

Estimating abundance of unmarked animal populations: accounting for imperfect detection and other sources of zero inflation

Francisco V. Dénes^{1,2,3,4*}, Luís Fábio Silveira² and Steven R. Beissinger³

¹Pós-graduação, Departamento de Zoologia, Instituto de Biociências, Universidade de São Paulo, Rua do Matão, travessa 14, 101, 05508-900 São Paulo, SP, Brasil; ²Seção de Aves, Museu de Zoologia da Universidade de São Paulo, Avenida Nazaré, 481, 04218-970 São Paulo, SP, Brasil; ³Department of Environmental Science, Policy & Management, and Museum of Vertebrate Zoology, University of California, Berkeley, CA 94720-3110, USA; and ⁴The Peregrine Fund, 5668 West Flying Hawk Lane, Boise, ID 83709, USA

Summary

1. Inference and estimates of abundance are critical for quantifying population dynamics and impacts of environmental change. Yet imperfect detection and other phenomena that cause zero inflation can induce estimation error and obscure ecological patterns.
2. Recent statistical advances provide an increasingly diverse array of analytical approaches for estimating population size to address these phenomena.
3. We examine how detection error and zero inflation in count data inform the choice of analytical method for estimating population size of unmarked individuals that are not uniquely identified. We review two established (GLMs and distance sampling) and nine emerging methods that use *N*-mixture models (Royle–Nichols model, and basic, zero inflated, temporary emigration, beta-binomial, generalized open-population, spatially explicit, single visit and multispecies) to estimate abundance of unmarked populations, focusing on their requirements and how each method accounts for imperfect detection and zero inflation.
4. Eight of the emerging methods can account for both imperfect detection and additional variation in population size in the forms of non-occupancy, temporary emigration, correlated detection and population dynamics.
5. Methods differ in sampling design requirements (e.g. count vs. detection/non-detection data, single vs. multiple visits, covariate data), and their suitability for a particular study will depend on the characteristics of the study species, scale and objectives of the study, and financial and logistical considerations.
6. Most emerging methods were developed over the past decade, so their efficacy is still under study, and additional statistical advances are likely to occur.

Key-words: abundance estimation, count data, detection, distance sampling, hierarchical model, multispecies abundance model, *N*-mixture model, population size

Introduction

Inference and estimates of abundance are critical for quantifying population dynamics and the impacts of environmental change. Conducting a census (i.e. counting all individuals) of almost any wild animal species is usually not possible. As a result, researchers have often used surrogate measurements thought to be proportional to population size, such as relative abundance (i.e. the count itself or density) or an index of abundance (e.g. individuals detected per unit effort, Buckland, Marsden & Green 2008). Traditional approaches to population size estimation usually adopt sampling designs that attempt to control for the factors influencing detection of indi-

viduals (e.g. observer, time of day, weather). These approaches have limited ability to incorporate the survey-, site- and species-level processes that differentially affect the detection of species or individuals (Iknayan *et al.* 2014). As such, it is common for the resulting counts to include a disproportionate number of absences (i.e. many zeroes), a circumstance called ‘zero inflation’ (Martin *et al.* 2005).

Imperfect detection and zero inflation in count data have been increasingly discussed in the ecological literature over the past decade (MacKenzie *et al.* 2002; Martin *et al.* 2005; Royle & Dorazio 2008). When unaccounted for, both can introduce considerable estimation error and obscure important ecological patterns (Wenger & Freeman 2008). For example, models that generate so-called detection-naïve estimates of abundance, such as a Poisson regression or a generalized linear model (GLM) with a Poisson error distribution, perform poorly in

*Correspondence author. E-mail: fvdenes@gmail.com

the presence of detection error (Sólymos, Lele & Bayne 2012). Furthermore, some environmental factors that influence population size (e.g. forest understorey density, and water depth or clarity) may also affect detection (Kéry 2008; Sólymos, Lele & Bayne 2012). Recognizing that perfect detection is rare in ecological data has led to the development of statistical methods to account for detection error in counts of unmarked populations (i.e. when animals are not individually identified) that formally accommodate the detection process (Buckland *et al.* 2001; Royle & Dorazio 2008).

Here, we review how detection error and zero inflation in count data of unmarked individuals inform the choice of analytical methods for estimating the size of unmarked populations. We begin by describing the sources of variation and types of zeroes that frequently arise in count data. Then, we review commonly used and recently developed methods to model abundance of unmarked populations. We focus on how each method accounts for imperfect detection and zero inflation. Finally, we discuss the performance and sampling design requirements of these methods in the context of surveying and counting unmarked populations. Our goal was to make recent statistical advances accessible for ecologists who wish to estimate population size.

Causes of variation and types of zeroes in count data

Variation in count data results from several distinct processes and understanding them is important to estimate abundance accurately (Martin *et al.* 2005). The number of individuals counted in a survey depends on the underlying (true) abundance of individuals and on their detectability. True abundance may vary among sites or sampling periods as a result of multiple ecological processes, including climatic variation, seasonal or environmental gradients, metapopulation dynamics, species interactions and density dependence. Identifying these processes and understanding their relative importance is a common goal of research programmes.

Uncertain detection probabilities – either caused by failure to detect an individual or by misidentification – have implications for assessments at the individual, survey, species and community levels. Differences in species traits, such as conspicuousness, behaviour, life history and rarity, can affect detection (Iknayan *et al.* 2014). In addition, individual traits, such as sex, age, or distance to the observer, may be important causes of non-detection. Detectability also varies due to survey-specific factors, such as effort, observer, weather, sampling method, or time of day or year. It can vary among sites, due to factors that influence visual or auditory detection regardless of observer, such as habitat structure or noise (Alldredge, Simons & Pollock 2007b). Importantly, the detection probability of a species at a site can also depend on its abundance at the site (McCarthy *et al.* 2013; Warren *et al.* 2013). In this context, a non-zero count is the product of the underlying non-zero abundance and the detection process. However, variability in count data can also be due to misidentification (i.e. a false posi-

tive). Perhaps because it is generally presumed that field surveys are performed by adequately trained and experienced scientists and technicians, most abundance estimation methods assume that false positives are absent from data sets, and there are considerably fewer studies employing analysis methods that account for misidentification errors (but see Royle & Link 2006; McClintock *et al.* 2010 and Miller *et al.* 2013b for applications to occupancy). The demand for methods that account for false positives is likely to expand as large-scale population studies and monitoring programmes make increasing use of citizen science data sets, where false-positive observations are probably more common (Miller *et al.* 2013b).

Counts of zero, on the other hand, can arise from several different mechanisms (Fig. 1). Zeroes due to ecological processes are true zeroes, in the sense that species are absent from the site (Fig. 1a). In many investigations this is interpreted to be a consequence of unsuitable habitat or competitive exclusion (Martin *et al.* 2005). However, a second type of true zero is induced by demographic stochasticity (Fig. 1a), when species fail to saturate all suitable habitats due to random local extinctions and dispersal limitation (Martin *et al.* 2005). Species rarity can increase the frequency of this type of zero in count data (Fig. 1b), due to a higher probability of local extinctions when populations are small.

Additional true zeros can also occur when the occupancy pattern and local abundance distribution of a species are the outcome of distinct processes, such as habitat selection operating at different spatial and temporal scales (Wenger & Freeman 2008). Consider, for example, animals that specialize in a spatially restricted resource, such as a hawk that mainly consumes aquatic prey. Except during dispersal and migration, this species is likely to be found only near rivers, lakes, ponds and other water bodies, that is, its potential occupancy is dependent on the presence of these habitats. Regional surveys may include sampling units in or near such habitats, but will also likely include dry sites where the species will be absent. In the wet sites, the abundance of the species might be strongly predicted by certain environmental factors (e.g. proportion of natural vegetation area, degree of fragmentation, land use). Such factors may favour the species in a dry site, but it will be absent because there is no body of water. Large numbers of these types of absences, which are more common in sparsely than in widely distributed species (Fig. 1a, non-occupancy true zeroes), induce error in analysis and confound interpretation because models will attempt to estimate abundance of the species in a habitat where the species does not occur. Sampling only sites that have potential to be occupied avoids this non-occupancy zero inflation, as does incorporating the factor determining occurrence of the species in models. However, this solution requires *a priori* knowledge of what habitats to survey, and it cannot be applied to multispecies surveys that broadly sample habitats (Wenger & Freeman 2008; Joseph *et al.* 2009).

False zeros, on the other hand, result from the observation process and can arise in two ways: (i) an individual may be

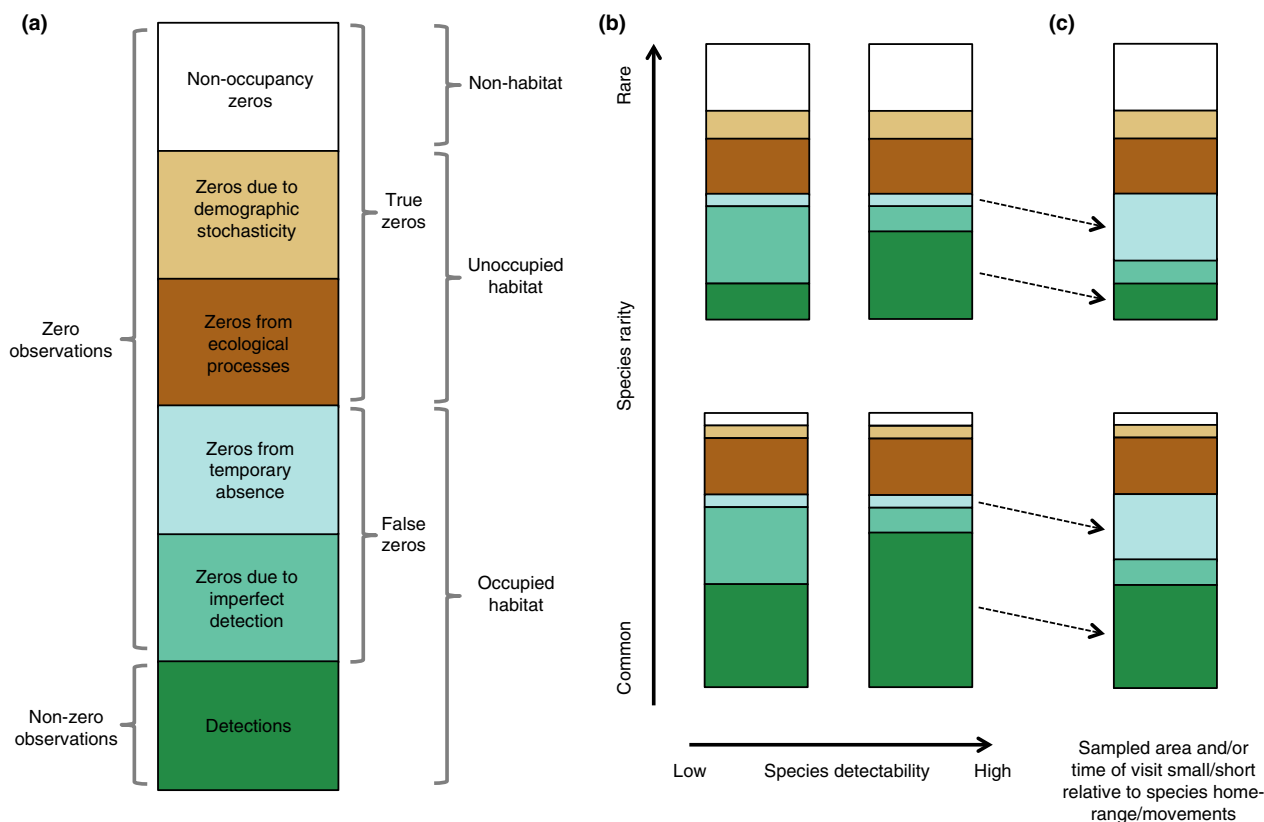


Fig. 1. Mechanisms that cause different types of zero observations in count surveys and how species rarity, detectability and sampling effort affect them. (a) False zeroes are due to either imperfect detection or temporary absence. True zeroes can occur when the sample unit is unoccupied by the species, due to demographic stochasticity or due to ecological mechanisms such as unsuitable habitat or interspecific competition. (b) For common and detectable species (lower right), the majority of zeroes can be expected to result from ecological processes. As species detectability decreases, new false zeroes arise due to detection error (lower left). Species rarity results in fewer detections (dark green bars), additional true zeroes arise from unoccupied sample units (white bars) and increased demographic stochasticity (beige bars). (c) When the area sampled and/or the time of visit are small/short relative to the species home range or movements, individuals may not be available for detection during the survey, resulting in additional false zeroes and fewer non-zero observations.

present at the site but be undetected by the observer (Fig. 1b), which is commonly called ‘detection error’ or ‘imperfect detection’ (MacKenzie *et al.* 2002; Royle & Dorazio 2008), and (ii) an individual of a mobile species with a large home range may regularly use a site but be absent from it at the time of survey, because it is visiting part of its home range outside of the sample unit. The latter may commonly occur if the sampling area is small and/or the length of visit is short relative to the movements of the species, and it is often termed ‘temporary emigration’ or ‘temporary absence’ (Fig. 1c, Tyre *et al.* 2003; Chandler, Royle & King 2011). Hibernation and aestivation underground could also be considered as a type of temporary emigration, because animals are present, but unavailable for detection on the surface. Such causes of false absence are likely to occur during animal surveys, because many species have large home ranges or are undetected due to cryptic behaviour and camouflage. As discussed above for non-zero observations, detection is also influenced by the type of habitat where an observation is made, time of day when it takes place, weather, distance to the observer and variation between observers (Buckland *et al.* 2001; Alldredge, Simons & Pollock 2007b; Iknayan *et al.* 2014).

Modelling methods for estimating abundance of unmarked animal populations

In this section, we review established and emerging methods to model abundance of unmarked populations, and focus on how they handle imperfect detection and zero inflation (Fig. 2, Table 1). We start with Poisson and negative binomial GLMs, which are simple and widely used methods for modelling non-normal data (including counts) that do not explicitly model detection. Next, we describe distance sampling methods that model detectability based on the distance between the observer and the animal. We then discuss hierarchical (N -mixture) models, which estimate detectability based on multiple visits, followed by the single-visit N -mixture models that employ covariate data to model detection instead of multiple visits. Finally, we describe multispecies N -mixture abundance models that account for detectability and make inference about the number of species not detected during surveys in the study region. They can also be used to estimate community measures such as species richness, diversity and similarity. We do not include methods that generate indices related to population size, such as time to detection models (Alldredge *et al.* 2007a;

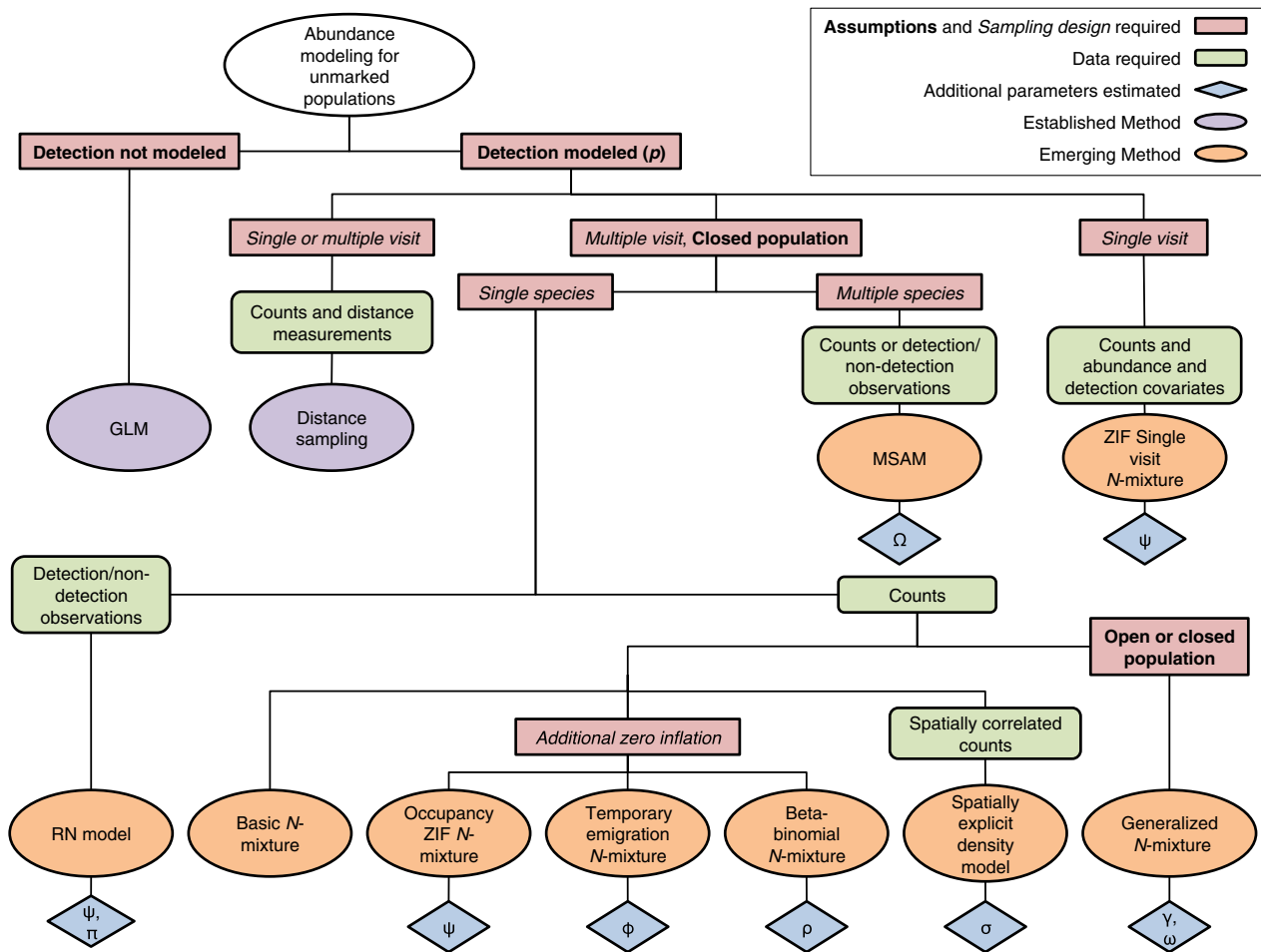


Fig. 2. Summary of the main modelling approaches for estimating abundance of unmarked animal populations described in the text. Red boxes represent important model assumptions (in bold) and sampling design requirements (in italic), green boxes represent the types of input data used by each model, lilac and orange ellipses represent established and emerging methods, respectively, and blue diamonds represent additional parameters estimated. ψ indicates models that estimate potential occupancy probability, ϕ indicates models that estimate probability of temporary emigration from the sample unit, and ρ indicates models that account for correlation in detection of individuals. π is site-level detection probability, γ and ω are arrival rate and survival probability parameters, respectively, σ is the spatial correlation in counts, and Ω is the probability that a species is present in the supercommunity.

McCarthy *et al.* 2013), rather than direct estimates of abundance.

ESTABLISHED METHODS

Poisson and negative binomial GLMs

Abundance of unmarked populations has often been estimated from count data using Poisson GLMs (Nelder & Wedderburn 1972; Zuur *et al.* 2009). A GLM is a generalization of the linear regression that allows the response variable to have error distributions other than Gaussian; in this case, it assumes the error term follows a Poisson distribution. Variation is described by environmental variables assigned as covariates using a log-link function. The choice of distribution arises from assumptions of how organisms are distributed in space in a homogeneous landscape. Potential departures from randomness arise as a result of ecological heterogeneity and are explained by covariates. The negative binomial (NB)

distribution allows the error distribution to vary independently of the mean by including an overdispersion parameter (k) and has been suggested as an alternative to account for extra Poisson variation (Zuur *et al.* 2009). It is also used when the assumption of independence of observations is violated, such as when individuals occur in aggregations. Negative binomial (NB) models are not strictly GLMs due to the lack of a canonical link function (e.g. log link for the Poisson GLM), but they can be fit using a small extension of the GLM approach by iteratively fitting the k parameter and then fitting the rest of the model with a fixed k parameter (Bolker 2008). Due to this similarity and for simplicity, we refer to both Poisson and NB models as GLMs, but readers should be careful when using this terminology in other contexts.

GLMs are frequently used with count data because they are easy to build, can be applied to the simplest count data sets (e.g. n sampling sites each visited once during a sampling period), work in both frequentist and Bayesian frameworks (Bayesian estimates under a vague [i.e. uninformative] prior

Table 1. Established and emerging methods for estimating abundance of unmarked animal populations, with associated data types and sampling requirements. *P* lists the strategy employed by each method to estimate individual-level detection probability (*P*), where π is site-level detection probability. ZIF/other variation indicates the parameters estimated when accounting for zero inflation or other additional variation, where ψ is the potential occupancy probability, $1 - \phi$ is the probability of temporary emigration from the sample unit, ρ is correlation in detection of individuals, γ is arrival rate of new individuals, ω indicates survival probability, σ is the spatial correlation in counts, and Ω is the probability that a species is present in the supercommunity. Covariate handling indicates terms (or levels) that accept covariates in each model

Method	Data	Sampling requirements	<i>P</i>	ZIF/other variation	Covariate handling	Examples
GLMs (Poisson, NB)	Counts	Single-visit surveys	None	–	Relative abundance	Nelder & Wedderburn (1972), Zuur <i>et al.</i> (2009)
Distance sampling	Counts and distance measurements	Single or multiple-visit surveys	Distance to observer	–	Detection, abundance	Buckland <i>et al.</i> (2001), Hedley & Buckland (2004), Royle, Dawson & Bates (2004)
RN model	Detection/non-detection observations	Multiple within-season surveys	Repeated surveys, closed population assumption and π	ψ	Abundance, detection and ZIF	Royle & Nichols (2003), Royle & Dorazio (2008)
Basic <i>N</i> -mixture	Counts	Multiple within-season surveys	Repeated surveys and closed population assumption	–	Abundance, detection	Royle (2004a,b), Royle & Dorazio (2008), Dail & Madsen (2011)
Zero-inflated <i>N</i> -mixture	Counts	Multiple within-season surveys	Repeated surveys and closed population assumption	ψ	Abundance, detection and ZIF	Wenger & Freeman (2008), Joseph <i>et al.</i> (2009)
Temporary emigration <i>N</i> -mixture	Counts	Multiple within-season surveys	Repeated surveys and closed population assumption	ϕ	Abundance, detection and temporary emigration	Chandler, Royle & King (2011)
Beta-binomial <i>N</i> -mixture	Counts	Multiple within-season surveys	Repeated surveys and closed population assumption	ρ	Abundance, detection and detection correlation	Martin <i>et al.</i> (2011)
Generalized <i>N</i> -mixture	Counts	Multiple surveys (not necessarily within seasons)	Repeated surveys and population dynamics parameters	γ, ω	Abundance, detection, arrival and survival rates	Dail & Madsen (2011)
Spatially explicit density model	Spatially referenced counts	Multiple within-season surveys	Spatial correlation in repeated counts	σ	Abundance, detection and spatial correlation	Chandler & Royle (2013)
Single-visit <i>N</i> -mixture	Counts, abundance and detection covariates	Single-visit surveys	Covariates	ψ	Abundance, detection and ZIF	Sólymos, Lele & Bayne (2012)
Multispecies <i>N</i> -mixture	Detection/non-detection observations or counts	Multiple within-season surveys	Repeated surveys and closed population assumption	Ω	Abundance, detection and ZIF	Yamamura <i>et al.</i> (2011), Chandler <i>et al.</i> (2013)

will be numerically close to maximum likelihood estimates [Kéry & Schaub 2012]) and are available in many statistical platforms.

GLMs are appropriate tools for drawing inference on the factors affecting the relative abundance of a species, provided such factors do not also strongly influence detectability. However, when the objective of the study is to obtain an estimate of the abundance, or to draw inference on effects acting on abundance, GLMs cannot account for departures from the error distribution resulting from excessive true or false zeroes (Fig. 1). Consequently, if the data analysed contain this additional variation (i.e. zero inflation), use of the Poisson and negative binomial GLMs will be inadequate (Welsh *et al.* 1996; Joseph *et al.* 2009). Data transformations are ineffective in normalizing zero values because the excessive zeros are simply replaced by an equally frequent non-zero value (Martin *et al.* 2005). An alternative approach used in many studies is to truncate data sets, reducing or even eliminating the zero values before the GLM analysis. This can compromise the estimation of parameters, because some of the zeros removed will likely be true zeros and their corresponding covariate values will be ignored (Joseph *et al.* 2009).

Distance sampling

This is a group of methods in which distances from the survey points or lines to the clusters (or groups) of detected individuals are recorded and used to estimate density or abundance (Buckland *et al.* 2001; Buckland, Marsden & Green 2008). All individuals at zero distance from the sampling point or transect are assumed to be detected, and the probability of detecting an individual (P) is modelled to decline with increasing distance from the point or transect. The distribution of distances recorded with every sighting is used to estimate a detection function (i.e. the probability of detection in relation to distance from the observer), which in turn allows estimation of density and abundance (Buckland *et al.* 2001). Thus, distance sampling methods do not accommodate false zeroes originating from the detection process explicitly, but do so implicitly under the hypothesis that detectability is primarily related to the distance between animals and the observer. A major advantage of the distance sampling methods is that they do not require repeated sampling of sites over time to estimate the individual-level probability of detection, P , as opposed to most N -mixture models (see below).

There are six central assumptions of the basic forms of distance sampling: (i) objects on the line or point are detected with certainty; (ii) objects do not move in response to the observer during a survey or before detection (i.e. they are detected at their initial locations); (iii) distance measurements are exact; (iv) the position of detected individuals is independent of the survey point or line; (v) cluster sizes are recorded without error; and (vi) detections are independent events. Violations of these assumptions may result in biased estimates (Buckland *et al.* 2001; Buckland, Marsden & Green 2008). Distance sampling data are frequently analysed with the free program Distance,

which also provides guidelines for survey design (Thomas *et al.* 2010).

Considerable effort in distance sampling methodology has been focused on modelling covariate effects on P (Marques & Buckland 2003; Marques *et al.* 2007) and on abundance (Royle, Dawson & Bates 2004; Johnson, Laake & Ver Hoef 2010; Chelgren *et al.* 2011; Conn, Laake & Johnson 2012; Oedekoven *et al.* 2013). Royle, Dawson & Bates (2004), for example, proposed a modified distance sampling hierarchical model that treated counts as a function of P and site-specific abundance (N_i , at sampling unit i [points or line segments]). N_i was modelled as a Poisson (or negative binomial) random effect that could be related to covariates through a link function. Site-specific abundances were integrated from the likelihood function, while parameters of the abundance distribution (e.g. mean $[\lambda]$ for the Poisson and negative binomial distributions, and the overdispersion parameter $[\alpha]$ for the latter) were estimated from the data (Royle, Dawson & Bates 2004). This method is a hybrid between distance sampling and N -mixture models and illustrates the opportunities for improving established approaches by incorporating the emerging methods discussed below. Sillett *et al.* (2012) extended this approach to include covariate effects on the detection function. These models are freely available in the *Unmarked* package (Fiske & Chandler 2011) for the free statistical software R (R Development Core Team 2014). Another distance sampling advancement for estimating abundance of unmarked populations is the density surface model (Hedley & Buckland 2004; Katsanevakis 2007; Niemi & Fernández 2010). It can be used to assess the effects of environmental variables on the spatial distribution of individuals and to generate spatial predictions of abundance over larger or different areas from those originally surveyed. This group of methods has been extensively reviewed elsewhere (Miller *et al.* 2013a). A major advantage of density surface and hierarchical distance sampling models (Royle, Dawson & Bates 2004) is that they allow inference from samples that are not randomly placed in the study area, and thus may not be representative of it, which is an important requirement for extrapolating predictions obtained with conventional distance sampling.

EMERGING METHODS

Hierarchical (N-mixture) models for multiple visits

Hierarchical models handle variation in the observed data as a result of explicit *observation* and *state process* components (Royle & Dorazio 2008). Detection error is incorporated in the observation component, while the state process component incorporates the underlying ecological process (i.e. abundance or occupancy). This class of models includes approaches that generally require temporally replicated surveys (i.e. repeated measures) conducted at multiple locations, but a variant of these models can be applied to single-visit data sets as described in the next section.

The repeated measures design was initially applied by MacKenzie *et al.* (2002) and Tyre *et al.* (2003) to detection/non-detection data in occupancy models to estimate the pro-

portion of sites occupied (i.e. containing ≥ 1 individual) while estimating π , the site-level detection probability (i.e. the probability of detecting at least one individual at site i). Note that π differs from the individual-level detection probability (P) discussed above. If, however, individuals are detected independently, under binomial sampling these two detection probabilities can be related as $\pi_i = 1 - (1 - P)^{N_i}$ (Royle & Nichols 2003). Assuming abundance varies among sites, one can expect local variation in π because sites with higher abundance yield more 'net' detections and *vice versa*. This abundance-induced heterogeneity in the probability of detecting a species (i.e. greater at sites with more individuals present) led Royle & Nichols (2003) to propose a model (hereafter, the RN model) to estimate abundance from binary observations of detection/non-detection (i.e. apparent presence/absence of the species at a site).

In the RN model, the detection frequencies for each i site, defined as the number of times the species was detected in the repeated j samples, are assumed to follow a binomial distribution with number of trials equal to j and probability π_i , which in turn depends on the unknown N_i , a random variable with a specified distribution (e.g. Poisson). Thus, the model assumes that site-specific detection probabilities are functionally dependent on local abundance. In other words, it obtains information about the local abundance distribution directly from the apparent heterogeneity in π among sites induced by variation in abundance (Royle & Dorazio 2008). Structurally, the model is analogous to classical generalized random effect models, with N_i as the random effect. Measurable variables that are thought to influence detection at the individual level (based on $\pi_{ij} = 1 - (1 - p_{ij})^{N_i}$) and abundance may be added to the model as covariates using a link function (e.g. log for the Poisson in the abundance portion). As such, the RN model can account for the false zeroes from detection error and the true zeroes from processes that influence abundance. While it is not built with an explicit occupancy probability (ψ), it can be derived as $\psi = \Pr(N > 0)$ by defining unoccupied sites to have a zero detection probability (as a result of zero abundance), and then $1 - \psi$ becomes the point mass at zero abundance (Royle & Nichols 2003).

Unmarked populations are a promising area of application of the RN model. For example, many species occur in low densities and are territorial, favouring occupancy sampling instead of counts. However, the premise of functional dependence between detection probability and local abundance may not be a reasonable assumption in some situations, such as extremely rare or highly territorial species. In such cases, density might be so low that local abundance (N_i), given presence, is essentially constant (i.e. = 1) and thus heterogeneity in apparent detection probability among sites is null or negligible, undermining the fundamental proposition of binomial detection in the RN model.

Obviously, not all species are so rare that only occupancy sampling is possible and many data sets are composed of counts. The basic N -mixture model for count data combines a binomial GLM (for the observed counts) and a standard model (Poisson or negative binomial) for N_i . It assumes that

the population sampled is closed during the sampling period with respect to mortality, recruitment and movement (Royle 2004b). It also assumes that detections at a site are independent and that all individuals recorded at a given site, and time have the same detection probability (Royle 2004b). Counts at site i are regarded as a binomial process dependent upon P (i.e. the observation process) and N_i , a Poisson (or negative binomial) random variable (i.e. the underlying state process). Additional explanatory variables can be included in both the abundance and the detection models using standard generalized linear modelling techniques. The method generates parameter estimates of the abundance distribution across sites (e.g. λ in the case of a Poisson distribution, or λ and α for negative binomial) that allow evaluation of temporal changes or geographic comparisons. Total abundance can be estimated posteriorly if the sample units are of known area (Royle 2004b). These models yield estimates of P , thus accounting for detection error and the resulting false zeros (Royle 2004b). Estimates can be generated with either frequentist (maximum likelihood estimation) or Bayesian approaches (Kéry 2008; Kéry & Schaub 2012). This type of model can be fit using the *Unmarked* R package (Fiske & Chandler 2011) for frequentist analysis, or by Bayesian methods with WinBUGS (Lunn *et al.* 2000), OpenBUGS (Lunn *et al.* 2009) or Jags (Plummer 2003).

The Poisson and negative binomial distributions are usually adequate to model abundance at occupied sites in the N -mixture, but they perform poorly when the data include a large number of true zeros. This can occur when unoccupied sites are sampled (Wenger & Freeman 2008; Joseph *et al.* 2009). Extensions of the N -mixture models were proposed to accommodate excess true zeroes, in addition to the false zeros from detection error, by employing zero-inflated mixture distributions such as the zero-inflated Poisson (ZIP) to model abundance and potential occupancy simultaneously (Wenger & Freeman 2008; Joseph *et al.* 2009). The ZIP distribution is a mixture of a Poisson distribution, with a rate or average parameter (λ) and a Bernoulli distribution with a parameter (ψ) for the probability that a species is potentially present (i.e. 'potential occupancy'), but the site is unoccupied due to other factors acting on abundance (i.e. demographic stochasticity and ecological processes in Fig. 1).

These mixture models allow three sources for zeroes in data – the false zeroes from detection error, and the true zeroes from potential occupancy and abundance. As in the original N -mixture, counts in site i are the result of a binomial process dependent on P and N_i . The latter is, in turn, a product of the potential occupancy process (dependent on ψ) in addition to the abundance process (dependent on λ). Covariates can be used for the detection, presence or abundance terms with appropriate link functions, and the same covariate may be used for more than one process (e.g. per cent forest cover for detection and abundance). Like the basic N -mixture model, use of other zero-inflated distributions, such as the zero-inflated negative binomial (ZINB), for modelling the abundance-potential occupancy portion is possible (Wenger & Freeman 2008; Joseph *et al.* 2009). However, this can often result in unrealistic parameter estimates despite good model fit, so should be used with

caution (Joseph *et al.* 2009). Zero-inflated N -mixture models inherit the assumptions of the original N -mixture model – population closure, independence of observations and choice of error distribution.

Recent developments in N -mixture models focus on handling violations of assumptions or sampling requirements: (i) Chandler, Royle & King (2011) developed a variation of the N -mixture model to explicitly account for temporary emigration by adding a binomial zero-inflation level, whereby each individual is considered to be within the survey plot with probability ϕ (and $1-\phi$ is the probability of temporary emigration). Variation in ϕ can be modelled as a function of site and survey-specific variables, although often this variation will more likely be a result of disparity between the sizes of survey plots relative to the species' home range (Chandler, Royle & King 2011); (ii) Martin *et al.* (2011) developed a beta-binomial N -mixture model to account for non-independent (correlated) detection of individuals. This can occur when the behaviour of an animal affects the probability of detection of other individuals by the observer (e.g. when singing behaviour in birds and amphibians elicits a response from a neighbouring individual). The binomial distribution in the observation model (with parameters N and P) is replaced by the beta-binomial (with N , P and the correlation parameter ρ); larger values of ρ indicate stronger correlation (Martin *et al.* 2011).

Dail & Madsen (2011) proposed a generalized N -mixture method that can be used to formally test the closure assumption through estimation of parameters of population dynamics. When applied to data from an open population (e.g. annual counts), the model estimates arrival rate of new individuals (γ , which includes both births and immigrants) and survival probability (ω), which can in turn be used to estimate abundance while accounting for imperfect P . No additional information is required to estimate γ and ω : the generalized model structure replaces the assumption that N is constant at each site during the sampling period with estimation of these parameters based on variation among sampling occasions at each site. In addition, because the model does not assume a closed population, it is not necessary for the repeated visits to be performed within a single season. As a result, multiple-season data sets with a single visit per season can be analysed (e.g. the North American Breeding Bird Study). Description of the generalization procedure is extensive, so we refer readers to the original article for details (Dail & Madsen 2011). The model is freely available in the *unmarked* R package (Fiske & Chandler 2011).

Yet another recent development of N -mixture models for unmarked populations is the spatially explicit density model (Chandler & Royle 2013), which is an adaptation of spatial capture–recapture (SCR) models for estimating density of marked individuals (Efford 2004; Borchers & Efford 2008). SCR models estimate density or population size based on estimation of the activity centres of individuals in the sampled area from encounter histories using capture/recapture data. Activity centres, which cannot be directly observed, are the spatial average of an individual's locations during a time period. The model uses the information from the spatial coordinates of

traps with captured individuals to determine the locations of the activity centres, while capture probability (the equivalent of detection probability for capture/recapture data) is regarded as a function of the distance between survey (i.e. trap) locations and activity centres, similar to the detection function in distance sampling (Buckland *et al.* 2001). The spatially explicit density model has a structure similar to the SCRs, but uses the spatial correlation (σ) in temporally replicated counts to estimate the number and location of the activity centres instead of individual identification. Because spatial correlation between counts is required, sample locations need to be in close proximity to one another relative to the size of a home range to allow individuals to be detected at multiple locations over the repeated visits. An interesting aspect of the method is that, unlike most other approaches for estimating abundance or density, spatial correlation is not viewed as an obstacle for inference, but instead informs estimation of distribution and population size (Chandler & Royle 2013).

Single-visit N -mixture models

By using variation in counts between visits to adjust for detection error, many of the multiple-visit methods described in the previous section can produce more accurate estimates of abundance than estimators that do not explicitly model detection (e.g. GLMs). This, however, is not the only way to model detectability with N -mixture models, as exemplified by the distance sampling and spatially explicit density models that use spatial information to estimate the detection function (Chandler & Royle 2013). Detection error can also be modelled using only a single visit to each site, for both presence/absence and count data, if covariates that affect detection and abundance are available (Lele, Moreno & Bayne 2012; Sólymos, Lele & Bayne 2012), such as those discussed above in the section 'Causes and types of zeros in count data'.

The binomial–ZIP mixture model for analysing single-visit count data in the presence of zero inflation and detection error replaces both the need for repeated visit data and the assumption of population closure required by multiple-visit approaches with the use of non-overlapping sets of covariates that affect detection and/or abundance (Sólymos, Lele & Bayne 2012). The detection and abundance covariate sets need to differ by at least one continuous covariate (i.e. they can share covariates provided that at least one continuous covariate is unique to either set). As in the multiple-visit zero-inflated N -mixture approach, the model is built with detection, abundance and zero-inflation terms, and their respective error distributions, link functions and covariates. Because the abundance at each i location is unknown, the likelihood function involves summation over all possible values of N_i . Direct maximization of the likelihood function in this condition can lead to considerable confounding between the zero-inflation parameter and the intercept parameter in the detection model, so conditional likelihood is employed. In this procedure, maximization is conditioned on a sufficient statistic for the zero-inflation parameter, and a conditional distribution of the data is used to estimate the detection and

abundance parameters. To estimate the zero-inflation parameter, a new random variable is constructed with the non-zero observations [$W_i = I(Y_i > 0)$, where I is an indicator function such that $I(Y_i > 0) = 1$ if $Y_i > 0$ and $I(Y_i > 0) = 0$ otherwise], and the likelihood function for the detection, abundance and zero-inflation parameters is written based on the distribution of W_i . Then, with values of the abundance and detection parameters fixed at their conditional likelihood-based estimates, the function is maximized with respect to the zero-inflation parameter to obtain its estimate. This new likelihood function does not involve infinite summation and hence is easy to maximize. Thus, conditional likelihood separates the parameter space and reduces the extent of confounding (Sólymos, Lele & Bayne 2012).

Like the multiple-visit binomial-ZIP mixture model described in the previous section, the single-visit N -mixture model includes an ecological process level, an observation level and an additional zero-inflation level. Thus, it can account for detection error and other forms of zero inflation, such as those derived from non-occupancy or temporary absence. The method is implemented in the R package *detect* (Sólymos, Moreno & Lele 2013). The *pcount* procedure in package *unmarked* also permits modelling abundance with single-visit data, but with direct maximization of the likelihood.

Multispecies N -mixture abundance models (MSAMs)

Although abundance is a valuable attribute for evaluation and comparisons, parameters such as species richness, diversity and similarity are also important to understand community-level variability (Dorazio & Royle 2005; Dorazio *et al.* 2006; Yamaura *et al.* 2012; Iknayan *et al.* 2014). Yet, methods for estimating these community metrics, many of which require measures of species abundance, have remained separate until recently. Multispecies abundance models (MSAMs) are an extension of multiple-visit single-species abundance models that analyse the detection histories (i.e. the repeated counts) of all species encountered. The detection histories are used to inform the estimation of diversity, richness and derived metrics, including the number of species that were present in the community, but were not detected at any site – a measure that is useful for communities dominated by rare species (Dorazio & Royle 2005; Iknayan *et al.* 2014). MSAMs are still in their infancy, so there are very few applications to date (Yamaura *et al.* 2011, 2012; Chandler *et al.* 2013). However, MSAMs draw much of their structure from multispecies occupancy models (MSOMs), which have a longer history of use (see Iknayan *et al.* 2014 for a review).

Initial MSAMs were developed by combining two different modelling frameworks based on detection/non-detection data (Yamaura *et al.* 2011): (i) the RN model for estimating abundance (Royle & Nichols 2003) and (ii) an MSOM that allows for estimation of species richness and community composition at a given site by accounting for both undetected species and variability in occupancy and detectability among species (Dorazio & Royle 2005; Dorazio *et al.* 2006). The model

requires repeated visits at multiple sites to collect detection/non-detection data for each species. These detection histories are then linked to species abundance based on the RN model, and variation in detectability and abundance across sites can be modelled as a function of site-specific covariates for each species. Further development of MSAMs allows use of counts of individuals instead of detection/non-detection data (Yamaura *et al.* 2012), as in single-species N -mixture models for count data.

For well-detected species, model parameters including covariate coefficients can be independently estimated, but this is not the case for rare species due to insufficient data. The MSAM has an additional hierarchical level that treats each parameter as an independent, normally distributed random effect across species, that is, the value of each parameter for each species is assumed to be drawn from a normal distribution with mean and standard deviation that represent the mean response across species and the standard deviation among species (Yamaura *et al.* 2011; Iknayan *et al.* 2014). The mean and standard deviation community parameters are termed hyperparameters.

Data augmentation is used in MSAMs to estimate the number of species present in the community, but not detected at any site. In addition to the ecological process (i.e. abundance) and observation (i.e. detection) levels already present in single-species hierarchical models, the multispecies approach has a supercommunity (data-augmentation process) level, with parameter Ω . The supercommunity comprises the observed species (s) and an arbitrary but known number (m) of unobserved species. The inclusion rate (Ω) is the probability a species that belongs to the supercommunity is sampled (Royle & Dorazio 2008; Iknayan *et al.* 2014). In the data-augmentation approach, m all-zero detection histories are added to those of the s observed species, and used as input data for the model. An indicator variable that separates the data into species present (detected or not), and those not present in the community (and hence not detected) is also added to the model (Royle & Dorazio 2008). This variable is indexed by Ω , which is the parameter estimated. The number of species in the region (i.e. gamma diversity) is obtained by multiplying the estimate of Ω by the sum of s and m .

Both MSAM models (detection/non-detection and count-based) accommodate, but do not require, the inclusion of site and survey-specific covariates and also allow for estimation of community-level metrics derived from richness, such as diversity and similarity indices like species-accumulation curves (Royle & Dorazio 2008; Iknayan *et al.* 2014). The supercommunity level of the model, by modelling the presence or absence of each species, already handles the zero inflation derived from non-occupancy. Moreover, use of ZIP or ZINB distributions for the ecological process level should be possible to account for additional zero inflation (e.g. temporary absence), but to our knowledge have not yet been implemented. Because data augmentation requires estimation of the parameters by Markov chain Monte Carlo (MCMC), current implementation of multispecies abundance models is mostly restricted to Bayesian inference programs.

Discussion

In this section we summarize and compare the performance of the emerging methods for estimating abundance described above, focusing on estimation bias and sample size requirements. We conclude by discussing study design and how different analytical methods are best suited for particular situations.

METHOD PERFORMANCE

Detection-based estimation approaches generally performed well when evaluated against simulated data sets (Table 2). They typically estimated population size without strong bias, except in scenarios when detection probability was low and few sites were sampled a small number of times. An encouraging result from the RN model, which likely applies to the other methods, is that bias from small sample size can be countered by increasing the number of visits to each site (from 3 to 5–10) or whether the species has a high detection probability ($P \geq 0.3$, Royle & Nichols 2003). When sample size was small,

which sometimes skewed the mean of the abundance estimator, the median and mode were close to their true values (Royle 2004b). Similarly, in the spatially explicit density model, low spatial correlation in counts ($\sigma = 0.5$) results in a biased mean of population size estimates (5–10%), but their mode is unbiased (Chandler & Royle 2013). Simulations with the generalized N -mixture model showed that population dynamics were falsely detected at very low rates (< 1%) in closed population scenarios ($[\gamma, \omega] = [0, 1]$), suggesting the model performs adequately as a test of the closed population assumption (Dail & Madsen 2011).

Few field studies have applied multiple estimation methods to compare their relative performance. Abundances were higher when estimated with N -mixture models compared to estimates derived from territory mapping of birds (Kéry, Royle & Schmid 2005) and from distance sampling transects of desert tortoises (Zylstra, Steidl & Swann 2010), but not compared to direct observations of lizards (Doré *et al.* 2011). Precision (i.e. confidence intervals) of estimates varied among methods; estimates from N -mixture models had higher precision compared

Table 2. Summary of published simulation studies for emerging methods of estimating abundance, with simulation settings, effects and considerations for sample size. R is number of sites, T is number of visits, λ is the mean of the Poisson/NB distribution, P indicates detection probability, $1 - \phi$ is the probability of temporary emigration, ρ is correlation in detection of individuals, γ is arrival rate of new individuals, ω indicates survival probability, σ is the spatial correlation in counts and ψ is the potential occupancy probability

Method	Effects on abundance estimation	Sample size effects	Simulation settings	References
RN model	Bias of ~10–15% when $P = 0.1$;	Small at $R \leq 100$; Larger T and/or P counters small R	Wide (R, T) range, $\lambda = (4.61, 2.30, 1.61, 1.20, 0.92, 0.69, 0.51)$, $P = (0.1, 0.2)$, NB mean/variance ratios = (1.2, 2, 4)	Royle & Nichols (2003)
Basic N -mixture	Small positive bias when $R (P) = 20 (0.25)$	Adequate at $R = 20$	$R = (20, 50)$, $T = (3, 5, 10)$, $\lambda = (2, 5)$, $P = (0.25, 0.50)$	Royle (2004b)
Temporary emigration N -mixture	Unbiased in random emigration and spatially explicit emigration scenarios	Adequate at $R = 100$, lower values not investigated	$R = 100$, $T = 3$, $\lambda = 3.14$, $\phi = 0.5$, random emigration scenario vs. spatially explicit scenario (emigration dependent on home range size = [0, 10, 30])	Chandler, Royle & King (2011)
Beta-binomial N -mixture	Better performance than the Basic N -mixture in the scenario with correlated detection (bias of 2% vs. 123%, respectively)	Adequate at $R = 200$, lower values not investigated	$R = 200$, $T = 3$, $\lambda = 4.6$, no correlation scenario with $P = 0.5$, correlation scenario $\rho = (0.15, 0.3)$ and $P = (0.7, 0.5)$, zero-inflated scenario with $\psi = 0.25$.	Martin <i>et al.</i> (2011)
Generalized N -mixture	Less biased than the Basic N -mixture in open population scenarios ($[\gamma, \omega] \neq [0, 1]$)	Adequate at $R = 20$; better results with $R = 100$	$(R, T) = (100, 10)$, $(R, T) = (20, 5)$, $\lambda = (2, 5)$, $P = (0.25, 0.50)$, $\gamma = (0, 1, 2)$, $\omega = (1, 0.8, 0.5)$	Dail & Madsen (2011)
Spatially explicit density model	Mean of estimates is biased (5–10%) when $\sigma = 0.5$, but mode is unbiased. With higher σ , estimates are unbiased. Bias is also reduced when effective encounter rate is increased.	15×15 sample point grid adequate, lower values not investigated	15×15 sample point grid, $\sigma = (0.5, 0.75, 1)$, effective encounter rate = (2.5, 5). Population size of (27, 45, 75)	Chandler & Royle (2013)
Single-visit N -mixture	Overestimated with $R = 100$, converged to true values as R increased	Small at $R < 100$	$R = (100, 300, 500, 700, 1000)$, $\lambda = (2.13, 5.25)$, $P = (0.25, 0.65)$, $\psi = (0, 0.25)$	Sólymos, Lele & Bayne (2012)

to other methods for lizards (Doré *et al.* 2011), but not for tortoises (Zylstra, Steidl & Swann 2010). Couturier *et al.* (2013) extended the tortoise studies to compare abundance estimates from capture–recapture, distance sampling and *N*-mixture models, conducted simulations to assess bias between capture–recapture and *N*-mixture models, and computed a power analysis to evaluate the ability of the three methods to detect changes in abundance. The capture–recapture method resulted in abundance estimates 1.75 and 2.19 times greater than distance sampling and *N*-mixture models, respectively. Simulations showed that the *N*-mixture models resulted in estimations that were biased high when detection probabilities were < 0.5, whereas capture–recapture estimations were unbiased. That the *N*-mixture method showed less precision than distance sampling and capture–recapture in a species with low detectability is not surprising, given the additional information provided by distance measurements and individual marking. None of the methods were precise enough to detect small (< 1%/year) population changes (Couturier *et al.* 2013). Martin *et al.* (2011) applied binomial and beta-binomial *N*-mixture models (both with a Poisson abundance distribution) in a Bayesian approach to aerial survey data of manatees, in which correlated surfacing behaviour caused non-independent detection of individuals and assessed their fit using posterior predictive distributions (Gelman, Meng & Stern 1996). They found that the former model did not fit the data, whereas there was no evidence of lack of fit for the latter model.

CONSIDERATIONS FOR STUDY DESIGN AND CHOICE OF ANALYTIC FRAMEWORK

The modelling approaches reviewed here can be used to investigate many kinds of ecological questions – not only hypothe-

ses related to population size and the factors that affect it, but also questions about detection and community characteristics (see Table 3 for examples). Banks-Leite *et al.* (2014) argued that, because covariates of detection probability can be controlled through sampling design, unadjusted estimates of abundance generated with Poisson and negative binomial GLMs can be used for estimating population trends and to identify environmental factors that influence population size, despite the fact that these methods do not explicitly model detection. While careful sampling design can reduce the effects of covariates thought to influence detection probability (e.g. time of day, season, weather, habitat, trap type, observer, etc.), in the absence of modelling detection probabilities it may be difficult to distinguish between model outcomes that reliably reflect ecological processes and those that are related to detectability effects. We focus the remainder of the Discussion on methods that explicitly incorporate detection probability (Fig. 2, Tables 1 and 2).

Estimation methods can be divided into those that require one visit and those that require multiple visits (Fig. 2, Table 1). In the latter, sampling logistics often impede resurveys on the same day. Other factors, such as species’ behavioural traits, may also require that observers return to sites on different days to ensure independence of surveys. Visiting a site on several dates within a sampling season increases travel costs and personnel time. For a given budget, this will likely reduce the number of sites that can be surveyed, potentially decreasing the generality of the study (Lele, Moreno & Bayne 2012). This issue becomes especially important when sampling species that occur in low densities, because surveying a large number of sites is crucial to obtain enough non-zero observations for reliable population inferences. Trade-offs between the number of visits and number of sites surveyed must be considered

Table 3. Common questions asked in ecological studies, and the established and emerging methods for estimating abundance of unmarked animal populations that can be used to address them. GLM includes both Poisson and NB GLMs. DS is distance sampling. For the *N*-mixture methods, ZI is the zero-inflated model, Temp is the temporary emigration model, Beta-bin is the beta-binomial model, Gen is generalized model, Space is the spatially explicit density model, 1 visit is the single-visit model and Multispec is the multispecies approach

Objective	Questions often asked in ecological studies	GLM	DS	Type of <i>N</i> -mixture model							
				Basic	ZI	Temp	Beta-bin	Gen	Space	1 visit	Multispec
Detection	How is detection of individuals affected by covariates?		x	x	x	x	x	x	x	x	x
	Is detection of individuals non-independent due to correlated behaviour?						x				
Population	How is relative abundance affected by covariates?	x									
	How is abundance or density affected by covariates?		x	x	x	x	x	x	x	x	x
	What is the population trend?	x	x	x	x	x	x	x	x	x	x
	What is the estimated abundance of individuals?		x	x	x	x	x	x	x	x	x
	What is the probability of non-occupancy/temporary emigration?					x	x				x
Community	What are the arrival and survival rates							x			
	What is the species richness and diversity?										x
	How do covariates affect species richness/diversity?										x

carefully with regard to the objectives of the study. Monitoring populations in small to moderately sized areas may be a more likely scenario for use of multiple-visit methods than regional-scale studies. For the latter, single-visit methods may be a cost-effective option. Below, we first discuss the single-visit methods and return to the multiple-visit approaches later.

Single-visit methods (Fig. 2, Table 1) replace repeated visits with auxiliary data to estimate detection probability. The most widely used single-visit method is distance sampling, which assumes that the distance to observer measurements are made with minimal error and that detection at the survey point or line is perfect (Buckland *et al.* 2001). In practice, distances to observer can be difficult to accurately measure if surveys rely on aural detections (Allredge, Simons & Pollock 2007b), or if surveys include a considerable proportion of fast-moving individuals (e.g. birds in flight). To illustrate the latter, we describe encounter data for 10 common raptor species detected during roadside strip-transects in open habitats in west-central Brazil (Dénes, F.V. unpublished data). Eight of the 10 species had more than 20% of individuals detected in flight, suggesting that a large proportion of the distance to observer measurements might not be reliable for density estimation with distance sampling. Thus, it may be useful to conduct an assessment of the mode of detection (e.g. the ratio of flying vs. stationary, or aural vs. visual detections) when designing a study before deciding to adopt distance sampling; data sets in which detections are frequently of non-stationary, or calling individuals are probably better suited for other methods.

Point or transect counts are commonly performed along roads, rivers or paths, especially counts focusing on larger, highly mobile species (e.g. raptors, parrots and some waterfowl), due to ease of access and time and resource limitations (Buckland, Marsden & Green 2008). Many factors along such features can lead to atypical density, including an increased frequency of perch sites (e.g. power and telephone posts and lines, fences, road-signs), more food for scavenging species (i.e. road-kills) and increased edge habitats (Buckland, Marsden & Green 2008). Although this can present a problem for any analytical method when extrapolating over a larger region, distance sampling is particularly vulnerable because it relies so heavily on the spatial distribution of individuals in the vicinity of the sample point or line. Alternative strategies have been suggested to handle this problem, such as placing transects perpendicular to roads or paths (Buckland, Marsden & Green 2008), using multiple independent observers to estimate detection probability (Marques *et al.* 2010), recording locations of detected individuals at snapshot moments (Buckland 2006) or recording distances to detected cues (e.g. songbursts) rather than individuals (Buckland 2006), but they are not always feasible or logistically viable. In such situations, it might be better to replace conventional distance sampling with an alternate detection-based method to model count data (Fig. 2, Table 1). An interesting option still in the realm of distance sampling is to employ either the hierarchical distance sampling model of Royle, Dawson & Bates (2004) or a density surface model approach (Hedley &

Buckland 2004; Miller *et al.* 2013a), both of which allow extrapolation of detection-adjusted abundance predictions generated from samples lacking random placement that may not be representative of the study area.

The *N*-mixture method for single-visit count data (Sólymos, Lele & Bayne 2012) requires non-overlapping sets of covariates for detection and abundance. Most research and monitoring studies designed to compare abundance between different conditions or sampling periods are likely to collect detection and abundance covariates. For example, detection covariates often include time of day, weather, habitat and observer, which are frequently available for most data sets. If, however, the proper covariates are not collected and modelled, abundance estimates will be inaccurate regardless of the number of sampling visits (Sólymos, Lele & Bayne 2012). Hence, the need to collect covariates to use the single-visit approach can hardly be considered an objection to its use. When limited budget or sampling logistics preclude the use of the multiple-visit *N*-mixture or distance sampling methodologies, the single-visit method accommodating detection error and zero inflation provides an interesting alternative.

Focusing on the number of occupied sample units using multiple visits instead of on the number of animals in each unit has been suggested to reduce the efforts and costs associated with monitoring (MacKenzie *et al.* 2002; Royle & Nichols 2003). The RN model follows this general approach to produce an estimate of population size (Fig. 2, Table 1), and it is especially useful for situations where little information is available beyond detection/non-detection data (Royle & Nichols 2003). For example, detection of multiple individuals seldom occurs when monitoring of rare or very elusive species (with the caveat that cases of extreme rarity can be a problem even for the RN method). When it is not possible to ensure that individuals are not counted more than once during a visit, reduction of the count data to detection/non-detection for use in the RN model may be a useful approach.

When reliable multiple-visit count data are available, several *N*-mixture methods (Fig. 2, Table 1) allow researchers to model not only abundance and detection but also additional phenomena, such as temporary emigration, non-independent detection of individuals, population closure (a basic premise of the multiple-visit approach itself) and population dynamics parameters (i.e. arrival rate of new individuals [births and immigrants] and survival probability; Table 2). The versatility of this framework should encourage researchers to collect multiple-visit count data whenever their research system and budget allow.

Recently, the use of single-species models that adjust for imperfect detection in situations where communities are dominated by rare species, such as tropical rain forests, has been questioned due to their requirement for high detection or recapture probabilities, which are unlikely to be obtained for rare species (Banks-Leite *et al.* 2014). The multispecies *N*-mixture framework uses data from the entire sample (common and rare species) to inform a community-level distribution of detection and occupancy or abundance probabilities (via hyper-parameters) from which estimates for each species,

including rare ones, are obtained. This ‘information sharing’ allows data to be used more efficiently compared with single-species models, and individual estimates are improved (Dorazio & Royle 2005; Iknayan *et al.* 2014). A great advantage of multispecies abundance N -mixture models, which to our knowledge currently exists only in the multiple-visit framework, is the ability to combine abundance and richness estimates while explicitly incorporating observation error (i.e. imperfect detection) in count data (Fig. 2, Table 1). However, this type of hierarchical model is relatively new, so its limitations and performance remain to be more widely evaluated. In addition, implementation of the models is still complex and beyond the programming expertise of many potential users, although example code and software are freely available on the Internet (Iknayan *et al.* 2014).

Conclusions

Population studies with the goals of informing conservation actions and management decisions often need to rigorously estimate abundance and how it is influenced by multiple factors. We believe this involves recognizing that, regardless of observer expertise, count data are subject to detection error and multiple forms of zero inflation (Fig. 1). Thus, adopting a sampling scheme and analytical method that account for these phenomena is necessary to address most questions that involve making inference or obtaining estimates about population size (Table 3). Of the 11 kinds of analytical methods used to estimate population size that we have reviewed (Table 1), 8 can account for both imperfect detection and additional variation in population size in the forms of non-occupancy, temporary emigration, correlated detection and population dynamics. All eight have been developed within the past decade, and their efficacy is still under study (Table 2). We expect developments to continue in this important area of statistical ecology.

Methods for estimating population size vary in their degree of suitability for a research project, depending on the taxa sampled, whether surveys are composed of a single or multiple species, the availability and feasibility of collecting ancillary data, and the scale and objectives of the study. Estimation methods can also impose differing financial and logistical costs, depending upon the requirement for repeated visits (Fig. 2, Table 1). Researchers should be aware of these differences when designing studies, so that limited funds and effort are spent efficiently.

Acknowledgements

FVD was supported by São Paulo Research Foundation (FAPESP) grants 2010/08528-6 and 2012/13195-1 and by The Peregrine Fund. SRB was supported by NSF grants CNH 1115069 and DEB-1051342. Reviews by Marc Kéry, Andy Royle, José L. Tella, the Beissinger Lab, the Seção de Aves do Museu de Zoologia da Universidade de São Paulo and two anonymous reviewers greatly improved this manuscript.

Data accessibility

No data was used in this review.

References

- Allredge, M.W., Simons, T.R. & Pollock, K.H. (2007b) Factors affecting aural detections of songbirds. *Ecological Applications*, **17**, 948–955.
- Allredge, M.W., Pollock, K.H., Simons, T.R., Collazo, J.A. & Shriner, S.A. (2007a) Time-of-detection method for estimating abundance from point-count surveys. *The Auk*, **124**, 653–664.
- Banks-Leite, C., Pardini, R., Boscolo, D., Cassano, C.R., Püttker, T., Barros, C.S. & Barlow, J. (2014) Assessing the utility of statistical adjustments for imperfect detection in tropical conservation science. *Journal of Applied Ecology*, **51**, 849–859.
- Bolker, B.M. (2008) *Ecological Models and Data in R*. Princeton University Press, Princeton and Oxford.
- Borchers, D.L. & Efford, M.G. (2008) Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics*, **64**, 377–385.
- Buckland, S.T. (2006) Point-transect surveys for songbirds: robust methodologies. *The Auk*, **123**, 345–357.
- Buckland, S.T., Marsden, S.J. & Green, R.E. (2008) Estimating bird abundance: making methods work. *Bird Conservation International*, **18**, 91–108.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford.
- Chandler, R.B. & Royle, J.A. (2013) Spatially explicit models for inference about density in unmarked or partially marked populations. *The Annals of Applied Statistics*, **7**, 936–954.
- Chandler, R.B., Royle, J.A. & King, D.I. (2011) Inference about density and temporary emigration in unmarked populations. *Ecology*, **92**, 1429–1435.
- Chandler, R.B., King, D.I., Raudales, R., Trubey, R., Chandler, C. & Chávez, V.J.A. (2013) A small-scale land-sparing approach to conserving biological diversity in tropical agricultural landscapes. *Conservation Biology*, **27**, 785–795.
- Chelgren, N.D., Samora, B., Adams, M.J. & McCreary, B. (2011) Using spatio-temporal models and distance sampling to map the space use and abundance of newly metamorphosed western toads (*Anaxyrus* spp.). *Herpetological Conservation and Biology*, **6**, 175–190.
- Conn, P.B., Laake, J.L. & Johnson, D.S. (2012) A hierarchical modeling framework for multiple observer transect surveys. *PLoS ONE*, **7**, e42294.
- Couturier, T., Cheylan, M., Bertolero, A., Astruc, G. & Besnard, A. (2013) Estimating abundance and population trends when detection is low and highly variable: a comparison of three methods for the Hermann's tortoise. *The Journal of Wildlife Management*, **77**, 454–462.
- Dail, D. & Madsen, L. (2011) Models for estimating abundance from repeated counts of an open metapopulation. *Biometrics*, **67**, 577–587.
- Dorazio, R.M. & Royle, J.A. (2005) Estimating size and composition of biological communities by modeling the occurrence of species. *Journal of the American Statistical Association*, **100**, 389–398.
- Dorazio, R.M., Royle, J.A., Söderström, B. & Glimskär, A. (2006) Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology*, **87**, 842–854.
- Doré, F., Grillet, P., Thirion, J., Besnard, A. & Cheylan, M. (2011) Implementation of a long-term monitoring program of the ocellated lizard (*Timon lepidus*) population on Oleron Island. *Amphibia-Reptilia*, **32**, 159–166.
- Efford, M. (2004) Density estimation in live-trapping studies. *Oikos*, **106**, 598–610.
- Fiske, I. & Chandler, R. (2011) Unmarked: an R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software*, **43**, 1–23.
- Gelman, A., Meng, X.L. & Stern, H.S. (1996) Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica*, **6**, 733–807.
- Hedley, S.L. & Buckland, S.T. (2004) Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics*, **9**, 181–199.
- Iknayan, K.J., Tingley, M.W., Furnas, B.J. & Beissinger, S.R. (2014) Detecting diversity: emerging methods to estimate species diversity. *Trends in Ecology & Evolution*, **29**, 97–106.
- Johnson, D.S., Laake, J.L. & Ver Hoef, J.M. (2010) A model-based approach for making ecological inference from distance sampling data. *Biometrics*, **66**, 310–318.
- Joseph, L.L.N., Elkin, C., Martin, T.G. & Possingham, H.P. (2009) Modeling abundance using N -mixture models: the importance of considering ecological mechanisms. *Ecological Applications*, **19**, 631–642.
- Katsanevakis, S. (2007) Density surface modelling with line transect sampling as a tool for abundance estimation of marine benthic species: the *Pinna nobilis* example in a marine lake. *Marine Biology*, **152**, 77–85.

- Kéry, M. (2008) Estimating abundance from bird counts: binomial mixture models uncover complex covariate relationships. *The Auk*, **125**, 336–345.
- Kéry, M., Royle, J. & Schmid, H. (2005) Modeling avian abundance from replicated counts using binomial mixture models. *Ecological Applications*, **15**, 1450–1461.
- Kéry, M. & Schaub, M. (2012) *Bayesian Population Analysis Using WinBUGS: A Hierarchical Perspective*. Academic Press, Waltham, Massachusetts.
- Lele, S.R., Moreno, M. & Bayne, E. (2012) Dealing with detection error in site occupancy surveys: what can we do with a single survey? *Journal of Plant Ecology*, **5**, 22–31.
- Lunn, D., Thomas, A., Best, N. & Spiegelhalter, D. (2000) WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, **10**, 325–337.
- Lunn, D., Spiegelhalter, D., Thomas, A. & Best, N. (2009) The BUGS project: evolution, critique and future directions. *Statistics in Medicine*, **28**, 3049–3067.
- MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew Royle, J. & Langtimm, C.A. (2002) Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, **83**, 2248–2255.
- Marques, F.F.C. & Buckland, S.T. (2003) Incorporating covariates into standard line transect analyses. *Biometrics*, **59**, 924–935.
- Marques, T.A., Thomas, L., Fancy, S.G. & Buckland, S.T. (2007) Improving estimates of bird density using multiple-covariate distance sampling. *The Auk*, **124**, 1229–1243.
- Marques, T.A., Buckland, S.T., Borchers, D.L., Tosh, D. & McDonald, R.A. (2010) Point transect sampling along linear features. *Biometrics*, **66**, 1247–1255.
- Martin, T.G., Wintle, B.A., Rhodes, J.R., Kuhnert, P.M., Field, S.A., Low-Choy, S.J., Tyre, A.J. & Possingham, H.P. (2005) Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters*, **8**, 1235–1246.
- Martin, J., Royle, J.A., Mackenzie, D.I., Edwards, H.H., Kéry, M. & Gardner, B. (2011) Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. *Methods in Ecology and Evolution*, **2**, 595–601.
- McCarthy, M.A., Moore, J.L., Morris, W.K., Parris, K.M., Garrard, G.E., Vesk, P.A. *et al.* (2013) The influence of abundance on detectability. *Oikos*, **122**, 717–726.
- McClintock, B.T., Bailey, L.L., Pollock, K.H. & Simons, T.R. (2010) Experimental investigation of observation error in anuran call surveys. *Journal of Wildlife Management*, **74**, 1882–1893.
- Miller, D.L., Burt, M.L., Rexstad, E.A. & Thomas, L. (2013a) Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution*, **4**, 1001–1010.
- Miller, D.A.W., Nichols, J.D., Gude, J.A., Rich, L.N., Podruzny, K.M., Hines, J.E. & Mitchell, M.S. (2013b) Determining occurrence dynamics when false positives occur: estimating the range dynamics of wolves from public survey data. *PLoS ONE*, **8**, e65808.
- Nelder, J.A. & Wedderburn, R.W.M. (1972) Generalized Linear Models. *Journal of the Royal Statistical Society, Series A*, **135**, 370–384.
- Niemi, A. & Fernández, C. (2010) Bayesian spatial point process modeling of line transect data. *Journal of Agricultural, Biological, and Environmental Statistics*, **15**, 327–345.
- Oedekoven, C.S., Buckland, S.T., Mackenzie, M.L., Evans, K.O. & Burger, L.W. (2013) Improving distance sampling: accounting for covariates and non-independency between sampled sites. *Journal of Applied Ecology*, **50**, 786–793.
- Plummer, M. (2003) JAGS: a program for analysis of bayesian graphical models using gibbs sampling. *Proceedings of the 3rd International Workshop in Distributed Statistical Computing (DSC 2003)*, March 20–22 (eds K. Hornik, F. Leisch & A. Zeileis), pp. 1–10, Technische Universität, Vienna, Austria.
- R Development Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. URL <http://www.R-project.org> [accessed 15 May 2014].
- Royle, J.A. (2004a) Generalized estimators of avian abundance from count survey data. *Animal Biodiversity and Conservation*, **27**, 375–386.
- Royle, J.A. (2004b) N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, **60**, 108–115.
- Royle, J.A., Dawson, D.K. & Bates, S. (2004) Modeling abundance effects in distance sampling. *Ecology*, **85**, 1591–1597.
- Royle, J.A. & Dorazio, R. (2008) *Hierarchical Modeling and Inference in Ecology*. Academic Press, London, UK.
- Royle, J.A. & Link, W.A. (2006) Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, **87**, 835–841.
- Royle, J.A. & Nichols, J.D. (2003) Estimating abundance from repeated presence-absence data or point counts. *Ecology*, **84**, 777–790.
- Sillett, T.S., Chandler, R.B., Royle, J.A., Kéry, M. & Morrison, S.A. (2012) Hierarchical distance-sampling models to estimate population size and habitat-specific abundance of an island endemic. *Ecological Applications*, **22**, 1997–2006.
- Sólymos, P., Lele, S. & Bayne, E. (2012) Conditional likelihood approach for analyzing single visit abundance survey data in the presence of zero inflation and detection error. *Environmetrics*, **23**, 197–205.
- Sólymos, P., Moreno, M. & Lele, S.R. (2013) “Detect”: analyzing wildlife data with detection error. R package version 0.3-2. URL <http://cran.r-project.org/package=detect> [accessed 13 March 2014].
- Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A. & Burnham, K.P. (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, **47**, 5–14.
- Tyre, A.J., Tenhumberg, B., Field, S.A., Niejalke, D., Parris, K. & Possingham, H.P. (2003) Improving precision and reducing bias in biological surveys: estimating false-negative error rates. *Ecological Applications*, **13**, 1790–1801.
- Warren, C.C., Veech, J.A., Weckerly, F.W., O'Donnell, L. & Ott, J.R. (2013) Detection heterogeneity and abundance estimation in populations of Golden-cheeked Warblers (*Setophaga chrysoparia*). *The Auk*, **130**, 677–688.
- Welsh, A.H., Cunningham, R.B., Donnelly, C.F. & Lindenmayer, D.B. (1996) Modelling the abundance of rare species: statistical models for counts with extra zeros. *Ecological Modelling*, **88**, 297–308.
- Wenger, S.J. & Freeman, M.M.C. (2008) Estimating species occurrence, abundance, and detection probability using zero-inflated distributions. *Ecology*, **89**, 2953–2959.
- Yamaura, Y., Andrew Royle, J., Kuboi, K., Tada, T., Ikeno, S. & Makino, S. (2011) Modelling community dynamics based on species-level abundance models from detection/nondetection data. *Journal of Applied Ecology*, **48**, 67–75.
- Yamaura, Y., Royle, J.A., Shimada, N., Asanuma, S., Sato, T., Taki, H. & Makino, S. (2012) Biodiversity of man-made open habitats in an underused country: a class of multispecies abundance models for count data. *Biodiversity and Conservation*, **21**, 1365–1380.
- Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A. & Smith, G.M. (2009) *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.
- Zylstra, E.R., Steidl, R.J. & Swann, D.E. (2010) Evaluating survey methods for monitoring a rare vertebrate, the Sonoran Desert Tortoise. *Journal of Wildlife Management*, **74**, 1311–1318.

Received 4 November 2014; accepted 17 December 2014
Handling Editor: Nick Isaac