e.g., [43–45]). The MSAM is a

multispecies version of the N -mixture model that has been used to estimate the abundance of individual species [46–50]. It is structurally similar to the MSOM (Box 1), but replaces the ecological and observation process in the MSOM with distributions suitable for counts. Community-level hyperparameters of the MSAM represent: (i) the abundance of a species in the metacommunity; and (ii) the probability of detecting an individual for each species.

Beyond the need for repeated surveys, MSAMs lack the ancillary data requirements characteristic of distance sampling and capture–recapture methods for estimating abundance [51]. They do, however, require assumptions about the probability distribution governing the true number of individuals of each species occurring at each site. Moreover, for each species at a site, the detection of each individual is assumed to be independent of that of other individuals. This can lead to overinflation of counts and contrasts with the approach taken in distance sampling [21], which models the detectability of groups of individuals. However, a beta-binomial distribution can be applied to address nonindependence of detections of individuals [52,53].

Richness and diversity estimation in the multispecies detection-based modeling framework

MSOMs have been applied to diversity estimation in two situations: (i) when all possible species in the metacommunity or region were detected at least once during sampling, which can be modeled with the MSOM structure discussed above [13]; and (ii) where metacommunity or regional species richness exceeds the total number of species detected across all sites [31]. To model richness when the pool of potential species is unknown (a common occurrence), MSOMs employ ‘data augmentation’ [36,54] to estimate the number of species never detected during sampling in addition to those that were detected at least once (Figure 2). As discussed in Box 1 and illustrated in Box 2, this is done by augmenting the detection histories of

observed species with an arbitrarily large number of all-zero records that represent the detection histories of hypothetical, unobserved species. The model then estimates how many hypothetical species are likely to occur in the sampling region but were missed by all surveys, in addition to the number of species that were detected at least once (see [55] for additional details on data augmentation). The total number of species estimated by the data-augmentation process is comparable to the asymptote of the species-accumulation curve in a homogeneous landscape [36,54], but allows for the incorporation of detection and occupancy covariates into the estimation of the asymptote. Thus, data augmentation provides one way to estimate diversity at the scale of the metacommunity [32] or region (gamma diversity).

Any occurrence-based descriptor of diversity and its associated measures of uncertainty can be derived directly within the MSOM framework for any subset of samples, such as sites, groupings of species, metacommunities, or regions (e.g., [7,8,42]). The multispecies approach retains the identity of encountered species throughout the modeling process and allows estimation of species-specific detection and occupancy probabilities, which can be derived from general or group-specific (e.g., taxonomic or functional) covariates [42,56]. Currently MSOMs must be written by the user and run within a programming environment or precompiled Markov chain Monte Carlo (MCMC) programs such as WinBUGS [57], Jags [58], or OpenBUGS [59], all of which use the Bayesian inference Using Gibbs Sampling (BUGS) language. Statistical programs such as R (<http://www.R-project.org>) can often interface with MCMC programs. Code for various models is available in the BUGS language (e.g., [7,13,36,41,54]).

There are a limited but growing number of studies that have applied MSOMs to estimate diversity. Birds have been the primary target [7,35,54,60], but applications include amphibians [61,62], freshwater and reef fish [56,63,64], and invertebrates [31,32,65]. MSOM estimation

Box 2. An example illustrating multispecies modeling

Here we provide a brief overview of the modeling process used to estimate detection-based community metrics derived from a MSOM developed by Zipkin *et al.* [54] (Figure 1). Data were collected to determine the effects of two different management treatments on breeding-bird diversity. A single-season MSOM was used to account for species-specific differences in detection and occupancy by site. The model runs using the freely available software R (<http://www.R-project.org>) and WinBUGS [57]. Code and data are available online (<https://sites.google.com/site/communitymodeling/software-code>).

Inputs

Temporally and spatially replicated surveys are required to calculate detection-based community metrics. The number of surveys need not be the same for all sites. One or more explanatory covariates can be incorporated at the species, site, or survey level (e.g., Julian day is a survey-level covariate of detection). Multispecies survey data are organized into a 3D array: survey (k) \times site (j) \times species (i). This array includes the encounter histories for all observed species. An arbitrarily large number of all-zero encounter histories are appended to this array in a process known as data augmentation (Box 1). Continuous covariates are typically standardized, which helps with overall MCMC performance and simplifies the comparison of covariate effects when interpreting model output.

Model specification and MCMC sampling

Occupancy (ψ) and detectability (p) parameters can be generalized to vary according to explanatory covariates. This is usually handled as a linear model on the logit scale, where regression coefficients are modeled as species-specific random effects derived from a community-level distribution. Before initiating MCMC sampling, prior distributions are specified for the hyperparameters of the community-level distributions, for the hyperdistributions themselves, and for Ω , the parameter that determines membership in the metacommunity from the superpopulation (Box 1). In practice, non-informative priors are often used [36].

Model output of posterior estimates and summary analyses

The model returns posterior estimates for species-specific occupancy and detection probabilities and the species-specific effects of covariates. Posterior distributions for any derived community-level metrics can also be returned. In an MSOM, these metrics can include richness of the metacommunity, richness of individual sites, or richness of sets of sites, as well as richness of different functional groups at those spatial levels. Similarity indices of beta diversity can be also be derived [7]. These estimates can be used in summary analyses to investigate relationships with covariates or to compare metacommunities.

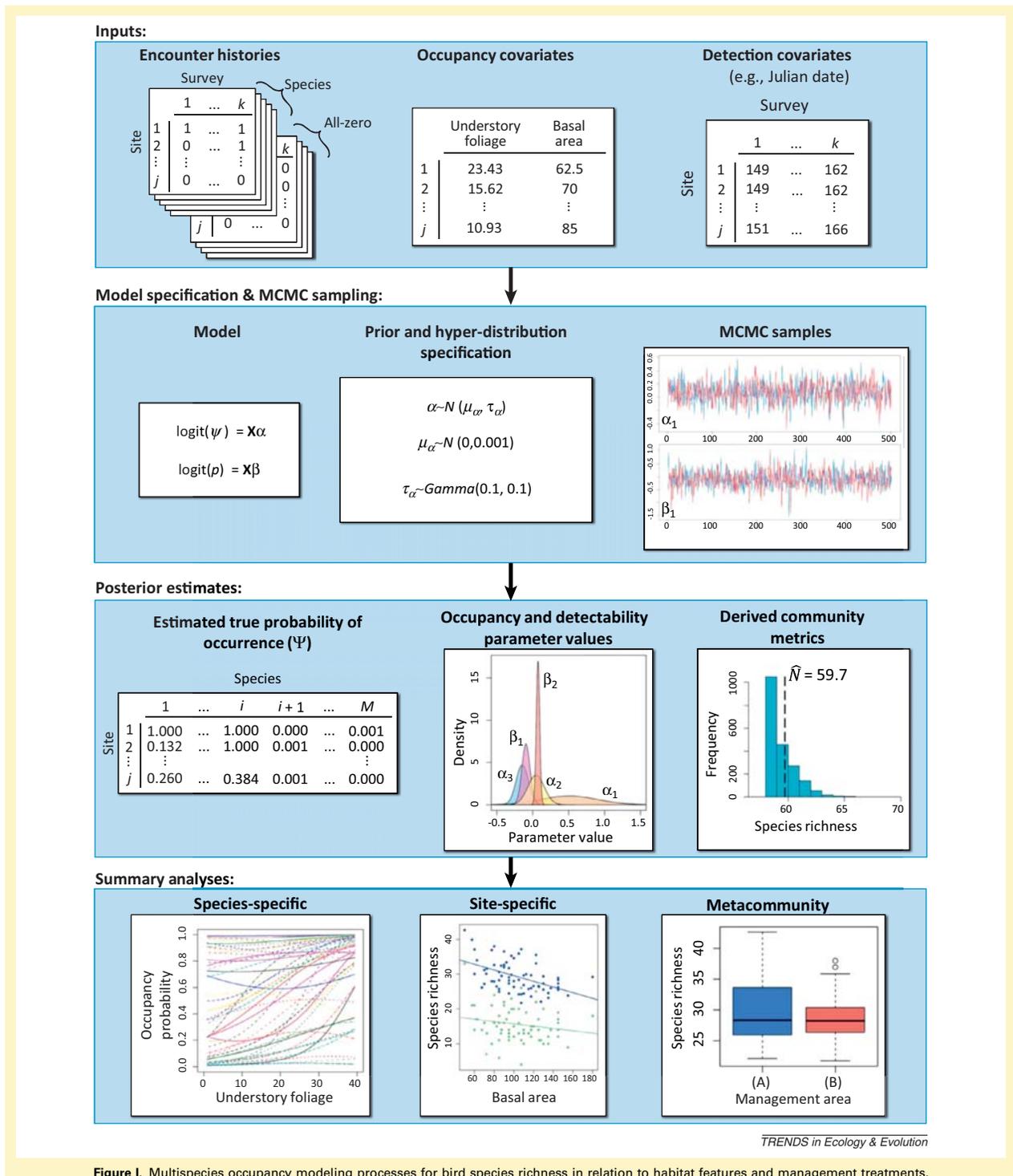


Figure I. Multispecies occupancy modeling processes for bird species richness in relation to habitat features and management treatments.

of beta-diversity indices has been done only with the Jaccard and Sørensen indices [7,41,54,63], but the approach could be applied to any of the large number of beta-diversity indices used with occurrence-based data (e.g., complementarity).

Abundance-based estimates of diversity from MSAMs are in their infancy. Only two studies have extended the

N-mixture model to a multispecies framework to evaluate community responses of birds and bees to land-use change, using guild-level hyperparameters and data augmentation to differentiate estimated abundances and species richness between early successional and mature forest species. Chandler *et al.* [67] modeled abundance-based estimates

of the Chao–Jaccard similarity index and Shannon diversity to compare the conservation values of competing agricultural systems in a tropical forest landscape. Although diversity estimation with MSAMs has been very limited to date, these models could be used to estimate directly any abundance-based diversity metric, including Hill numbers [68], by including them as derived quantities in the model (Box 2). An advantage of this approach is that posterior distributions calculated by the MCMC algorithm can be used to describe straightforwardly the uncertainty associated with the estimates of these diversity metrics.

Comparison of detection-based estimators of diversity with traditional estimators

Compared with detection-based estimators of diversity, traditional approaches for estimating diversity have limited ability to incorporate or accommodate the survey-, site-, and species-level processes that differentially affect the detection of species or individuals (Figure 1). The most popular species-richness estimators, the Jackknife [69] and Chao estimators [70,71], correct for bias due to sampling effort and species' rarity by extrapolating the proportion of species encountered infrequently to species encountered more frequently to estimate the number of undetected species [31,72,73]. Although this reduces bias compared with uncorrected species counts, mathematical methods for incorporating additional features of the study design, such as site- or survey-specific covariates affecting species' detection (Figure 1), are not straightforward or often not possible [36]. Beta-diversity indices of community similarity are frequently calculated without correcting for imperfect detection [73] and their performance in the face of undetected species is variable [45,74]. The Chao–Jaccard/Sorensen index [45,73] is one of the few beta-diversity estimators that corrects for sampling bias by estimating the contribution of undetected species using the probability that individuals drawn randomly from the sample belong to a species shared by the two assemblages, based on species' relative abundances or occurrences.

Critically, this index assumes that the detection rate of species is driven primarily by relative abundance and that detectability does not otherwise differ among observers, species, or sites [73], factors that often affect detection (Figure 1) and that can be accommodated in MSOMs and MSAMs. Alpha- and gamma-diversity indices, such as the Shannon–Wiener and Simpson indices and Hill numbers, which combine richness and relative abundance [75,76], face estimation issues similar to beta diversity. Chao and Shen [77] and Chao *et al.* [68] developed indices and estimators to correct for sampling bias in the Shannon–Wiener index and Hill numbers, respectively. They improve diversity estimation, but appear unable to correct for most site- and species-level causes of imperfect detection. None of the traditional estimators of diversity is able to differentiate whether the estimated number of undetected species represents: (i) species that were detected at other sites; or (ii) species that were not detected in the sample as a whole (Figure 2).

Direct comparisons of traditional- and detection-based estimates of species diversity are limited. Two studies compared the second-order Jackknife estimator to a simple MSOM that was built without site- or survey-level covariates [40,63]. Both models assumed that species identity was the only factor contributing to variation in detectability [40]. MSOM richness estimates were more precise than the second-order Jackknife estimator and generated comparable estimates across surveys with dual observers [40]. Jackknife estimates of richness at the site-level were sometimes erratic, generating reasonable estimates for some sites but unreasonably high, inaccurate estimates for others [40,63] (Figure 3A). These differences might reflect more efficient use of data by MSOMs compared with the Jackknife technique. However, comparisons of the methods have not been extensive and this is an area with ample opportunity for further evaluation. Diversity estimates from MSOMs and MSAMs could be compared with estimates from traditional methods using the same simulated and 'exhaustively sampled' real-world datasets.

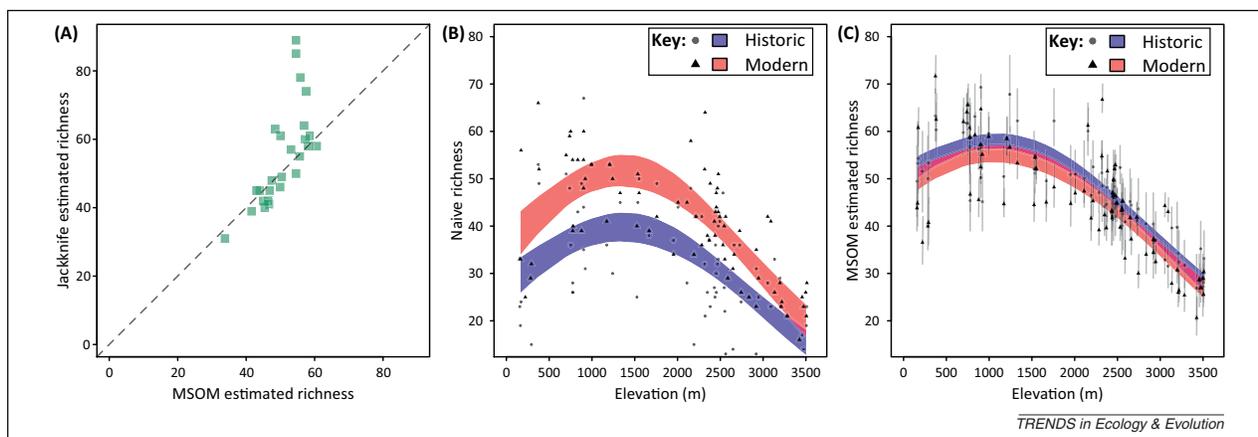


Figure 3. Examples of improved inferences for species richness estimation from multispecies occupancy models (MSOMs). (A) Comparison of bird species richness estimated from MSOMs versus the traditional Jackknife estimator. Estimators match well when richness is low, but Jackknife estimates diverged at higher richness values and are likely to represent an overestimation. Data taken from Table 1 in [36]. (B,C) Conclusions about change in species richness over the past century for birds of the Sierra Nevada are reversed when raw species counts are corrected for detectability using a MSOM. Uncorrected counts (B) suggested species richness increased between surveys conducted by Joseph Grinnell and colleagues from 1911 to 1929 and resurveys conducted by Tingley and Beissinger from 2003 to 2009, whereas MSOM estimates (C) accounting for detectability differences showed that richness declined. (B) and (C) adapted from Figure 1 in [7].

To date, no study has compared beta-diversity indices of any kind with estimates corrected for heterogeneous detectability derived from a MSOM or MSAM.

We do know, however, that richness estimated with MSOMs and MSAMs is always greater than when estimated from raw species counts [31], because MSOMs supplement raw species counts with the expected number of undetected species [36]. Using raw counts to measure species diversity can mask temporal trends that are detected with MSOM estimation [7] (Figure 3B,C) or alter the explanatory power of ecological covariates [54].

Concluding remarks: the future of diversity metrics

Producing accurate estimates of species richness and diversity requires accounting for factors that affect imperfect detection (Figure 1) and recognizing the categories of detected and undetected species at surveyed sites (Figure 2). Hierarchical multispecies models that incorporate both the detection process and the occurrence state provide a promising way forward, because they lead to a more process-driven estimate of diversity through the delineation of the biological and sampling processes. However, detection-based estimators of diversity generally require multiple surveys at a site to estimate detectability, make assumptions about the closure of populations and about the veracity of species identification [78], and are still relatively new so their limitations and performance need further evaluation (Box 3). Moreover, these hierarchical models are relatively complex and beyond the experience of many potential users. Although the software that runs them is freely available and example code for commonly used models is found on the Internet, more user-friendly implementations of MSOMs and MSAMs would facilitate their use and acceptance. A package that fully implements these models within the R environment would be a strong first step toward this goal. Above and beyond this, the development of specialized software to run these models would increase the ease of implementation and encourage widespread use, much like EstimateS (<http://purl.oclc.org/estimates>) did for traditional measures of species diversity or Presence (<http://www.mbr-pwrc.usgs.gov/software/presence.html>) did for single-species occupancy models.

Although the use of detection-based estimators of diversity is increasing, further innovations that address various issues already incorporated into single-species occupancy models are needed. As discussed in Box 3, this includes correcting for false positives [78,79] and evaluating model performance and issues relating to closure. Regarding the latter, Sólymos *et al.* [80] recently proposed a 'conditional-likelihood model' of abundance estimation that avoids the closure assumption and need for multiple surveys at a site by adjusting detection error through covariates that affect detection and occupancy. It performed well against simulated data sets, but remains to be applied in a multispecies framework. Another important advance was the development of a dynamic, multi-season MSOM that models change in occupancy between sampling periods as local colonization and extinction parameters, which provides a method to calculate beta diversity and evaluate metacommunity dynamics [32].

Box 3. Assumptions and limitations of MSOMs

The main assumptions of occupancy models and MSOMs are: (i) sites are closed to extinction and colonization over replicated surveys ('closure'); (ii) species are correctly identified; and (iii) the probability of detection and occupancy at a site are independent of detection and occupancy at another site (usually assured by a minimum distance between sites).

Closure might be the most commonly violated assumption of occupancy models because species might be unavailable for detection due to daily or seasonal movements. When closure is violated, modeled detectability is biased low and occupancy is overestimated [87]. This, in turn, has the potential to inflate richness estimates, which might not be problematic if inflation occurs equally across all compared units. If the closure violation varies among species, however, diversity metrics could be differentially biased by species or site. Kéry *et al.* [65] developed a MSOM that allows for temporary emigration between sampling periods, which could reduce this bias. The potential consequences of closure violation on species-diversity metrics deserve further exploration.

With their focus on false negatives, most occupancy models do not model false positives (but see [78,79,88]). Surveys that rely on visual and aural cues are particularly susceptible to misclassification [24,26] and small errors in classification can lead to large biases in occupancy [78,79]. Distance from observer, observer error, and simultaneous vocalizations of multiple species are the leading causes of misclassification. Observer experience and training instructions did not reduce the prevalence of errors [79]. Models have been proposed to account for both false negatives and false positives [78], but issues with parameter identifiability arise when they are applied to data with heterogeneous detection [89]. A more recent parameterization of the original model, which can be applied to data with two or more detection methods, might solve some of the identifiability issues [79].

The hierarchical modeling approach of MSOMs, particularly within a Bayesian framework, offers its own challenges. There are few options for choosing among competing models that contain different parameterizations. Information-theoretic methods [90] and model-selection indices, such as the deviance information criterion (DIC) [91], cannot be reliably calculated for hierarchical models [92,93]. Although no single model-selection approach for hierarchical models has yet seen wide application [13], it is an area of ongoing development and advancement (e.g., [94]). Model selection might present challenges, but variable selection can be done through various methods [95].

When applied to European butterflies, it illustrated the potential of MSOMs to evaluate competing theories of metacommunity ecology [81]. Similarly, development of MSAMs offers more than just the ability to improve estimation of abundance-based measures of diversity. They create new opportunities for evaluating the shapes and functional forms of species-abundance distributions [82,83] and offer applications to niche ecology [84,85] and neutral theories [1,2].

Acknowledgments

This work was supported by the Berkeley Initiative in Global Change Biology, NSF grants CNH 1115069 and DEB-1051342 to S.R.B. and a David H. Smith Postdoctoral Fellowship to M.W.T. Reviews by Marc Kéry, Paul Craze, Perry de Valpine, the Beissinger laboratory, and three anonymous reviewers greatly improved this manuscript.

References

- 1 Hubbell, S.P. (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press
- 2 Harte, J. (2011) *Maximum Entropy and Ecology: A Theory of Abundance, Distribution, and Energetics*, Oxford University Press
- 3 Leibold, M.A. *et al.* (2004) The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.* 7, 601–613

- 4 Brown, J.H. (1995) *Macroecology*, University of Chicago Press
- 5 Margules, C.R. and Pressey, R.L. (2000) Systematic conservation planning. *Nature* 405, 243–253
- 6 Moilanen, A. *et al.* (2009) *Spatial Conservation Prioritization: Quantitative Methods and Computational Tools*, Oxford University Press
- 7 Tingley, M.W. and Beissinger, S.R. (2013) Cryptic loss of montane avian richness and high community turnover over 100 years. *Ecology* 94, 598–609
- 8 Burton, A.C. *et al.* (2012) Hierarchical multi-species modeling of carnivore responses to hunting, habitat and prey in a West African protected area. *PLoS ONE* 7, e38007
- 9 Bried, J.T. *et al.* (2012) Bias of reduced-effort community surveys for adult Odonata of lentic waters. *Insect Conserv. Divers.* 5, 213–222
- 10 Shefferson, R.P. *et al.* (2003) Life history trade-offs in a rare orchid: the costs of flowering, dormancy, and sprouting. *Ecology* 84, 1199–1206
- 11 Chen, G. *et al.* (2013) Imperfect detection is the rule rather than the exception in plant distribution studies. *J. Ecol.* 101, 183–191
- 12 MacKenzie, D.I. *et al.* (2006) *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*, Academic Press
- 13 Royle, J.A. and Dorazio, R.M. (2008) *Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations and Communities*, Academic Press
- 14 Dorazio, R.M. and Royle, J.A. (2005) Estimating size and composition of biological communities by modeling the occurrence of species. *J. Am. Stat. Assoc.* 100, 389–398
- 15 He, F. and Gaston, K.J. (2003) Occupancy, spatial variance, and the abundance of species. *Am. Nat.* 162, 366–375
- 16 Royle, J.A. and Nichols, J.D. (2003) Estimating abundance from repeated presence–absence data or point counts. *Ecology* 84, 777–790
- 17 Dorazio, R.M. (2007) On the choice of statistical models for estimating occurrence and extinction from animal surveys. *Ecology* 88, 2773–2782
- 18 Bailey, L.L. *et al.* (2009) Modeling co-occurrence of northern spotted and barred owls: accounting for detection probability differences. *Biol. Conserv.* 142, 2983–2989
- 19 Richmond, O.M.W. *et al.* (2010) Two-species occupancy models: a new parameterization applied to co-occurrence of secretive rails. *Ecol. Appl.* 20, 2036–2046
- 20 Waddle, J.H. *et al.* (2010) A new parameterization for estimating co-occurrence of interacting species. *Ecol. Appl.* 20, 1467–1475
- 21 Buckland, S.T. *et al.* (2001) *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*, Oxford University Press
- 22 Buckland, S.T. *et al.* (2007) Line transect methods for plant surveys. *Biometrics* 63, 989–998
- 23 Jensen, A. and Meilby, H. (2012) Assessing the population status of a tree species using distance sampling: *Aquilaria crassna* (Thymelaeaceae) in Northern Laos. *Int. J. For. Res.* 2012, 265831
- 24 Simons, T.R. *et al.* (2007) Experimental analysis of the auditory detection process on avian point counts. *Auk* 124, 986–999
- 25 Alldredge, M.W. *et al.* (2007) Factors affecting aural detections of songbirds. *Ecol. Appl.* 17, 948–955
- 26 McClintock, B.T. *et al.* (2010) Experimental investigation of observation error in anuran call surveys. *J. Wildl. Manag.* 74, 1882–1893
- 27 Kéry, M. and Schaub, M. (2012) *Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective*, Academic Press
- 28 Ferrier, S. and Guisan, A. (2006) Spatial modelling of biodiversity at the community level. *J. Appl. Ecol.* 43, 393–404
- 29 Ahumada, J.A. *et al.* (2011) Community structure and diversity of tropical forest mammals: data from a global camera trap network. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 366, 2703–2711
- 30 Williams, M.R. (2009) Butterflies and day-flying moths in a fragmented urban landscape, south-west Western Australia: patterns of species richness. *Pac. Conserv. Biol.* 15, 32–46
- 31 Dorazio, R.M. *et al.* (2006) Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology* 87, 842–854
- 32 Dorazio, R.M. *et al.* (2010) Models for inference in dynamic metacommunity systems. *Ecology* 91, 2466–2475
- 33 Ovaskainen, O. and Soininen, J. (2010) Making more out of sparse data: hierarchical modeling of species communities. *Ecology* 92, 289–295
- 34 Gelfand, A.E. *et al.* (2005) Modelling species diversity through species level hierarchical modelling. *Appl. Statist.* 54, 1–20
- 35 Zipkin, E.F. *et al.* (2009) Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling. *J. Appl. Ecol.* 46, 815–822
- 36 Kéry, M. and Royle, J.A. (2009) Inference about species richness and community structure using species-specific occupancy models in the national Swiss breeding bird survey MHB. In *Modeling Demographic Processes in Marked Populations* (Thomson, D.L. *et al.*, eds), pp. 639–656, Springer
- 37 Clark, J.S. *et al.* (2005) Hierarchical Bayes for structured, variable populations: from recapture data to life-history prediction. *Ecology* 86, 2232–2244
- 38 Link, W.A. and Sauer, J.R. (1996) Extremes in ecology: avoiding the misleading effects of sampling variation in summary analyses. *Ecology* 77, 1633–1640
- 39 DeWan, A. and Zipkin, E. (2010) An integrated sampling and analysis approach for improved biodiversity monitoring. *Environ. Manag.* 45, 1223–1230
- 40 Kéry, M. and Royle, J.A. (2008) Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys. *J. Appl. Ecol.* 45, 589–598
- 41 Dorazio, R.M. *et al.* (2011) Modern methods of estimating biodiversity from presence–absence surveys. In *Biodiversity Loss in a Changing Planet* (Grillo, O. and Venora, G., eds), pp. 277–302, InTech
- 42 Ruiz-Gutiérrez, V. *et al.* (2010) Occupancy dynamics in a tropical bird community: unexpectedly high forest use by birds classified as non-forest species. *J. Appl. Ecol.* 47, 621–630
- 43 Shannon, C. (1948) A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423
- 44 Simpson, E.H. (1949) Measurement of diversity. *Nature* 163, 688
- 45 Chao, A. *et al.* (2006) Abundance-based similarity indices and their estimation when there are unseen species in samples. *Biometrics* 62, 361–371
- 46 Royle, J.A. (2004) *N*-mixture models for estimating population size from spatially replicated counts. *Biometrics* 60, 108–115
- 47 Dodd, C.K. and Dorazio, R.M. (2004) Using counts to simultaneously estimate abundance and detection probabilities in a salamander community. *Herpetologica* 60, 468–478
- 48 Kéry, M. *et al.* (2005) Modeling avian abundance from replicated counts using binomial mixture models. *Ecol. Appl.* 15, 1450–1461
- 49 Wenger, S.J. and Freeman, M.C. (2008) Estimating species occurrence, abundance, and detection probability using zero-inflated distributions. *Ecology* 89, 2953–2959
- 50 Joseph, L.N. *et al.* (2009) Modeling abundance using *N*-mixture models: the importance of considering ecological mechanisms. *Ecol. Appl.* 19, 631–642
- 51 Kéry, M. (2008) Estimating abundance from bird counts: binomial mixture models uncover complex covariate relationships. *Auk* 125, 336–345
- 52 Martin, J. *et al.* (2011) Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. *Methods Ecol. Evol.* 2, 595–601
- 53 Dorazio, R.M. *et al.* (2013) Estimating abundance while accounting for rarity, correlated behavior, and other sources of variation in counts. *Ecology* 94, 1472–1478
- 54 Zipkin, E.F. *et al.* (2010) Multi-species occurrence models to evaluate the effects of conservation and management actions. *Biol. Conserv.* 143, 479–484
- 55 Royle, J.A. and Dorazio, R.M. (2012) Parameter-expanded data augmentation for Bayesian analysis of capture–recapture models. *J. Ornithol.* 152, 521–537
- 56 Cheal, A.J. *et al.* (2012) Spatial variation in the functional characteristics of herbivorous fish communities and the resilience of coral reefs. *Ecol. Appl.* 23, 174–188
- 57 Lunn, D.J. *et al.* (2000) WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat. Comput.* 10, 325–337
- 58 Plummer, M. (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing* (Hornik, K. *et al.*, eds), pp. 20–22 TU Wien
- 59 Lunn, D. *et al.* (2009) The BUGS project: evolution, critique and future directions. *Stat. Med.* 28, 3049–3067

- 60 Russell, R.E. *et al.* (2009) Modeling the effects of environmental disturbance on wildlife communities: avian responses to prescribed fire. *Ecol. Appl.* 19, 1253–1263
- 61 Walls, S.C. *et al.* (2011) Estimating occupancy dynamics in an anuran assemblage from Louisiana, USA. *J. Wildl. Manag.* 75, 751–776
- 62 Zipkin, E.F. *et al.* (2012) Evaluating the predictive abilities of community occupancy models using AUC while accounting for imperfect detection. *Ecol. Appl.* 22, 1962–1972
- 63 Holtrop, A.M. *et al.* (2010) Estimating sampling effort required for characterizing species richness and site-to-site similarity in fish assemblage surveys of Wadeable Illinois streams. *Trans. Am. Fish. Soc.* 139, 1421–1435
- 64 MacNeil, M.A. *et al.* (2008) Accounting for detectability in reef-fish biodiversity estimates. *Mar. Ecol. Prog. Ser.* 367, 249–260
- 65 Kéry, M. *et al.* (2009) Species richness and occupancy estimation in communities subject to temporary emigration. *Ecology* 90, 1279–1290
- 66 Yamaura, Y. *et al.* (2012) Biodiversity of man-made open habitats in an underused country: a class of multispecies abundance models for count data. *Biodivers. Conserv.* 21, 1365–1380
- 67 Chandler, R.B. *et al.* (2013) Small-scale land-sparing approach to conserving biological diversity in tropical agricultural landscapes. *Conserv. Biol.* 27, 785–795
- 68 Chao, A. *et al.* (2013) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol. Monogr.* <http://dx.doi.org/10.1890/13-0133.1>
- 69 Burnham, K.P. and Overton, W.S. (1979) Robust estimation of population size when capture probabilities vary among animals. *Ecology* 60, 927–936
- 70 Chao, A. (1984) Nonparametric estimation of the number of classes in a population. *Scand. J. Stat.* 11, 265–270
- 71 Chao, A. (1987) Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* 43, 783–791
- 72 Mao, C.X. and Colwell, R.K. (2005) Estimation of species richness: mixture models, the role of rare species, and inferential challenges. *Ecology* 86, 1143–1153
- 73 Chao, A. *et al.* (2005) A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol. Lett.* 8, 148–159
- 74 Beck, J. *et al.* (2013) Undersampling and the measurement of beta diversity. *Methods Ecol. Evol.* 4, 370–382
- 75 Magurran, A.E. (2004) *Measuring Biological Diversity*, Blackwell
- 76 Ellison, A.M. (2010) Partitioning diversity. *Ecology* 91, 1962–1963
- 77 Chao, A. and Shen, T.J. (2003) Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environ. Ecol. Stat.* 10, 429–443
- 78 Royle, J.A. and Link, W.A. (2006) Generalized site occupancy models allowing for false positive and false negative errors. *Ecology* 87, 835–841
- 79 Miller, D.A. *et al.* (2011) Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification. *Ecology* 92, 1422–1428
- 80 Sólymos, P. *et al.* (2012) Conditional likelihood approach for analyzing single visit abundance survey data in the presence of zero inflation and detection error. *Environmetrics* 23, 197–205
- 81 Holyoak, M. *et al.* (2005) *Metacommunities: Spatial Dynamics and Ecological Communities*, University of Chicago Press
- 82 Preston, F.W. (1948) The commonness, and rarity, of species. *Ecology* 29, 254–283
- 83 Wilson, J.B. (1991) Methods for fitting dominance/diversity curves. *J. Veg. Sci.* 2, 35–46
- 84 King, C.E. (1964) Relative abundance of species and MacArthur's model. *Ecology* 45, 716–727
- 85 Whittaker, R.H. (1965) Dominance and diversity in land plant communities: numerical relations of species express the importance of competition in community function and evolution. *Science* 147, 250–260
- 86 MacKenzie, D.I. *et al.* (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83, 2248–2255
- 87 Rota, C.T. *et al.* (2009) Occupancy estimation and the closure assumption. *J. Appl. Ecol.* 46, 1173–1181
- 88 Aing, C. *et al.* (2011) A Bayesian hierarchical occupancy model for track surveys conducted in a series of linear, spatially correlated, sites. *J. Appl. Ecol.* 48, 1508–1517
- 89 Fitzpatrick, M.C. *et al.* (2009) Observer bias and the detection of low-density populations. *Ecol. Appl.* 19, 1673–1679
- 90 Burnham, K.P. and Anderson, D.R. (2002) *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*, Springer Verlag
- 91 Spiegelhalter, D.J. *et al.* (2002) Bayesian measures of model complexity and fit. *J. R. Statist. Soc. B* 64, 583–639
- 92 Celeux, G. *et al.* (2006) Deviance information criteria for missing data models. *Bayesian Anal.* 1, 651–673
- 93 Millar, R.B. (2009) Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes' Factors. *Biometrics* 65, 962–969
- 94 Mattsson, B.J. *et al.* (2013) Explaining local-scale species distributions: relative contributions of spatial autocorrelation and landscape heterogeneity for an avian assemblage. *PLoS ONE* 8, e55097
- 95 O'Hara, R.B. and Sillanpää, M.J. (2009) A review of Bayesian variable selection methods: what, how and which. *Bayesian Anal.* 4, 85–117