Received: 1 September 2017 | Revised: 19 January 2018 | Accepted: 22 January 2018

DOI: 10.1111/1755-0998.12765

RESOURCE ARTICLE

WILEY MOLECULAR ECOLOGY RESOURCES

A model-based solution for observational errors in laboratory studies

Brian M. Brost[1] | Brittany A. Mosher[2] | Kristen A. Davenport[3]

[1]Marine Mammal Laboratory, Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, WA, USA

[2]Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO, USA

[3]Department of Microbiology, Immunology, and Pathology, Colorado State University, Fort Collins, CO, USA

Correspondence
Brian M. Brost, Marine Mammal Laboratory, Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, WA, USA.
Email: brian.brost@noaa.gov

Abstract

Molecular techniques for detecting microorganisms, macroorganisms and infectious agents are susceptible to false-negative and false-positive errors. If left unaddressed, these observational errors may yield misleading inference concerning occurrence, prevalence, sensitivity, specificity and covariate relationships. Occupancy models are widely used to account for false-negative errors and more recently have even been used to address false-positive errors, too. Current modelling options assume false-positive errors only occur in truly negative samples, an assumption that yields biased inference concerning detection because a positive sample could be classified as such not because the target agent was successfully detected, but rather due to a false-positive test result. We present an extension to the occupancy modelling framework that allows false-positive errors in both negative and positive samples, thereby providing unbiased inference concerning occurrence and detection, as well as reliable conclusions about the efficacy of sampling designs, handling protocols and diagnostic tests. We apply the model to simulated data, showing that it recovers known parameters and outperforms other approaches that are commonly used when confronted with observation errors. We then apply the model to an experimental data set on *Batrachochytrium dendrobatidis*, a pathogenic fungus that is implicated in the global decline or extinction of hundreds of amphibian species. The model-based approach we present is not only useful for obtaining reliable inference when data are contaminated with observational errors, but also eliminates the need for establishing arbitrary thresholds or decision rules that have hidden and unintended consequences.

KEYWORDS

eDNA, false negative, false positive, imperfect detection, occupancy, sensitivity, specificity

1 | INTRODUCTION

Molecular techniques are central to answering ecological questions concerning species distributions, diet composition, disease dynamics and exposing invasive, rare or cryptic species. Procedures for detecting microorganisms, macroorganisms and infectious agents are susceptible to two sources of observation error, namely false negatives and false positives. False-negative errors occur when the target agent (e.g., pathogen) is present in a sample but not detected (i.e., imperfect sensitivity), an event that may be attributable to sample variation (e.g., low quantity available to be detected), the presence of substances that inhibit molecular assays, or sample degradation (McClintock, Nichols et al., 2010; Schrader, Schielke, Ellerbroek, & Johne, 2012). Conversely, false positives are erroneous detections that may occur when nonspecific binding or cross-reactivity of reagents leads to spurious amplification. Contamination is another

frequent cause of false-positive errors, a source that may be unavoidable despite sound practices (McClintock, Nichols et al., 2010).

Ignoring false-negative and false-positive errors, or addressing them in an ad hoc manner, is common practice in the analysis of molecular data; however, failure to rigorously account for observation errors may yield misleading inference. For example, false-negative (positive) errors introduce negative (positive) biases in estimates of occurrence, prevalence, sensitivity and specificity and are even more problematic when interest lies in relationships between these entities and covariates (MacKenzie et al., 2002; Miller et al., 2011; Royle & Link, 2006). Indeed, even low error rates that are often considered innocuous (e.g., 1% of observations are errors) can lead to spurious inferences that impede mitigation efforts, reduce the efficacy of surveillance programmes or cause errant diagnoses (Lahoz-Monfort, Guillera-Arroita, & Tingley, 2016; McClintock, Bailey, Pollock, & Simons, 2010; Royle & Link, 2006; Tyre et al., 2003).

Bias caused by false-negative errors can be eliminated using a class of binomial mixture models that were originally developed to examine patterns of species occurrence when individuals are detected imperfectly (imagine detecting animals that are hidden in thick vegetation or that can be detected only when vocalizing). These "occupancy" models rely on replicate surveys to estimate the true but partially unobserved (i.e., latent) occurrence of the species while controlling for the detrimental effect of false-negative errors (MacKenzie et al., 2002; Tyre et al., 2003). Because laboratory protocols often include multiple replicates on the same sample (e.g., molecular assays performed in triplicate), occupancy models have also been used to account for false-negative errors in laboratory studies (Davenport et al., 2018; Lachish, Gopalaswamy, Knowles, & Sheldon, 2012; McClintock, Nichols et al., 2010; Mosher et al., 2017; Schmidt, Kéry, Ursenbacher, Hyman, & Collins, 2013).

To illustrate the general occupancy modelling framework, suppose a single sample is tested three times for a pathogen (i.e., three replicates). If the only potential observation error is a false negative, a 001 outcome indicates that the pathogen is present in the sample (a "1" indicates a positive test result) but it went undetected in two of three tests (indicated by a "0"). A 000 outcome, on the other hand, is ambiguous because two possibilities can explain the event: (i) the test results were accurate and the pathogen was actually absent from the sample, or (ii) the sample was in fact positive but the pathogen simply went undetected (a "false negative"). The underlying premise of occupancy modelling is that repeat tests and unambiguous state assignments (e.g., 001) are useful not only for estimating the probability of detection, but also for informing the true occurrence state of the ambiguous samples (i.e., 000) in a probabilistic and statistically valid manner.

In the presence of false-positive errors, samples that include positive replicates (e.g., 001) are no longer unambiguous, undermining a central assumption of the classical occupancy model (MacKenzie et al., 2002; Tyre et al., 2003). Royle and Link (2006) were the first to extend the occupancy modelling framework to accommodate false-positive errors, although limitations to their approach (e.g., unidentifiable parameters) have impeded its successful implementation. Fortunately, modifications that use alternative sampling designs and ancillary information can circumvent these problems. For example, Miller et al. (2011) took advantage of "certain" detections (i.e., methods that yield true positives only) for a subset of samples to facilitate parameter estimation. Trials on reference samples of known state have also been used to help inform rates of observational errors (Chambert, Miller, & Nichols, 2015; McClintock, Nichols et al., 2010).

Existing false-positive occupancy models only allow for imperfect specificity, or false-positive errors that occur when the target agent is not present but nevertheless "detected" (Chambert et al., 2015; McClintock, Nichols et al., 2010; Royle & Link, 2006); however, the mechanisms that yield spurious detections in negative samples (e.g., spurious amplification, contamination) also operate in samples that are truly positive. Accordingly, false-positive errors can also arise when the agent is present, and a positive sample could thus be correctly classified simply by mistake. Even though classifications that are correct but made in error do not affect statistical inference concerning the occurrence of the pathogen, these errors do result in biased inference concerning its detectability (Royle & Link, 2006). Consequently, reliable conclusions about detection or the efficacy of different sampling designs, handling protocols or diagnostic tests cannot be drawn from current modelling approaches when data are contaminated with false-positive errors.

We present a model-based approach that allows false-positive errors in both negative and positive samples, thereby providing unbiased inference concerning the occurrence and detection processes. The modelling framework is general and broadly applicable because it only requires repeat sampling and trials on negative controls (i.e., known negative samples), data that are already collected in many diagnostic protocols (e.g., Carey et al., 2006; Davenport et al., 2018; Lachish et al., 2012; Mosher et al., 2017). We first describe occupancy models that address imperfect sensitivity and specificity, which we subsequently extend to fully accommodate false-positive errors. We conduct a simulation study not only to evaluate the performance of the extended model, but also to compare it to alternative approaches used when confronted with observational errors. Lastly, we include an illustrative example using *Batrachochytrium dendrobatidis*, a pathogenic fungus that is implicated in the global decline or extinction of hundreds of amphibian species (Skerratt et al., 2007).

2 | MODEL FORMULATION

Separating detectability from the true occurrence process first requires differentiating between observations and the underlying latent occupancy (or occurrence) state. We denote the observations by y_{ij} for $i = 1, \ldots, N$ samples and $j = 1, \ldots, J_i$ replicates, where $y_{ij} = 1$ represents a positive test result and $y_{ij} = 0$ represents a negative result. We denote the latent state by z_i, where $z_i = 1$ indicates sample i is truly positive (i.e., occupied) and $z_i = 0$ indicates it is actually negative (i.e., unoccupied). If occurrence was observed

without error, the values for y_{ij} and z_i would be identical; however, in the presence of observation errors, y_{ij} is only sometimes equal to z_i. For example, the event $\{y_{ij} = 0, z_i = 1\}$ characterizes a false-negative error.

Occupancy models have a natural hierarchical formulation consisting of two basic components, an "observation" model and a "process" model (Figure 1). The observation model accounts for imperfect detection, whereas the process model provides inference concerning the true but (partially) unobserved occupancy state. The occupancy models considered here differ only in the errors they accommodate and are thus differentiated by their observation models alone. Therefore, we first describe observation models that accommodate false-negative errors and poor specificity, followed by a more general false-positive model. Then, we describe model components that are common across the different specifications (e.g., the process model). See Appendices S1–S3 for additional details regarding model implementation.

2.1 | Classical occupancy model

The classical occupancy model is predicated on the assumption that false negatives are the only observational errors that can occur. Therefore, its use presumes positive test results unmistakably indicate the pathogen is present in a sample (false positives cannot occur), whereas negative test results do not necessarily indicate absence (MacKenzie et al., 2002; Tyre et al., 2003). Conditional on the latent occupancy state (z_i), the observation model that only allows for false-negative errors (i.e., imperfect sensitivity) is specified as

$$y_{ij} \sim \begin{cases} \text{Bernoulli}\,(p_{ij}), & z_i = 1 \\ 0, & z_i = 0 \end{cases}, \tag{1}$$

where p_{ij} is the probability of detecting the pathogen given that it is present (i.e., $\Pr(y_{ij} = 1 \mid z_i = 1)$). In other words, for a sample that is occupied ($z_i = 1$), an observation y_{ij} is a Bernoulli outcome with parameter p_{ij} (Figure 1a). Conversely, y_{ij} must equal 0 if the sample is unoccupied ($z_i = 0$). Accordingly, observations of nonoccurrence ($y_{ij} = 0$) are ambiguous because they can arise in either occupancy state (i.e., when $z_i = 1$ or $z_i = 0$).

2.2 | Occupancy model for poor specificity

Existing false-positive occupancy models are formulated to accommodate poor specificity, or spurious detections in the unoccupied state only (i.e., when $z_i = 0$). In this case, Equation 1 is modified as:

$$y_{ij} \sim \begin{cases} \text{Bernoulli}\,(\pi_{ij}), & z_i = 1 \\ \text{Bernoulli}\,(\phi), & z_i = 0 \end{cases}, \tag{2}$$

where ϕ is the probability of obtaining a false-positive error (see below for a description of π_{ij}). In contrast to Equation 1, which required $y_{ij} = 0$ in the unoccupied state ($z_i = 0$), the observation model presented in Equation 2 allows for positive test results ($y_{ij} = 1$) in a negative sample with probability ϕ. Consequently, observations of occurrence ($y_{ij} = 1$) and nonoccurrence ($y_{ij} = 0$) are now ambiguous because they can arise in either occupancy state (Figure 1b).

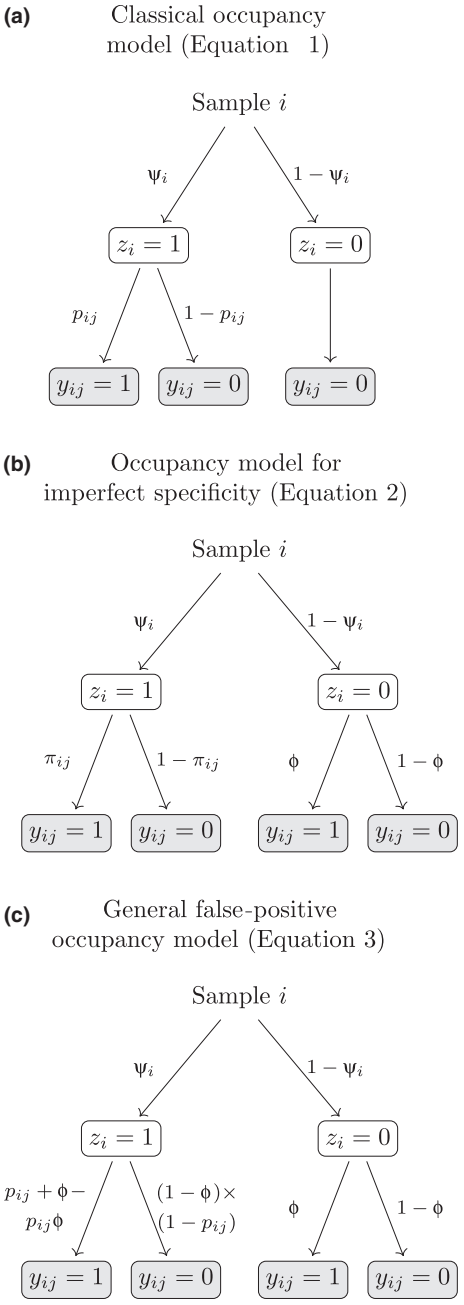

FIGURE 1 Diagrams of (a) the classical occupancy model (Equation 1), (b) the occupancy model that accommodates false-positive errors in the unoccupied state only (Equation 2), and (c) an extension that accommodates false positives in both occupancy states (Equation 3). The variable z_i denotes the true occupancy state for sample i, whereas y_{ij} represents the test outcomes that are observed with error. The arrows are labelled according to the probabilities associated with the occupancy states or observational outcomes. For example, the probability that a sample is occupied ($z_i = 1$) is denoted by ψ_i. Detection probability is denoted by p_{ij}, the probability of a false-positive error is represented by ϕ, and π_{ij} combines the effects of p_{ij} and ϕ

It is often assumed π_{ij} is a "correct" detection probability, consistent with p_{ij} in Equation 1; however, this presumes all positive outcomes in occupied samples are attributable to the detection process. In practice, the mechanisms that yield spurious detections (e.g., contamination and spurious amplification) are independent of the true occurrence state (z_i), and false-positive errors are not exclusive to samples that are truly negative (as implied by Equation 2 where ϕ only appears in the marginal distribution of y_{ij} given $z_i = 0$). Thus, π_{ij} is an *apparent* detection probability because it combines the effects of the detection process and the mechanisms that produce false-positive errors (and is analogous to apparent survival in capture–recapture studies, which combines the effects of survival and study area fidelity; Lebreton, Burnham, Clobert, & Anderson, 1992).

2.3 | General false-positive occupancy model

The occupancy model we present extends Equation 2 to explicitly allow false-positive errors in both the occupied ($z_i = 1$) and unoccupied ($z_i = 0$) states (Figure 1c). The augmented observation model is specified as

$$y_{ij} \sim \begin{cases} \text{Bernoulli}\,(p_{ij} + \phi - p_{ij}\phi), & z_i = 1 \\ \text{Bernoulli}\,(\phi), & z_i = 0, \end{cases} \tag{3}$$

where the probability of obtaining a positive test result given an occupied sample (i.e., $\Pr(y_{ij} = 1 \mid z_i = 1)$) is now expressed as the probability of two independent events, namely observing a genuine detection *or* a false-positive error (Royle & Link, 2006). That is, the probability of observing a positive test result given an occupied sample is a function of both detection (p_{ij}) and false-positive error probabilities (ϕ). Conversely, the probability of observing a negative test result given an occupied sample is

$$\Pr(y_{ij} = 0 \mid z_i = 1) = (1 - \phi)(1 - p_{ij}), \tag{4}$$

where $1 - \phi$ is the probability of not obtaining a false-positive error and $1 - p_{ij}$ is the probability of failing to detect the pathogen (Figure 1c). The marginal distribution for unoccupied samples (i.e., $z_i = 0$) is unchanged from Equation 2.

2.4 | Shared components

2.4.1 | Process model

The process model associated with Equations 1–3 is specified as a Bernoulli random variable

$$z_i \sim \text{Bernoulli}\,(\psi_i), \tag{5}$$

where ψ_i is the true probability of occurrence (i.e., $\Pr(z_i = 1)$). Importantly, p_{ij} and ψ_i are uniquely identifiable only if replicate tests ($J_i > 1$) are available for at least some samples. It is also assumed that the occupancy status (z_i) does not change within replicates of the same sample (MacKenzie et al., 2002).

2.4.2 | Heterogeneity

Occurrence and detection may vary as some function of measurable sample- or replicate-level characteristics. We incorporate covariates related to detection and occurrence probabilities into the model using a logit link

$$p_{ij} = \text{logit}^{-1}(\mathbf{w}_{ij}'\boldsymbol{\alpha}) \tag{6}$$

$$\psi_i = \text{logit}^{-1}(\mathbf{x}_i'\boldsymbol{\beta}), \tag{7}$$

where $\mathbf{w}_{ij}$ is a vector of covariates affecting detection in the jth replicate of sample i, $\mathbf{x}_i$ is a vector of covariates affecting the occupancy of sample i, and $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are vectors of coefficients to be estimated. The parameter π_{ij} (Equation 2) can be modelled in a fashion similar to p_{ij}. Because occupancy describes a sample-level state, the probability of occurrence cannot vary with replicate-level characteristics.

2.4.3 | False-positive error rate

The probability of a false-positive error (ϕ) can be obtained through trial tests on known negative samples (i.e., negative controls; Chambert et al., 2015; McClintock, Nichols et al., 2010). Given the number of samples tested (M), as well as the number of positive outcomes observed (v), we model the negative control data as

$$v \sim \text{Binomial}\,(M, \phi). \tag{8}$$

Furthermore, we incorporate estimation of ϕ (Equation 8) directly into the general parameter estimation framework associated with the false-positive occupancy models (Equations 2 and 3; also see Appendices S2 and S3). Alternatively, ϕ may be obtained from previous studies and treated as a known quantity; however, this approach does not allow uncertainty in the false-positive error rate to propagate through other model components, nor does it accommodate error rates that may be study-specific.

2.4.4 | Prior distributions

We adopt a BAYESIAN framework for analysis and inference that permits retention of the latent variables (i.e., z_i), and thus specify prior distributions for the unknown parameters. We assume $\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \sigma_\alpha^2\mathbf{I})$, $\boldsymbol{\beta} \sim \mathcal{N}\left(\boldsymbol{\mu}_\beta, \sigma_\beta^2\mathbf{I}\right)$, and for the false-positive occupancy models, $\phi \sim \text{Beta}(a, b)$.

3 | MODEL APPLICATIONS

3.1 | Model evaluation using simulated data

We performed a simulation study to examine the effect of observational errors on inference concerning parameters related to detection ($\boldsymbol{\alpha}$) and occupancy ($\boldsymbol{\beta}$). Our simulated data sets consisted of $N = 100$

samples and $J_i = 3$ replicates per sample. The false-positive error rate (ϕ) was set to either 0.05 or 0.10, and $M = 50$ samples were used to simulate the negative control data (Equation 8). Values for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ were chosen such that probabilities of detection and occurrence were roughly in the intervals [0.05, 0.70] and [0.1, 0.9], respectively. We selected these settings in an attempt to simulate data that may realistically be encountered in laboratory studies (e.g., Davenport et al., 2018; Guillera-Arroita, Lahoz-Monfort, van Rooyen, Weeks, & Tingley, 2017; Miller et al., 2012; Mosher et al., 2017; Royle & Link, 2006); however, many parameters can be adjusted to produce different scenarios (e.g., low occupancy and high detection). Therefore, we provide all necessary simulation code (see Supporting Information) and encourage researchers to adjust sample sizes, detection probabilities, false-positive error rates, and occupancy probabilities to mimic their particular application.

We compared an occupancy model that accounts for false-positive errors in both occupancy states (Equation 3) to the simpler occupancy models that ignore false-positive errors (Equation 1) or only allow false positives in the unoccupied state (Equation 2). We also evaluated an ad hoc "classification rule" approach that is commonly used when confronted with observational errors. In this approach, a sample is considered positive if $\geq q$ replicates are positive, and the sample-level outcomes (i.e., $\tilde{y}_i = 1_{\{\sum_j y_{ij} \geq q\}}$, where the indicator function $1_{\{arg\}}$ equals 1 if *arg* is satisfied and 0 otherwise) are subsequently modelled using logistic regression. Because detection probability is often assumed to be perfect under this classification rule approach, the associated parameter estimates are generally interpreted as affecting occurrence. Therefore, we report logistic regression parameter estimates alongside inference concerning $\boldsymbol{\beta}$ obtained from the occupancy models; however, analysing sample-level (rather than replicate-level) outcomes actually leads to inference concerning a random variable that confounds occurrence and detection (Royle & Dorazio, 2008). Each model was fit to 1,000 simulated data sets using Markov chain Monte Carlo (MCMC) algorithms written in R (provided in the Supporting Information; R Development Core Team, 2015). All inferences (BAYESIAN posterior means and 95% credible intervals) were based on 10,000 MCMC samples after convergence, which was determined based on potential scale reduction factors <1.1 (Gelman & Rubin, 1992) and visual inspection of trace plots (Hobbs & Hooten, 2015).

The false-positive occupancy model we present (Equation 3) provides approximately unbiased inference concerning parameters related to detection ($\boldsymbol{\alpha}$) and occupancy ($\boldsymbol{\beta}$) in the presence of both false-positive and false-negative errors (Figure 2). The model accounting for false-positive errors in the unoccupied state only (Equation 2) provides reliable inference concerning $\boldsymbol{\beta}$, but yields biased and overly confident inference for detection with actual coverage probabilities that are below nominal coverages (Table 1, Figure 2). Similarly, strategies that are based on arbitrary classification rules (i.e., $q = 1$ or 2) or otherwise ignore false-positive errors (Equation 1) provide wholly unreliable inference (Table 1, Figure 2). Unless false-positive errors are incorporated into both occupancy states (Equation 3), bias in parameter estimates increases (and actual coverage probabilities decrease) as ϕ increases (Table 1, Figure 2).

3.2 | Case study: *Batrachochytrium dendrobatidis*

Chytridiomycosis is an amphibian disease caused by *Batrachochytrium dendrobatidis* (Bd), a fungal pathogen that is implicated in the decline or extinction of more than 200 amphibian species worldwide (Skerratt et al., 2007). Swabbing amphibian skin is the recommended method for detecting Bd (Hyatt et al., 2007); however, relying on swabs limits studies of amphibian-Bd dynamics to known host populations and collection can be difficult when host species are rare. Consequently, detecting the pathogen in its free-living form is necessary to inform management decisions (e.g., reintroduction sites) and understand host–pathogen dynamics, metapopulation dynamics and long-term species persistence (Mosher, Bailey, Hubbard, & Huyvaert, 2018).

Coupling water filtration and qPCR is a promising option for detecting Bd in the aquatic environment. Mosher et al. (2017) conducted an experiment to evaluate the efficacy of this approach, and to quantify the effect of pathogen concentration and water type on Bd detection probability. They randomly assigned samples to experimental groups and analysed DNA extracts in triplicate after the removal of PCR inhibitors (Mosher et al., 2017). The use of thresholds to eliminate questionable samples in previous studies (Shin, Bataille, Kosch, & Waldman, 2014; Venesky, Liu, Sauer, & Rohr, 2014), as well as observed "detections" in a 0 Bd zoospore/ml control group (Mosher et al., 2017), indicates water filtration is prone to false-positive errors. Consequently, Mosher et al. (2017) classified replicates as positive only if they exceeded the highest qPCR copy number observed in the control group and analysed the resulting data using classical occupancy models (e.g., Equation 1).

We used the extended occupancy model (Equation 3) to quantify the effect of Bd concentration on detection probability in $N = 144$ natural water samples while fully accounting for false-positive errors. For purposes of comparison, we also fit the classical occupancy model (Equation 1) to the data after imposing the copy number threshold (as in Mosher et al., 2017). Given that samples were experimentally inoculated with Bd (and our focus here was specifically on pathogen detection), we fixed the latent occupancy state to $z_i = 1$ for all samples during model fitting. Six of 72 (M) negative control samples returned positive outcomes, yielding an empirical false-positive error rate of 8%. The probability of detecting Bd increased with Bd concentration under both analytical scenarios (Figure 3a); however, compared to our model, detection probability was estimated to be 10% lower, on average, when the classical occupancy model and thresholded data set were used. The largest difference between estimated detection probabilities occurred in the 0.05 zoospore/ml concentration samples (−24% lower), and parameters related to detection ($\boldsymbol{\alpha}$) were estimated to be as much as 170% lower under the classical occupancy model. In the context of study design, the false-positive modelling framework (Equation 3) suggests four replicates are necessary to obtain a 95% probability of

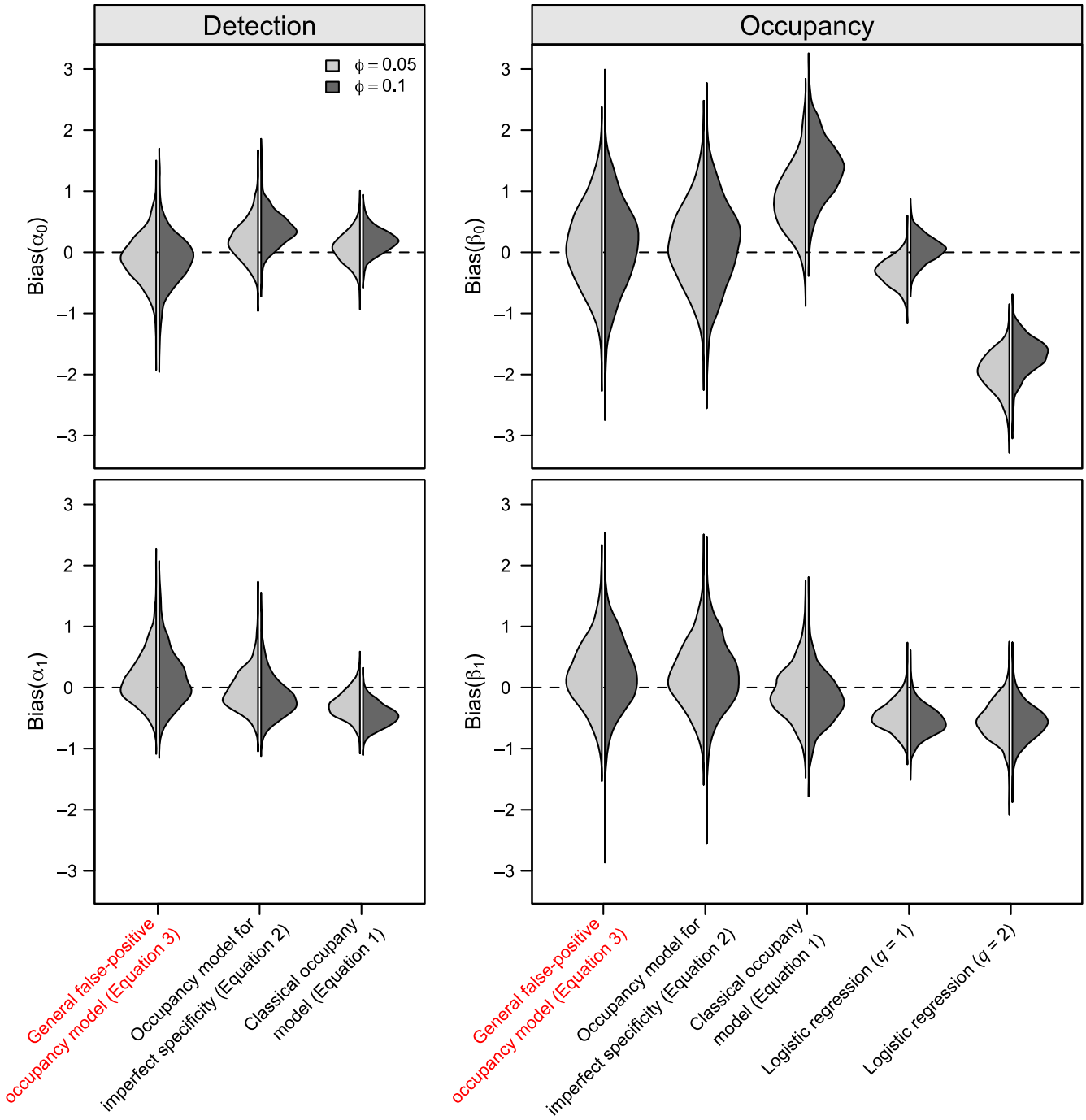

FIGURE 2 Bias in parameters related to detection (**α**, left) and occupancy (**β**, right) when data contaminated with false-negative and false-positive errors are modelled using various approaches. Violins depict the distribution of bias observed in point estimates (posterior means) over 1,000 simulated data sets. Light grey violins represent bias observed when ϕ, the false-positive error rate, was set to 0.05; dark grey violins represent results obtained when $\phi = 0.1$ [Colour figure can be viewed at wileyonlinelibrary.com]

observing one or more genuine Bd detections (p^*) in a 0.05 zoospore/ml concentration sample (Figure 3b). Conversely, the classical occupancy model implies six replicates are necessary to achieve the same probability, an artefact of an experimentally derived threshold that is meant to address false-positive errors, but also removes true detections and thus depresses detection probabilities (Mosher et al., 2017).

4 | DISCUSSION

Diagnosticians are rarely able to detect microorganisms, DNA or pathogens with perfect accuracy. If not properly incorporated into inferential methods, false-negative and false-positive errors can lead to unfounded conclusions about occurrence, effects, prevalence, incidence, dynamics and detectability (McClintock, Nichols et al., 2010).

TABLE 1 Actual coverage of 95% credible intervals for parameters related to detection (α) and occurrence (β) when data contaminated with false-negative and false-positive errors are modelled using various approaches

	$\phi = 0.05$				$\phi = 0.10$			
Estimation method	α_0	α_1	β_0	β_1	α_0	α_1	β_0	β_1
General false-positive occupancy model (Equation 3)	0.96	0.94	0.94	0.97	0.97	0.94	0.94	0.98
Occupancy model for imperfect specificity (Equation 2)	0.92	0.92	0.93	0.97	0.78	0.88	0.95	0.97
Classical occupancy model (Equation 1)	0.95	0.64	0.55	0.95	0.86	0.28	0.10	0.96
Logistic regression ($q = 1$)	—	—	0.66	0.45	—	—	0.94	0.33
Logistic regression ($q = 2$)	—	—	0.00	0.57	—	—	0.00	0.48

Coverages that are below 0.95 indicate the estimation method does not accurately reflect uncertainty in the parameter estimate. Coverage was calculated as the proportion of intervals that contain the true parameter value over 1,000 simulated data sets.

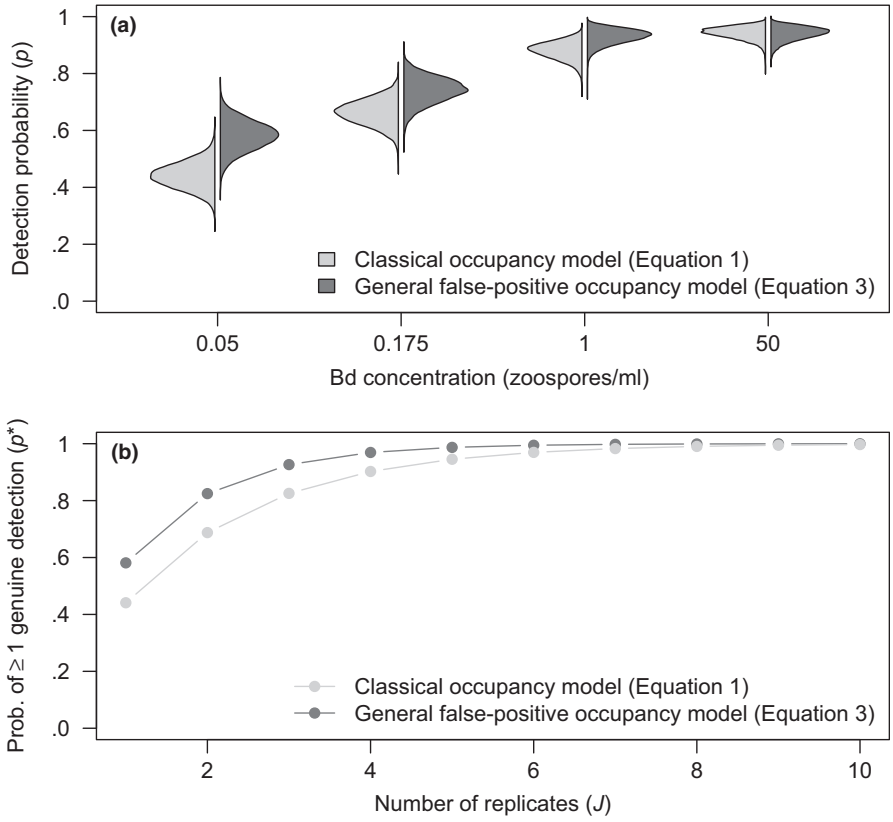

FIGURE 3 Estimated probabilities of detecting Bd in natural water samples. (a) Posterior distributions of detection probabilities (p) across four experimental Bd concentrations and two analytical approaches. The general false-positive occupancy model (dark grey) was fit to the raw detection/nondetection data and incorporated a negative control data set to fully accommodate false-positive errors in both occupancy states (Equation 3). The classical occupancy model (light grey; Equation 1) was fit to the data after imposing a qPCR copy number threshold to address false-positive errors. (b) The probability of observing at least one genuine detection given some number of replicates tested ($p^* = 1 - (1 - p)^J$) for the 0.05 Bd zoospore/ml concentration

Occupancy models formally separate the detection and occurrence processes, thereby providing a natural framework within which observational errors can be addressed in a rigorous, statistically valid manner (Royle & Dorazio, 2008). To our knowledge, the occupancy model presented in Equation 3 represents the only modelling framework that provides unbiased inference for both occupancy and detection when data are contaminated with false-negative and false-positive errors.

Occupancy models were initially developed to account for uncertainty induced by false-negative errors and relied exclusively on repeated sampling (MacKenzie et al., 2002; Tyre et al., 2003). Royle and Link (2006) provided an approach for extending the general

framework to accommodate false-positive errors, and subsequent developments in false-positive occupancy modelling incorporate additional information on p_{ij} and ϕ to overcome limitations in the original model (Chambert et al., 2015; Guillera-Arroita et al., 2017; McClintock, Nichols et al., 2010; Miller et al., 2011). We use negative control samples as ancillary information to inform the rate of false-positive errors (ϕ), an approach that is broadly applicable because tests on reference samples of known state are common in laboratory protocols. Even though detection types or methods that provide "certain" detections (i.e., true positives only) have also been used to inform error rates (Miller et al., 2011), this approach may be impractical because "gold standard" methods that only yield unambiguous detections are often unavailable, expensive or otherwise inefficient.

Previous models only allow false-positive errors to occur when the sample is truly negative (e.g., Equation 2; Chambert et al., 2015; Guillera-Arroita et al., 2017; McClintock, Nichols et al., 2010). Models formulated under this assumption provide reliable inference concerning parameters related to occurrence ($\boldsymbol{\beta}$) because false-positive outcomes in occupied samples represent correct (but accidental) classifications (Table 1, Figure 2; Royle & Link, 2006); however, they do not provide reliable inference concerning detection. In this case, π_{ij} (Equation 2) does not represent a true detection probability, but rather the "net" rate at which positive samples are correctly classified either due to the real detection process or due to a false-positive error (Royle & Link, 2006). Consequently, the application of models that assume false-positive errors only occur in the unoccupied state is not suitable for addressing questions regarding the efficacy of alternate sampling designs and laboratory procedures (Table 1, Figure 2; e.g., Davenport et al., 2018; Pilliod, Goldberg, Arkle, & Waits, 2013; Wilcox et al., 2013, 2016; Williams, Huyvaert, & Piaggio, 2017). By letting false-positive errors arise in both occupancy states, p_{ij} in the model we present (Equation 3) is a "correct" detection probability, thereby providing accurate inference concerning detection and enabling reliable optimization of diagnostic procedures (Table 1, Figure 2). For example, selecting a procedure (e.g., method to concentrate, extract and amplify DNA, handling protocol) that maximizes the difference between true detection probabilities (p_{ij}, not π_{ij}) and false-positive error rates (ϕ) is one approach that can be used to increase the precision of occupancy estimates (ψ_i), the parameters typically of most interest. Ignoring false-positive errors altogether, or hiding both types of observational errors in a preprocessing stage, yields parameter estimates that are biased and overly confident (Table 1, Figure 2).

We allow heterogeneity in detection and occupancy probabilities by modelling p_{ij} and ψ_i as functions of covariates (Equations 6 and 7). Heterogeneity in the false-positive error rate can also be modelled in a similar manner, if appropriate (e.g., $\phi_{ij} = \text{logit}^{-1}(\mathbf{u}_{ij}'\boldsymbol{\gamma})$). Indeed, detection and false-positive error rates are likely affected by different covariates, a circumstance that cannot be fully incorporated into the simpler false-positive modelling framework (Equation 2). Reliance on Equation 2 could thus yield misleading results even if occupancy is the sole quantity of interest because unmodelled observation error can bias parameter estimates (McClintock, Bailey et al., 2010). This issue in particular deserves further attention. Along these lines, some laboratory techniques provide numerical values associated with detection. For example, qPCR provides an estimated abundance of DNA sequences and spectrophotometers permit calculating the mass of particles or molecules. Although numeric values like these have been used to establish thresholds intended to eliminate questionable detections (e.g., Shin et al., 2014; Venesky et al., 2014), the same information could be used more appropriately as covariates to help inform observation error probabilities (i.e., abundance-induced heterogeneity; Royle & Nichols, 2003).

The observation model we specified in Equation 3 could be expanded to include a latent indicator variable that denotes a false-positive error, just as the latent variable z_i indicates the true occupancy state (Appendix S4). Such an approach is useful for estimating which observations may indeed be false positives. Our interest here is inference concerning detection (p_{ij}) and occupancy (ψ_i). Therefore, we focus on the "marginalized" version (Equation 3), wherein the indicator variables have been removed from the likelihood (Berger, Liseo, & Wolpert, 1999), because it has a reduced parameter space and yields a MCMC algorithm that is quicker to converge (Finley, Banerjee, & Gelfand, 2015).

The occupancy modelling framework has been extended to address a wide variety of questions and scenarios that are often encountered in the field of ecology. For example, occupancy models have been modified to assess occurrence at multiple spatial scales (Nichols et al., 2008), evaluate the co-occurrence of multiple species (Dorazio, Kéry, Royle, & Plattner, 2010), incorporate additional latent occupancy states (MacKenzie, Nichols, Seamans, & Gutiérrez, 2009) and examine spatial and temporal dynamics (Johnson, Conn, Hooten, Ray, & Pond, 2012; Royle & Kéry, 2007). We anticipate similar scenarios where data are contaminated not only with false-negative errors, but also false positives (e.g., Guillera-Arroita et al., 2017). The model we propose (Equation 3) could be extended to each of these scenarios and used, for example, to examine temporal dynamics in disease occurrence or its occurrence at multiple scales (e.g., Davenport et al., 2018).

5 | CONCLUSION

Even though known negative and positive samples are often used to optimize diagnostic tests for sensitivity and specificity, such calibrations do not ensure accurate results. Improved handling protocols and sampling designs can help reduce false negatives and false positives, but these measures are unlikely to eliminate all sources of error (McClintock, Nichols et al., 2010). When confronted with conflicting outcomes among replicate tests or errant results in control samples, practitioners have little guidance on how to proceed and often turn to ad hoc approaches that have hidden and unintended consequences. Model-based approaches like the one we present are not only useful for obtaining reliable inference when data are

contaminated with observational errors, but also eliminate the need for arbitrary thresholds and decision rules.

Occupancy models are becoming more accessible to practitioners through various software platforms (e.g., Presence, MARK, R package *unmarked*); however, the model we present (Equation 3) currently requires customized code for implementation. We provide this code and details regarding our MCMC algorithm to facilitate the analysis of other data sets (see Appendices and Supporting Information). Parameter estimation could also be performed using maximum likelihood or specialized software for BAYESIAN analysis (e.g., BUGS, JAGS or Stan; Carpenter et al., 2017; Lunn, Thomas, Best, & Spiegelhalter, 2000; Plummer, 2003; Royle & Dorazio, 2008). Nonetheless, we acknowledge that occupancy models present an additional complexity in analysing data. We encourage collaborations between researchers and statisticians to overcome analytical challenges like those addressed here.

ACKNOWLEDGEMENTS

We thank Phil Turk and Thierry Chambert for insightful conversations regarding this research and Brett McClintock and an anonymous reviewer for helpful comments on a previous version of the manuscript. The findings and conclusions in the manuscript are those of the authors and do not necessarily represent the views of the National Marine Fisheries Service. Reference to trade names does not imply endorsements by the National Marine Fisheries Service.

DATA ACCESSIBILITY

Simulation code, custom MCMC algorithms for model fitting and experimental data used for the *Batrachochytrium dendrobatidis* case study are provided in the Supporting Information.

AUTHOR CONTRIBUTIONS

B.M.B., B.A.M. and K.A.D. designed and performed the research. B.M.B. analysed the data. B.M.B., B.A.M. and K.A.D. wrote the manuscript.

REFERENCES

Berger, J. O., Liseo, B., & Wolpert, R. L. (1999). Integrated likelihood methods for eliminating nuisance parameters. *Statistical Science*, *14*, 1–22.

Carey, C., Bruzgul, J. E., Livo, L. J., Walling, M. L., Kuehl, K. A., Dixon, B. F., … Rogers, K. B. (2006). Experimental exposures of boreal toads (*Bufo boreas*) to a pathogenic chytrid fungus (*Batrachochytrium dendrobatidis*). *EcoHealth*, *3*, 5–21. https://doi.org/10.1007/s10393-005-0006-4

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*, 1–32.

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

Davenport, K. A., Mosher, B. A., Brost, B. M., Henderson, D. M., Denkers, N. D., Nalls, A. V., … Hoover, E. A. (2018). Assessment of chronic wasting disease prion shedding in deer saliva with occupancy modeling. *Journal of Clinical Microbiology*, *56*, e01243-17.

Dorazio, R. M., Kéry, M., Royle, J. A., & Plattner, M. (2010). Models for inference in dynamic metacommunity systems. *Ecology*, *91*, 2466–2475. https://doi.org/10.1890/09-1033.1

Finley, A., Banerjee, S., & Gelfand, A. (2015). spBayes for large univariate and multivariate point-referenced spatio-temporal data models. *Journal of Statistical Software*, *63*, 1–28.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*, 457–472. https://doi.org/10.1214/ss/1177011136

Guillera-Arroita, G., Lahoz-Monfort, J. J., van Rooyen, A. R., Weeks, A. R., & Tingley, R. (2017). Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods in Ecology and Evolution*, *8*, 1081–1091. https://doi.org/10.1111/2041-210X.12743

Hobbs, N. T., & Hooten, M. B. (2015). *Bayesian models: A statistical primer for ecologists*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400866557

Hyatt, A. D., Boyle, D. G., Olsen, V., Boyle, D. B., Berger, L., Obendorf, D., … Colling, A. (2007). Diagnostic assays and sampling protocols for the detection of *Batrachochytrium dendrobatidis*. *Diseases of Aquatic Organisms*, *73*, 175–192. https://doi.org/10.3354/dao073175

Johnson, D. S., Conn, P. B., Hooten, M. B., Ray, J. C., & Pond, B. A. (2012). Spatial occupancy models for large data sets. *Ecology*, *94*, 801–808.

Lachish, S., Gopalaswamy, A. M., Knowles, S. C. L., & Sheldon, B. C. (2012). Site-occupancy modelling as a novel framework for assessing test sensitivity and estimating wildlife disease prevalence from imperfect diagnostic tests. *Methods in Ecology and Evolution*, *3*, 339–348. https://doi.org/10.1111/j.2041-210X.2011.00156.x

Lahoz-Monfort, J. J., Guillera-Arroita, G., & Tingley, R. (2016). Statistical approaches to account for false-positive errors in environmental DNA samples. *Molecular Ecology Resources*, *16*, 673–685. https://doi.org/10.1111/1755-0998.12486

Lebreton, J., Burnham, K. P., Clobert, J., & Anderson, D. R. (1992). Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. *Ecological Monographs*, *62*, 67–118. https://doi.org/10.2307/2937171

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, *10*, 325–337. https://doi.org/10.1023/A:1008929526011

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. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2

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., Nichols, J. D., Bailey, L. L., MacKenzie, D. I., Kendall, W. L., & Franklin, A. B. (2010). Seeking a second opinion: Uncertainty in disease ecology. *Ecology Letters*, *13*, 659–674. https://doi.org/10.1111/j.1461-0248.2010.01472.x

Miller, D. A., 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

Mosher, B. A., Bailey, L. L., Hubbard, B. A., & Huyvaert, K. P. (2018). Inferential biases linked to unobservable states in complex occupancy models. *Ecography*, *41*, 32–39. https://doi.org/10.1111/ecog.02849

Mosher, B. A., Huyvaert, K. P., Chestnut, T., Kerby, J. L., Madison, J. D., & Bailey, L. L. (2017). Design- and model-based recommendations for detecting and quantifying an amphibian pathogen in environmental samples. *Ecology and Evolution*, *7*, 10952–10962. https://doi.org/10.1002/ece3.3616

Nichols, J. D., Bailey, L. L., O'Connell, A. F. Jr, Talancy, N. W., Campbell Grant, E. H., Gilbert, A. T., … Hines, J. E. (2008). Multi-scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology*, *45*, 1321–1329. https://doi.org/10.1111/j.1365-2664.2008.01509.x

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., & Waits, L. P. (2013). Estimating occupancy and abundance of stream amphibians using environmental DNA from filtered water samples. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*, 1123–1130. https://doi.org/10.1139/cjfas-2013-0047

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch & A. Zeileis (Eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, Vienna, Austria*. Retrieved from https://www.r-project.org/conferences/DSC-003/Proceedings/Plummer.pdf

R Development Core Team (2015). *R: A language and environment for statistical computing* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities*. Amsterdam, the Netherlands: Academic Press.

Royle, J. A., & Kéry, M. (2007). A Bayesian state-space formulation of dynamic occupancy models. *Ecology*, *88*, 1813–1823. https://doi.org/10.1890/06-0669.1

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

Royle, J. A., & Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, *84*, 777–790. https://doi.org/10.1890/0012-9658(2003)084[0777:EAFRPA]2.0.CO;2

Schmidt, B. R., Kéry, M., Ursenbacher, S., Hyman, O. J., & Collins, J. P. (2013). Site occupancy models in the analysis of environmental DNA presence/absence surveys: A case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution*, *4*, 646–653. https://doi.org/10.1111/2041-210X.12052

Schrader, C., Schielke, A., Ellerbroek, L., & Johne, R. (2012). PCR inhibitors – Occurrence, properties and removal. *Journal of Applied Microbiology*, *113*, 1014–1026. https://doi.org/10.1111/j.1365-2672.2012.05384.x

Shin, J., Bataille, A., Kosch, T. A., & Waldman, B. (2014). Swabbing often fails to detect amphibian chytridiomycosis under conditions of low infection load. *PLoS One*, *9*, 1–8.

Skerratt, L. F., Berger, L., Speare, R., Cashins, S., McDonald, K. R., Phillott, A. D., … Kenyon, N. (2007). Spread of chytridiomycosis has caused the rapid global decline and extinction of frogs. *EcoHealth*, *4*, 125. https://doi.org/10.1007/s10393-007-0093-5

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

Venesky, M. D., Liu, X., Sauer, E. L., & Rohr, J. R. (2014). Linking manipulative experiments to field data to test the dilution effect. *Journal of Animal Ecology*, *83*, 557–565. https://doi.org/10.1111/1365-2656.12159

Wilcox, T. M., McKelvey, K. S., Young, M. K., Jane, S. F., Lowe, W. H., Whiteley, A. R., & Schwartz, M. K. (2013). Robust detection of rare species using environmental DNA: The importance of primer specificity. *PLoS One*, *8*, 1–9.

Wilcox, T. M., McKelvey, K. S., Young, M. K., Sepulveda, A. J., Shepard, B. B., Jane, S. F., … Schwartz, M. K. (2016). Understanding environmental DNA detection probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. *Biological Conservation*, *194*, 209–216. https://doi.org/10.1016/j.biocon.2015.12.023

Williams, K. E., Huyvaert, K. P., & Piaggio, A. J. (2017). Clearing muddied waters: Capture of environmental DNA from turbid waters. *PLoS One*, *12*, 1–17.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Brost BM, Mosher BA, Davenport KA. A model-based solution for observational errors in laboratory studies. *Mol Ecol Resour*. 2018;18:580–589. https://doi.org/10.1111/1755-0998.12765