

RESEARCH ARTICLE

Two-species occupancy modelling accounting for species misidentification and non-detection

Thierry Chambert^{1,2}  | Evan H. Campbell Grant³ | David A. W. Miller¹ |
James D. Nichols² | Kevin P. Mulder^{4,5}  | Adrienne B. Brand³

¹Department of Ecosystem Science and Management, Pennsylvania State University, University Park, PA, USA

²Patuxent Wildlife Research Center, United States Geological Survey, Laurel, MD, USA

³S.O. Conte Anadromous Fish Laboratory, Patuxent Wildlife Research Center, United States Geological Survey, Turners Falls, MA, USA

⁴Center for Conservation Genomics, Smithsonian Conservation Biology Institute, National Zoological Park, Washington, DC, USA

⁵Research Center in Biodiversity and Genetic Resources, CIBIO/InBIO, Vairão, Portugal

Correspondence

Thierry Chambert
Email: thierry.chambert@gmail.com

Handling Editor: Pedro Peres-Neto

Abstract

1. In occupancy studies, species misidentification can lead to false-positive detections, which can cause severe estimator biases. Currently, all models that account for false-positive errors only consider omnibus sources of false detections and are limited to single-species occupancy.
2. However, false detections for a given species often occur because of the misidentification with another, closely related species. To exploit this explicit source of false-positive detection error, we develop a two-species occupancy model that accounts for misidentifications between two species of interest. As with other false-positive models, identifiability is greatly improved by the availability of unambiguous detections at a subset of site \times occasions. Here, we consider the case where some of the field observations can be confirmed using laboratory or other independent identification methods ("confirmatory data").
3. We performed three simulation studies to (1) assess the model's performance under various realistic scenarios, (2) investigate the influence of the proportion of confirmatory data on estimator accuracy and (3) compare the performance of this two-species model with that of the single-species false-positive model. The model shows good performance under all scenarios, even when only small proportions of detections are confirmed (e.g. 5%). It also clearly outperforms the single-species model.
4. We illustrate application of this model using a 4-year dataset on two sympatric species of lungless salamanders: the US federally endangered Shenandoah salamander *Plethodon shenandoah*, and its presumed competitor, the red-backed salamander *Plethodon cinereus*. Occupancy of red-backed salamanders appeared very stable across the 4 years of study, whereas the Shenandoah salamander displayed substantial turnover in occupancy of forest habitats among years.
5. Given the extent of species misidentification issues in occupancy studies, this modelling approach should help improve the reliability of estimates of species distribution, which is the goal of many studies and monitoring programmes. Further developments, to account for different forms of state uncertainty, can be readily undertaken under our general approach.

KEYWORDS

false positive, observation error, *Plethodon cinereus*, *Plethodon shenandoah*, site colonization, site extinction, species distribution modelling, species misidentification

1 | INTRODUCTION

Estimation of species occurrence is central to many conservation efforts and scientific endeavours (Ferrier, 2002; Gaston, 2009; MacKenzie et al., 2006; Rushton, Ormerod, & Kerby, 2004). Indeed, interactions with sympatric species can be important drivers of a species' distribution dynamic (Wisz et al., 2013). For instance, competition between two species occupying similar niches might limit the range distribution, by limiting expansion or triggering range contraction, of one or both species (Sexton, McIntyre, Angert, & Rice, 2009). Reliable estimates of species occurrence and co-occurrence require accounting for observational errors, such as the imperfect detection of animals in the wild (MacKenzie et al., 2002; Tyre et al., 2003). When dealing with two species, accounting for detection is even more important because observation uncertainty exists for both species, which might introduce stronger biases on estimates of occurrence probabilities as well as on the species interaction parameter estimates (MacKenzie, Bailey, & Nichols, 2004). Occupancy models were explicitly developed to account for observation errors when estimating species occurrence. Generally, these estimators focus on only false absences, which occur when the species is present but not detected. Since the seminal paper of MacKenzie et al. (2002), occupancy methods have been extended to incorporate temporal dynamics (MacKenzie, Nichols, Hines, Knutson, & Franklin, 2003), species co-occurrence (MacKenzie et al., 2004), multistate variables (MacKenzie, Bailey, Hines, & Nichols, 2011; MacKenzie, Nichols, Seamans, & Gutiérrez, 2009) and community and regional patterns of occurrence (Dorazio, Royle, Söderström, & Glimskär, 2006; Miller & Grant, 2015).

While the set of tools available to account for false absences has been well developed, much less effort has gone into modelling false-positive detections. False positives occur when a species is recorded as detected despite not being present at a site. The first attempt to deal with false positives within an occupancy framework was developed by Royle and Link (2006). They developed the basic estimator and provided a detailed outline of the challenges encountered when both false-negative and false-positive detection errors occur. Subsequent developments have focused on methods to increase identify ability of these two sources of observational heterogeneity. Examples include combining data with a second unambiguous source (Chambert, Miller, & Nichols, 2015; Miller et al., 2011, 2013), validating observations post hoc using a second method (e.g. DNA for scat samples), including multiseason dynamics in the model (Miller et al., 2013, 2015), spatial autocorrelation (Pacifi et al., 2017) and including information about the relative intensity of detections (Miller et al., 2011). In all these cases, including additional information improves the ability to generate accurate and unbiased estimates of species occurrence (Miller et al., 2015).

One unexplored source of information for identifying when false-positive errors are likely to occur comes from the process by which misidentifications usually occur. In most studies, a false-positive detection occurs because one species is misidentified as a second species. When these misidentification errors occur between only

two species, one can surmise that the probability of a false-positive detection for a species A is zero when a species B does not occur at the site and will be non-zero only when species B does occur. Thus, the distribution of the second species should help define where misidentification does and does not occur. In addition, as mentioned above, failing to account for confusion in the identity of co-occurring species can lead to bias in estimating single-species distributions as well as patterns of co-occurrence and species interaction. However, the distribution of species B is itself an unknown parameter and accounting for the second species requires simultaneously estimating occupancy of both species.

A robust set of estimators has been developed for estimating patterns of occurrence and co-occurrence using two-species models (MacKenzie et al., 2004; Richmond, Hines, & Beissinger, 2010; Waddle et al., 2010). These methods have been extended to incorporate multiseason dynamics (Yackulic et al., 2014) and both habitat dynamics and species interactions (Davis et al., 2017; Miller, Brehme, Hines, Nichols, & Fisher, 2012). However, these methods all assume that the two species are not misidentified, hence ignoring the potential for false-positive detections.

In this paper, we develop a general approach for modelling false-positive errors when error rate for one species is dependent on the presence of another species. We provide a general parameterization for dealing with misidentification errors and explore if and how the model reduces bias and improves accuracy using both simulated and real-world datasets.

2 | MODEL DESCRIPTION

This model estimates site occupancy probability for two species, from detection/non-detection data, when misidentification between the two species of interest can occur. First, we consider the case of a single season occupancy study, and then describe the multiseason extension, which models colonization and extinction dynamics, for each species, between consecutive seasons (MacKenzie et al., 2003).

2.1 | Single season occupancy model

The sampling frame for data collection is as follows. In a single season, a selected number of sites $i = \{1, \dots, I\}$ is surveyed, repeatedly, J_i times. Individual surveys (i.e. sampling occasions) are denoted $j = \{1, \dots, J_i\}$. For the sake of simplicity, in what follows, we will consider an equal number of sampling occasions for all sites, such that $J_i = J$ for any site i . At each site x occasion $\{i, j\}$, field observers record whether they detected either of the two species of interest, which we will refer to as species A and species B. Field observation (y_{ij}) can thus take four values (hereafter, "observation states"): "O" if neither of the species was detected; "A" if only species A was detected; "B" if only species B was detected; "AB" if both species were detected. Similarly, we model the true, unknown occupancy status (z_i) of any site i as a multistate process

combining the presence/absence of both species, which leads to four possible values (hereafter, "occupancy states"): "0" if neither of the species is present; "A" if only species A is present; "B" if only species B is present; "AB" if both species are present. As with most occupancy models, we assume closure, which means that the occupancy state of a site remains constant across all sampling occasions.

We consider two different model formulations for the species' probabilities of occupancy and co-occurrence. The first model formulation, which follows MacKenzie et al. (2004), considers a *general* interaction between the two species. The probability of occupancy of each species is defined unconditionally on the other species' presence or absence, and the probability of co-occurrence is modelled through a common interaction term (γ). Under this formulation (hereafter, the *general interaction*), we define ψ_i^A and ψ_i^B as the unconditional probabilities of occupancy of species A and B, respectively, at site i . The probability of co-occurrence is defined as $\psi_i^{AB} = \psi_i^A \times \psi_i^B \times \gamma$, where γ is the interaction term, sometimes referred to as the species interaction factor (SIF; MacKenzie et al., 2004). When $\gamma = 1$, the two species occur totally independently of each other. On the other hand, $\gamma < 1$ indicates that the two species tend to avoid each other (e.g. competition), while $\gamma > 1$ indicates that they favour co-occurrence (e.g. similar habitat preferences). Under this *general interaction* model formulation, the vector of probabilities (Φ) describing the four mutually exclusive occupancy states (0, A, B and AB respectively) is:

$$\Phi = [1 - \psi^A - \psi^B + \psi^{AB} \quad \psi^A - \psi^{AB} \quad \psi^B - \psi^{AB} \quad \psi^{AB}].$$

Note that we removed the site subscripts (i) here for ease of reading, but this vector of probabilities could be made site-specific (Φ_i).

The second model formulation, which follows Richmond et al. (2010), considers that one species is dominant over the other. Under this formulation (hereafter referred to as the *dominance interaction*), occupancy of species A (the dominant species) is independent of the presence or absence of species B (the subordinate species). The unconditional probability of site occupancy of species A is defined as ψ_i^A . Occupancy of species B is conditional on whether or not species A is present at the site of interest. We define $\psi_i^{B|0}$ as the probability of occupancy of species B at sites where species A is not present, and $\psi_i^{B|A}$ as the probability of occupancy of species B at sites where species A is present. To be consistent with the notation used in the general interaction model, we redefine $\psi^{B|A} = \psi^{B|a} * \gamma$, where γ can be interpreted as the effect of presence of species A on the probability of site occupancy of species B. Here, $\gamma = 1$ indicates no effect, $\gamma < 1$ indicates a negative effect, such that species B is less likely to be present when species A occupies the site (e.g. one-way competitive effect), while $\gamma > 1$ indicates a positive effect of species A on the probability of occurrence of species B. Under this dominance interaction formulation, the multistate occupancy probability vector Φ is (here again, not including site-specific subscripts i):

$$\Phi = [(1 - \psi^A)(1 - \psi^{B|a}) \quad \psi^A(1 - \psi^{B|A}) \quad (1 - \psi^A)\psi^{B|a} \quad \psi^A\psi^{B|A}].$$

For both formulations, the two-species occupancy state process is then modelled as a multinomial process:

$$z_i \sim \text{Multinomial}(\Phi).$$

The observation process that links latent occupancy states (z_i) to the multistate data (y_{it}) consists of two types of processes: (1) the detection of a species (A and/or B) given that it is truly present at a site and (2) the false detection of a species (A or B) given that it is not present at a site, which, here, can only result from the misidentification of some specimen from the other species. It is this latter constraint on the process of misidentification that distinguishes this new class of models from existing false-positive models (Chambert et al., 2015; Miller et al., 2011, 2013). To model the observation state process, we define the following sets of parameters: (1) $p^A[p^B]$ is the probability of correctly detecting species A [B] at a site where only species A [B] is present; (2) $r^A[r^B]$ is the probability of correctly detecting species A [B] at a site where both species are present; (3) ω^A and ω^a [ω^B and ω^b] are the probabilities of erroneously detecting species B [A] at a site where only the other species, A [B], is present (i.e. some specimens of species A [B] are misidentified as species B [A]). We refer to these (ω) parameters as probabilities of species misidentification, and the distinction between ω^A and ω^a [ω^B and ω^b] is defined as follows. Uppercased parameter ω^A [ω^B] represents the probability that some specimens of species A [B] are misidentified as being species B [A] when species A [B] has also been correctly detected at that site x occasion. Lowercased parameter ω^a [ω^b] is the probability that some specimens of species A [B] are misidentified as being species B [A] when species A [B] has not been correctly detected at that site x occasion. This distinction allows us to account for the non-independence that might exist between true detection and misidentification. For instance, when several animals are caught and held next to each other, if some are easily identified as belonging to, let's say, species A, this might influence the identification of the remaining specimens, for which identity might be less clear-cut. In the analyses presented in this paper, we consider two cases: (1) the case where misidentification and correct detection are independent (i.e. $\omega^a = \omega^A$ and $\omega^b = \omega^B$) and (2) the case where the probability of misidentification is influenced, positively or negatively, by the occurrence of correct detections ($\omega^a \neq \omega^A$ and $\omega^b \neq \omega^B$).

The probabilities of field observation states ($y_{ij} = \{0, A, B, AB\}$), conditional on the true occupancy state z_i of a site, are defined by the probability matrix Π shown in Table 1. Field data are thus modelled as:

$$y_{it}|z_i \sim \text{Multinomial}(\Pi_{z_i}),$$

where Π_{z_i} is the probability vector corresponding to a specific occupancy state z_i (i.e. a row of matrix Π ; Table 1)

As one can see from matrix Π (Table 1), the three non-null states ($z_i = \{A, B, AB\}$) cannot be strictly disentangled from field observations only, such that none of the states can be unambiguously determined. Indeed, whenever we observe A, B or AB, any of these three states is possible. Without any additional assumption, this creates issues of

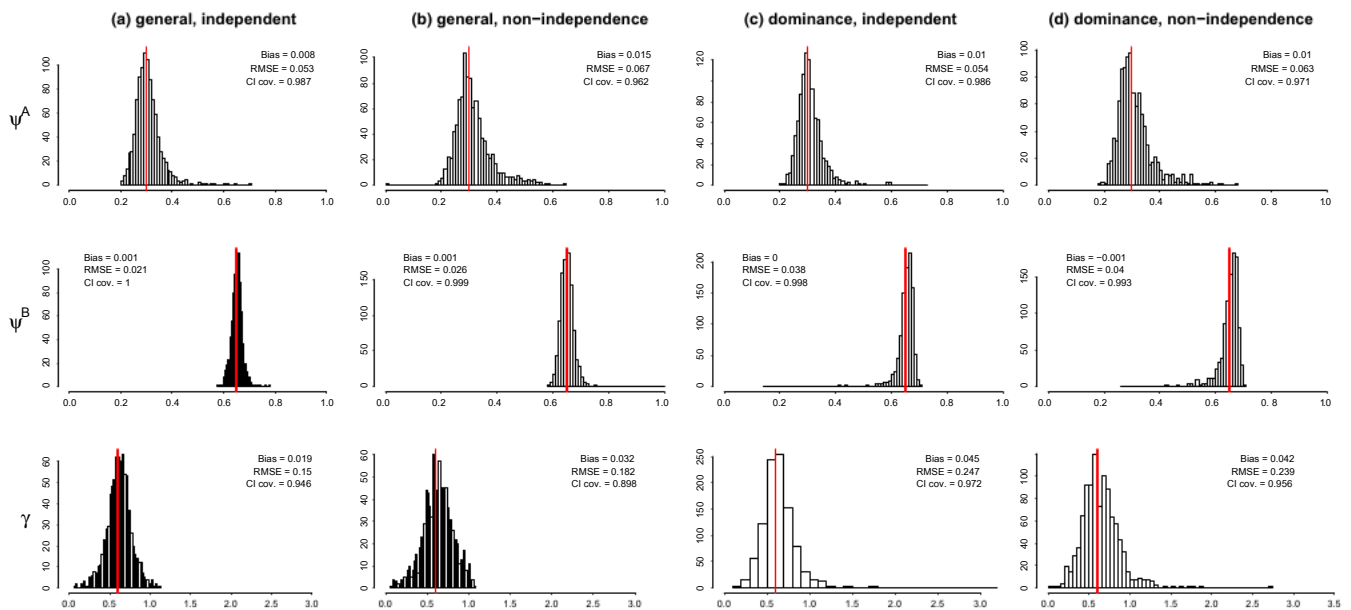


FIGURE 1 Summarized results of the first simulation study, which focused on four different model formulations: (a) general, independence model; (b) general, non-independence model; (c) dominance, independence model; (d) dominance, non-independence model. Estimation accuracy of the three occupancy parameters is shown: ψ^A (top row), ψ^B (middle row) and γ (bottom row). The estimator bias, root-mean-squared error (rmse) and the coverage of 95% CI's (CI cov.) is shown on each individual plot. The proportion of site x occasion for which some detections were confirmed was 30%. The red line represents the true parameter value used to generate the data

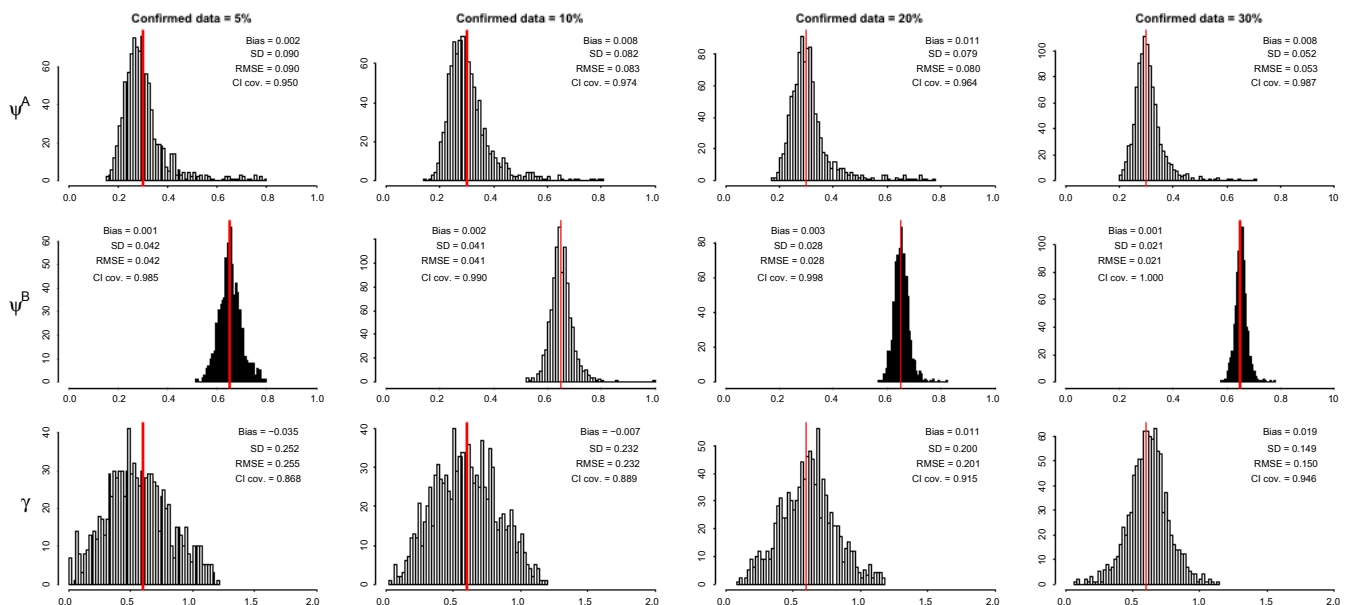


FIGURE 2 Results of the second simulation study. Model performance was assessed for the following proportions of confirmatory data: 5%, 10%, 20% and 30% (left to right columns on the graph). Data were simulated and analysed using the “general interaction, independence misidentification” model. Estimation accuracy of the three occupancy parameters is shown: ψ^A (top row), ψ^B (middle row) and γ (bottom row). The estimator bias, standard deviation (SD), root-mean-squared error (rmse) and the coverage of 95% CI's (CI cov.) is shown on each individual plot. The red line represents the true parameter value used to generate the data

they were virtually identical to those of the model with 15 parameters (Table 4), which reinforces our confidence that the issue with this latter model only concerned standard errors (and not point estimates of the parameters). As expected, the probability of occupancy of red-backed salamanders was much higher (almost threefold) than

that of the Shenandoah salamander ($\hat{\psi}^C = 0.41$ [0.30–0.53], $\hat{\psi}^S = 0.15$ [0.08–0.27], Table 4). The species interaction factor for 2011 was very close to 1.0 ($\hat{\gamma} = 0.98$ [0.50–1.95]) suggesting that both species occurred independently of each other. We interpret this latter result with caution because sampled sites cover an area much larger than

type of observation collected during the same or different surveys as the ambiguous observations. For instance, a large volume of ambiguous data might be provided through auditory detections, while infrequent visual observations provide fewer, but unambiguous detections (Miller et al., 2011). Similarly, in large predator surveys (Molinari-Jobin et al., 2012), it is common to collect any clue available, some of which are highly uncertain (pug marks, scats) while others are totally unambiguous (reports of a dead animal found, visual observation of the animal by an expert). On the other hand, they can be provided by the a posteriori validation of ambiguous detections, for instance, through the analysis of genetic material of captured animals, as in our salamander example. Other examples include the validation, under the microscope, of specimens collected in the field (insects, larvae), validation by an expert of a subset of acoustic detections provided by an automated call-recognition programme (Chambert et al., 2017, in press) and post hoc analysis of specimen photographs (Conn et al., 2013; McClintock et al., 2015).

Occupancy approaches have become a method of choice for many wildlife studies and monitoring programmes (Bailey, MacKenzie, & Nichols, 2014). Many of the assumptions and limitations of the seminal occupancy models (MacKenzie et al., 2002, 2006) have already been relaxed (Bailey et al., 2014). Species misidentification can be a major impairment to the success of such studies, because it is likely to cause large biases (McClintock, Bailey, Pollock, & Simons, 2010; Miller et al., 2011, 2015). The model described here, along with false-positive models previously developed (Chambert et al., 2015; Miller et al., 2011, 2013; Ruiz-Gutierrez et al., 2016), provide an important step to making occupancy approaches even more useful and reliable. Although such an approach can be data hungry and might require substantial field effort, it should still involve much less effort than what is typically required for individual-based monitoring, such as mark-recapture methods. Finally, it is important to note that this approach, which falls under a Hidden Markov Model framework, can easily be generalized to deal with any kind of state uncertainty, such as breeding or species residency states (Pradel, 2005). This model will likely be extended to further expand the occupancy toolkit available to wildlife ecologists.

ACKNOWLEDGEMENTS

Funding for this research was primarily provided by the U.S. Geological Survey–Amphibian Research and Monitoring Initiative (ARMI), and the National Park Service. We are very thankful to Richard Chandler for his help and early contribution to the development of the method. We are also grateful to all the field technicians who helped collecting the salamander data. This is ARMI contribution number 618.

AUTHORS' CONTRIBUTIONS

T.C., D.M. and J.N. conceived and developed the methodology; E.G. conceived the field sampling protocols and collected the data; K.M. and A.B. generated and analysed the genetic data; T.C. analysed the

final data; T.C., E.G., J.N. and D.M. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Data used in this paper: detection/non-detection data, for the two salamander species: *Plethodon cinereus* and *Plethodon shenandoah*. These data are available in a Dryad Digital Repository <https://doi.org/10.5061/dryad.88ks4g5> (Chambert et al., 2018).

ORCID

Thierry Chambert  <http://orcid.org/0000-0002-9450-9080>

Kevin P. Mulder  <http://orcid.org/0000-0001-6688-8848>

REFERENCES

- Bailey, L. L., MacKenzie, D. I., & Nichols, J. D. (2014). Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5, 1269–1279. <https://doi.org/10.1111/2041-210X.12100>
- Chambert, T., Grant, E. H. C., Miller, D. A. W., Nichols, J. D., Mulder, K. P., & Brand, A. B. (2018). Data from: Two-species occupancy modeling accounting for species misidentification and nondetection. *Dryad Digital Repository*, <https://doi.org/10.5061/dryad.88ks4g5>
- Chambert, T., Miller, D. A. W., & Nichols, J. D. (2015). Modeling false positive detections in species occurrence data under different study designs. *Ecology*, 96, 332–339. <https://doi.org/10.1890/14-1507.1>
- Chambert, T., Waddle, J. H., Miller, D. A. W., Walls, S. C., & Nichols, J. D. (2017). A new framework for analysing automated acoustic species-detection data: Occupancy estimation and optimization of recordings post-processing. *Methods in Ecology and Evolution*, <https://doi.org/10.1111/2041-210X.12910> (in press).
- Clement, M. J. (2016). Designing occupancy studies when false-positive detections occur. *Methods in Ecology and Evolution*, 7, 1538–1547. <https://doi.org/10.1111/2041-210X.12617>
- Conn, P. B., McClintock, B. T., Cameron, M. F., Johnson, D. S., Moreland, E. E., & Boveng, P. L. (2013). Accommodating species identification errors in transect surveys. *Ecology*, 94, 2607–2618. <https://doi.org/10.1890/12-2124.1>
- Davis, C. L., Miller, D. A. W., Walls, S. C., Barichivich, W. J., Riley, J. W., & Brown, M. E. (2017). Species interactions and the effects of climate variability on a wetland amphibian metacommunity. *Ecological Applications*, 27, 285–296. <https://doi.org/10.1002/eap.1442>
- 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. [https://doi.org/10.1890/0012-9658\(2006\)87\[842:ESRAAB\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[842:ESRAAB]2.0.CO;2)
- Ferrier, S. (2002). Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here? *Systematic Biology*, 51, 331–363. <https://doi.org/10.1080/10635150252899806>
- Gaston, K. J. (2009). Geographic range limits: achieving synthesis. *Proceedings of the Royal Society of London B: Biological Sciences*, 276, 1395–406.
- Jaeger, R. G. (1971). Competitive exclusion as a factor influencing the distributions of two species of terrestrial salamanders. *Ecology*, 52, 632–637. <https://doi.org/10.2307/1934151>
- MacKenzie, D. I., Bailey, L. L., Hines, J. E., & Nichols, J. D. (2011). An integrated model of habitat and species occurrence dynamics. *Methods in Ecology and Evolution*, 2, 612–622. <https://doi.org/10.1111/j.2041-210X.2011.00110.x>

- MacKenzie, D. I., Bailey, L. L., & Nichols, J. D. (2004). Investigating species co-occurrence patterns when species are detected imperfectly. *Journal of Animal Ecology*, 73, 546–555. <https://doi.org/10.1111/j.0021-8790.2004.00828.x>
- MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., & Franklin, A. B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, 84, 2200–2207. <https://doi.org/10.1890/02-3090>
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83, 2248–2255. [https://doi.org/10.1890/0012-9658\(2002\)083\[2248:ESORWD\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2)
- MacKenzie, D., Nichols, J., Royle, J., Pollock, K., Bailey, L., & Hines, J. (2006). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence*. San Diego, CA: Elsevier.
- MacKenzie, D. I., Nichols, J. D., Seamans, M. E., & Gutiérrez, R. J. (2009). Modeling species occurrence dynamics with multiple states and imperfect detection. *Ecology*, 90, 823–835. <https://doi.org/10.1890/08-0141.1>
- McClintock, B. T., Bailey, L. L., Pollock, K. H., & Simons, T. R. (2010). Unmodeled observation error induces bias when inferring patterns and dynamics of species occurrence via aural detections. *Ecology*, 91, 2446–2454. <https://doi.org/10.1890/09-1287.1>
- McClintock, B. T., Moreland, E. E., London, J. M., Dahle, S. P., Brady, G. M., Richmond, E. L., ... Boveng, P. L. (2015). Quantitative assessment of species identification in aerial transect surveys for ice-associated seals. *Marine Mammal Science*, 31, 1057–1076. <https://doi.org/10.1111/mms.12206>
- Miller, D. A. W., Bailey, L. L., Campbell Grant, E. H., McClintock, B. T., Weir, L. A., & Simons, T. R. (2015). Performance of species occurrence estimators when basic assumptions are not met: A test using field data where true occupancy status is known. *Methods in Ecology and Evolution*, 6, 557–565. <https://doi.org/10.1111/2041-210X.12342>
- Miller, D. A. W., Brehme, C. S., Hines, J. E., Nichols, J. D., & Fisher, R. N. (2012). Joint estimation of habitat dynamics and species interactions: Disturbance reduces co-occurrence of non-native predators with an endangered toad. *The Journal of Animal Ecology*, 81, 1288–1297. <https://doi.org/10.1111/j.1365-2656.2012.02001.x>
- Miller, D. A. W., & Grant, E. H. C. (2015). Estimating occupancy dynamics for large-scale monitoring networks: Amphibian breeding occupancy across protected areas in the northeast United States. *Ecology and Evolution*, 5, 4735–4746. <https://doi.org/10.1002/ece3.1679>
- Miller, D. A. W., Nichols, J. D., Gude, J. A., Rich, L. N., Podrutzny, K. M., Hines, J. E., & Mitchell, M. S. (2013). Determining occurrence dynamics when false positives occur: Estimating the range dynamics of wolves from public survey data. *PLoS ONE*, 8, e65808. <https://doi.org/10.1371/journal.pone.0065808>
- Miller, D. A. W., Nichols, J. D., McClintock, B. T., Grant, E. H. C., Bailey, L. L., & Weir, L. A. (2011). Improving occupancy estimation when two types of observational error occur: Non-detection and species misidentification. *Ecology*, 92, 1422–1428. <https://doi.org/10.1890/10-1396.1>
- Miller, D. A. W., Weir, L. A., McClintock, B. T., Grant, E. H. C., Bailey, L. L., & Simons, T. R. (2012). Experimental investigation of false positive errors in auditory species occurrence surveys. *Ecological Applications*, 22, 1665–1674. <https://doi.org/10.1890/11-2129.1>
- Molinari-Jobin, A., Kéry, M., Marboutin, E., Molinari, P., Koren, I., Fuxjäger, C., ... Breitenmoser, U. (2012). Monitoring in the presence of species misidentification: The case of the Eurasian lynx in the Alps. *Animal Conservation*, 15, 266–273. <https://doi.org/10.1111/j.1469-1795.2011.00511.x>
- Pacifici, K., Reich, B. J., Miller, D. A. W., Gardner, B., Stauffer, G., Singh, S., ... Collazo, J. A. (2017). Integrating multiple data sources in species distribution modeling: A framework for data fusion*. *Ecology*, 98, 840–850. <https://doi.org/10.1002/ecy.1710>
- Pillay, R., Miller, D. A. W., Hines, J. E., Joshi, A. A., & Madhusudan, M. D. (2014). Accounting for false positives improves estimates of occupancy from key informant interviews. *Diversity and Distributions*, 20, 223–235. <https://doi.org/10.1111/ddi.12151>
- Pradel, R. (2005). Multievent: An extension of multistate capture-recapture models to uncertain states. *Biometrics*, 61, 442–447. <https://doi.org/10.1111/j.1541-0420.2005.00318.x>
- Richmond, O. M. W., Hines, J. E., & Beissinger, S. R. (2010). Two-species occupancy models: A new parameterization applied to co-occurrence of secretive rails. *Ecological Applications*, 20, 2036–2046. <https://doi.org/10.1890/09-0470.1>
- Royle, J. A., & Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, 87, 835–841. [https://doi.org/10.1890/0012-9658\(2006\)87\[835:GSOMAF\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[835:GSOMAF]2.0.CO;2)
- Ruiz-Gutierrez, V., Hooten, M. B., & Grant, E. H. C. (2016). Uncertainty in biological monitoring: A framework for data collection and analysis to account for multiple sources of sampling bias. *Methods in Ecology and Evolution*, 7, 900–909. <https://doi.org/10.1111/2041-210X.12542>
- Rushton, S. P., Ormerod, S. J., & Kerby, G. (2004). New paradigms for modelling species distributions? *Journal of Applied Ecology*, 41, 193–200. <https://doi.org/10.1111/j.0021-8901.2004.00903.x>
- Sexton, J. P., McIntyre, P. J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, 40, 415–436. <https://doi.org/10.1146/annurev.ecolsys.110308.120317>
- 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. <https://doi.org/10.1890/02-5078>
- Waddle, J. H., Dorazio, R. M., Walls, S. C., Rice, K. G., Beauchamp, J., Schuman, M. J., & Mazzotti, F. J. (2010). A new parameterization for estimating co-occurrence of interacting species. *Ecological Applications*, 20, 1467–1475. <https://doi.org/10.1890/09-0850.1>
- Wis, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., ... Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, 88, 15–30. <https://doi.org/10.1111/j.1469-185X.2012.00235.x>
- Yackulic, C. B., Reid, J., Nichols, J. D., Hines, J. E., Davis, R., & Forsman, E. (2014). The roles of competition and habitat in the dynamics of populations and species distributions. *Ecology*, 95, 265–279. <https://doi.org/10.1890/13-0012.1>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Chambert T, Grant EHC, Miller DAW, Nichols JD, Mulder KP, Brand AB. Two-species occupancy modelling accounting for species misidentification and non-detection. *Methods Ecol Evol*. 2018;9:1468–1477. <https://doi.org/10.1111/2041-210X.12985>