Received: 3 March 2019 | Accepted: 4 July 2019

DOI: 10.1111/2041-210X.13262

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Cohesive framework for modelling plant cover class data

Kathryn M. Irvine[1] | Wilson J. Wright[2] | Erin K. Shanahan[3] | Thomas J. Rodhouse[4]

[1]U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, USA

[2]Department of Ecology, Montana State University, Bozeman, MT, USA

[3]Greater Yellowstone Network, U.S. National Park Service, Bozeman, MT, USA

[4]Upper Columbia Basin Network, U.S. National Park Service, Bend, OR, USA

Correspondence
Kathryn M. Irvine
Email: kirvine@usgs.gov

Funding information
National Park Service, Grant/Award Number: IAA P12PG70586 and P17AC00623; U.S. Fish and Wildlife Service, Grant/Award Number: IAA 450063656; Montana State University

Handling Editor: Jari Oksanen

Abstract

1. The study of plant distribution and abundance is a fundamental pursuit in ecology and conservation biology. Measuring plant abundance by visually assessing percent cover and recording a cover class is a common field method that yields ordinal data. Statistical models for ordinal data exist but entail cumbersome interpretations and sometimes restrictive assumptions.
2. We propose a Bayesian hierarchical framework for analysing cover class data that allows for linking ordinal observations to a latent beta distribution and accounts for zero inflation. Harnessing a latent beta distribution supports interpreting changes in abundance in terms of mean percent cover rather than odds ratios of cumulative cover classes as for cumulative link models. The zero augmentation allows for simultaneous inferences on both occurrence (distribution) and abundance. We show how our model can account for true and false zeros, misclassification of cover classes, multiple species and hierarchical sampling designs, using empirical examples and simulations.
3. Simulated observation errors, when ignored, led to models overestimating abundance and underestimating occurrence. Based on simulations, we found no substantial difference between mean percent cover estimates when analyzing ordinal cover classes versus continuous percent cover as the response. Our empirical datasets displayed high probability of detection (>0.85 on average for all species), likely due to the sampling design used and training of observers. Probability of occurrence was slightly underestimated for bare ground, *Artemisia tridentata*, *Elycap medusae*, and *Poa secunda* using a model that ignored imperfect detection. Estimated mean percent cover was not substantially impacted by ignoring measurement error for five plant species and bare ground.
4. Our modelling framework for cover class data allows for an explicit separation of distribution from abundance and, importantly, allows for interpreting species–environment relationships in terms of variation in mean percent cover as compared to cumulative odds ratios. The beta distribution inherently accommodates heteroscedasticity and skewness, statistical properties that are a consequence of spatially aggregated patterns common to plant survey data. Recording cover classes provides a reliable, efficient way to measure plants and our simulations suggest little loss of information compared to assuming continuous percent cover. We provide JAGS and Stan model code for implementation.

This article has been contributed to by US employees and their work is in the public Domain in the USA.

 © 2019 The Authors. Methods in Ecology and Evolution © 2019 British Ecological Society

KEYWORDS
beta regression, imperfect detection, measurement error, monitoring, observer errors, occupancy models, ordinal models, plant cover

1 | INTRODUCTION

The study of plant species distribution and abundance are fundamental to informing conservation and management. A common field method for plant surveys is to visually assess the percent cover within an areal plot and subsequently record a cover class value as an estimate of abundance (e.g. 0, 1, 2,..., 6). Cover class data are an example of ordered categorical or ordinal data. Statistical models appropriate for such data are available (e.g., cumulative link models, Agresti, 2010; Warton, Foster, De'ath, Stoklosa, & Piers, 2015), but recently an ordinal zero-augmented beta model (OZAB) was proposed as a way to directly link the cover classes to the partially observed percent cover of a plot (Herpigny & Gosselin, 2015; Irvine, Rodhouse, & Keren, 2016). The OZAB model allows for investigating the drivers of distribution (presence/absence) separate from environmental factors governing abundance and provides a more appealing interpretation of model parameters in terms of changes in mean percent cover compared to cumulative odds ratios (Irvine et al., 2016). Here we expand the OZAB model to account for imperfect detection and measurement errors in plant cover class data. We introduce this modelling framework as an ordinal zero-augmented beta with errors model (hereafter, OZABE). Additionally, our empirical examples demonstrate the flexibility of OZAB(E) to account for other analytical complications often encountered in plant survey data, such as cluster-type sampling designs and observations on multiple species.

The novelty of the OZAB framework is that we explicitly acknowledge that percent cover, which is the variable of primary interest, is partially observed for cover class data because for each plot the cover is known within an interval, such as between 2.5% to 5%. Using a Bayesian hierarchical framework, we specify the latent percent cover variable as arising from a zero-augmented beta distribution. The beta distribution has been advocated as suitable for plant percent cover for many reasons (see review in Damgaard & Irvine, 2019). Within the unit interval, the beta density can be J-, L-, or U-shaped providing tremendous flexibility for modelling skewness and heteroscedasticity both common attributes of plant cover datasets. Damgaard (2013) explored hypotheses related to changes in species spatial aggregation within a homogeneous site using the beta distribution. The zero-augmented portion of OZAB allows for modelling true zeros, in our case, plots with a recorded cover class of zero. A side benefit of the two-part model structure is that it provides simultaneous inferences about factors governing distribution *and* abundance (Irvine et al., 2016). For native species, distribution modelling can inform habitat suitability and species range shifts whereas changes in abundance can inform population growth and trends over time. Finally, the Bayesian hierarchical framework accommodates extensions to OZAB that account for observation errors, multiple species (a community or joint species model), and cluster-type sampling designs.

In plant surveys, observation errors can include both measurement or misclassification error (recorded cover classes differ from truth; e.g. Gorrod, Bedward, Keith, & Ellis, 2013; Helm & Mead, 2004; Killourhy, Crane, & Stehman, 2016) and imperfect detection (species is not observed even though present; e.g. Chen, Kéry, Plattner, Ma, & Gardner, 2013; Chen, Kéry, Zhang, & Ma, 2009; Clarke, Lewis, Brandle, & Ostendorf, 2012; Vittoz et al., 2010). Wright, Irvine, Warren, and Barnett (2017) developed a beta regression model for percent cover data that accounts for measurement error and imperfect detection. Estimates of both mean percent cover and proportion of an area occupied can be biased when observation errors are ignored (Wright et al., 2017). Our OZABE model accommodates observation errors in cover class data by utilizing the formulation of beta regression with observation errors as in Wright et al. (2017) for the partially observed (latent) variable.

Monitoring programs employ plot-based methods to survey plant communities and record information on multiple species simultaneously. In our example datasets, a targeted set of species were a priori selected for long-term monitoring because they are relevant to management decisions (Yeo, Rodhouse, Dicus, Irvine, & Garrett, 2009). This approach is slightly different than community plant surveys that record all species present in a plot. Joint species or community models are available for data collected as presence/absence or stem counts (Kéry & Royle, 2016). One option for cover class data is to use a multivariate cumulative probit model (e.g. Schliep et al., 2016), in which the latent variable is specified as multivariate Gaussian (normal) distribution. Alternatively, a Dirichlet model, a multivariate extension to the beta distribution, may be appropriate (e.g. Damgaard, 2015). We present a formulation consistent with other joint species or community models (Clark, Nemergut, Seyednasrollah, Turner, & Zhang, 2017; Warton, Blanchet, et al., 2015) except we specify that the latent percent cover variable is distributed as a zero-augmented beta. To demonstrate our approach, we analyse cover class data for four species of conservation concern (two invasive annual grass and two native bunchgrasses) and bareground collected at John Day Fossil Beds National Monument using our joint species OZAB(E) models. Variation in the patterns of co-occurrence and abundances of these species along environmental gradients in western American sagebrush steppe is a central concern among conservation practitioners.

We also consider how the OZAB(E) models can accommodate data collected following a hierarchical sampling design in which multiple units or areas are selected oftentimes for a specific management interest. Within each unit, multiple plots are randomly selected and the plant community or set of targeted species are surveyed. We present an example from Bighorn Canyon National Recreation

Area in which three units were selected for their management significance; two units within sagebrush steppe/grassland communities were identified as 'good' sage grouse habitat while a third was noted as unaltered juniper woodland (Tercek, Jean, Daley, & Legg, 2013). This cluster-type design is beneficial in rugged terrain because aggregating plots within core areas of interest can reduce overall travel costs. However, the design structure needs to be appropriately modelled otherwise standard errors may be underestimated.

Finally, we conduct a simulation study to assess the performance (bias and coverage properties) of OZAB and OZABE for estimating mean proportion cover and probability of occurrence. Further, we explore whether observation errors differentially affected parameter estimation if observers recorded percent cover on a continuous versus ordinal scale. One perceived benefit to using ordinal cover classes is a reduction in observer variation. We provide the code used to model the empirical datasets as Appendix S1 and the Stan and JAGS model files (Carpenter et al., 2016; Plummer, 2003) and datasets are accessible at https://doi.org/10.5066/P99JX2ND.

2 | MATERIALS AND METHODS

2.1 | Ordinal zero-augmented beta (OZAB) model

The ordinal zero-augmented beta model was proposed as an alternative to cumulative logit or probit models for ordinal data in Irvine et al. (2016) and Herpigny and Gosselin (2015). Following the notation in Irvine et al. (2016), the observed cover class data are of the grouped continuous type of ordinal response and can be considered a realization from a multinomial distribution with $J + 1$ categories, $[C_i \mid \boldsymbol{\pi}_i] = \text{Multinomial}(1, \pi_{0i}, ..., \pi_{Ji})$, for $i = 1, ..., n$ observations (here, descriptions of areal 1 m^2 plots) with category levels of 0, 1, ..., J and $\Pr(C_i = j) = \pi_{ji}$. In our case, following Daubenmire (1959), there were seven pre-defined categories ($J = 6$), but other classification schemes are possible (e.g. Milberg, Bergstedt, Fridman, Odell, & Westerberg, 2008). The Daubenmire scale used in our examples was a seven cover class scheme (0, 1, ..., 6) with zero representing absence of cover and the following interval divisions of areal proportion covered: >0–0.05, >0.05–0.25, >0.25–0.50, >0.50–0.75, >0.75–0.95, and >0.95–1.

In the OZAB model, we specify the latent variable (Y_i^*) distribution as a mixture of a continuous beta distribution and a point mass at 0, a zero-augmented beta (ZAB) distribution (Ospina & Ferrari, 2010), as follows:

$$\left[Y_i^* \mid \psi, \mu, \phi\right] = \begin{cases} 1-\psi, & \text{if } Y_i^* = 0, \\ (\psi)[Y_i \mid \mu, \phi], & \text{if } Y_i^* \in (0, 1), \end{cases} \tag{1}$$

where $[Y_i \mid \mu, \phi]$ is the beta distribution as parameterized for regression analysis by Ferrari and Cribari-Neto (2004) with $E(Y_i) = \mu$ and $\text{Var}(Y_i) = \frac{\mu(1-\mu)}{\phi+1}$, where $0 < \mu < 1$ and $\phi > 0$. For cover class data, μ is the mean proportion cover (or $\mu \times 100$ for mean percent cover) conditional on presence. The unconditional mean cover, if of interest, is $\psi\mu$. A re-parameterization of ϕ has been interpreted as a measure of spatial aggregation for plants, $\delta = \left(\frac{1}{1+\phi}\right)$ (Damgaard, 2013). As the patchiness or spatial aggregation of a plant within a homogeneous area increases, δ increases and the beta distribution tends to a U-shape for a fixed value μ (Figure 2 in Damgaard & Irvine, 2019). As with any ecological study, the interpretation of these parameters is dependent on the chosen grain size (plot size and shape) and spatial extent.

Then, following the usual latent variable motivation for an ordinal observation Agresti (2010), we specify that the observed cover class category is a discretized version of Y_i^* for observation i as follows,

$$C_i = \begin{cases} 0, & \text{if and only if } Y_i^* = 0 \\ j, & \text{if and only if } \theta_{j-1} < Y_i^* \le \theta_j \text{ for } j = 1, ..., J, \end{cases} \tag{2}$$

where $\boldsymbol{\theta}$ is a set of known values that define the breaks between each cover class. In our case, for example, the Daubenmire (1959) scale defines the vector as $\theta_0 = 0$, and $\{\theta_1, \cdots, \theta_5\} = \{0.05, 0.25, 0.5, 0.75, 0.95\}$ based on the breaks between classes, and $\theta_6 = 1$. By definition the multinomial probabilities are

$$\pi_{ij} = \begin{cases} 1-\psi & j=0 \\ \psi \times F_{\text{Beta}}(\theta_1) & j=1 \\ \psi \times (F_{\text{Beta}}(\theta_j) - F_{\text{Beta}}(\theta_{j-1})) & 2 \le j \le (J-1) \\ \psi \times (1 - F_{\text{Beta}}(\theta_{j-1})) & j=J \end{cases} \tag{3}$$

where F_{Beta} denotes the cumulative distribution function for the partially observed percent cover Y, which is beta distributed with parameters μ and ϕ. This model can be written hierarchically as follows, $[Z_i = 1] = \text{Bernoulli}(\psi)$ and $[Y_i | Z_i = 1] = \text{Beta}(\mu, \phi)$ (see the ecological state at plot i box displayed in Figure 1).

To understand how environmental variables may influence abundance and distribution, we extended the mean of proportion cover (μ) and probability of occurrence (ψ) by $\text{logit}(\mu_i) = \mathbf{W}_i\boldsymbol{\beta}$ and $\text{logit}(\psi_i) = \mathbf{L}_i\boldsymbol{\gamma}$, where $\mathbf{W}$ and $\mathbf{L}$ represent a matrix of covariates measured at each plot i (Irvine et al., 2016). In explanatory models, interpreting the effect of an environmental variable on the unconditional mean ($\psi\mu$) is complicated because covariates are included and their effects estimated conditionally (i.e. distribution separate from abundance). Therefore, interpreting an effect on the unconditional mean should be considered carefully in conjunction with assessing impacts separately.

2.2 | Ordinal zero-augmented beta with errors (OZABE) Model

Next we consider how to incorporate imperfect detection and measurement error into the observation process for plant cover class data. Essentially, we modify the zero-process model allowing for the possibility that a plant was present but not detected by using a zero-inflated binomial model, similar to occupancy models for detection/non-detection data (MacKenzie et al., 2018). Then the abundance

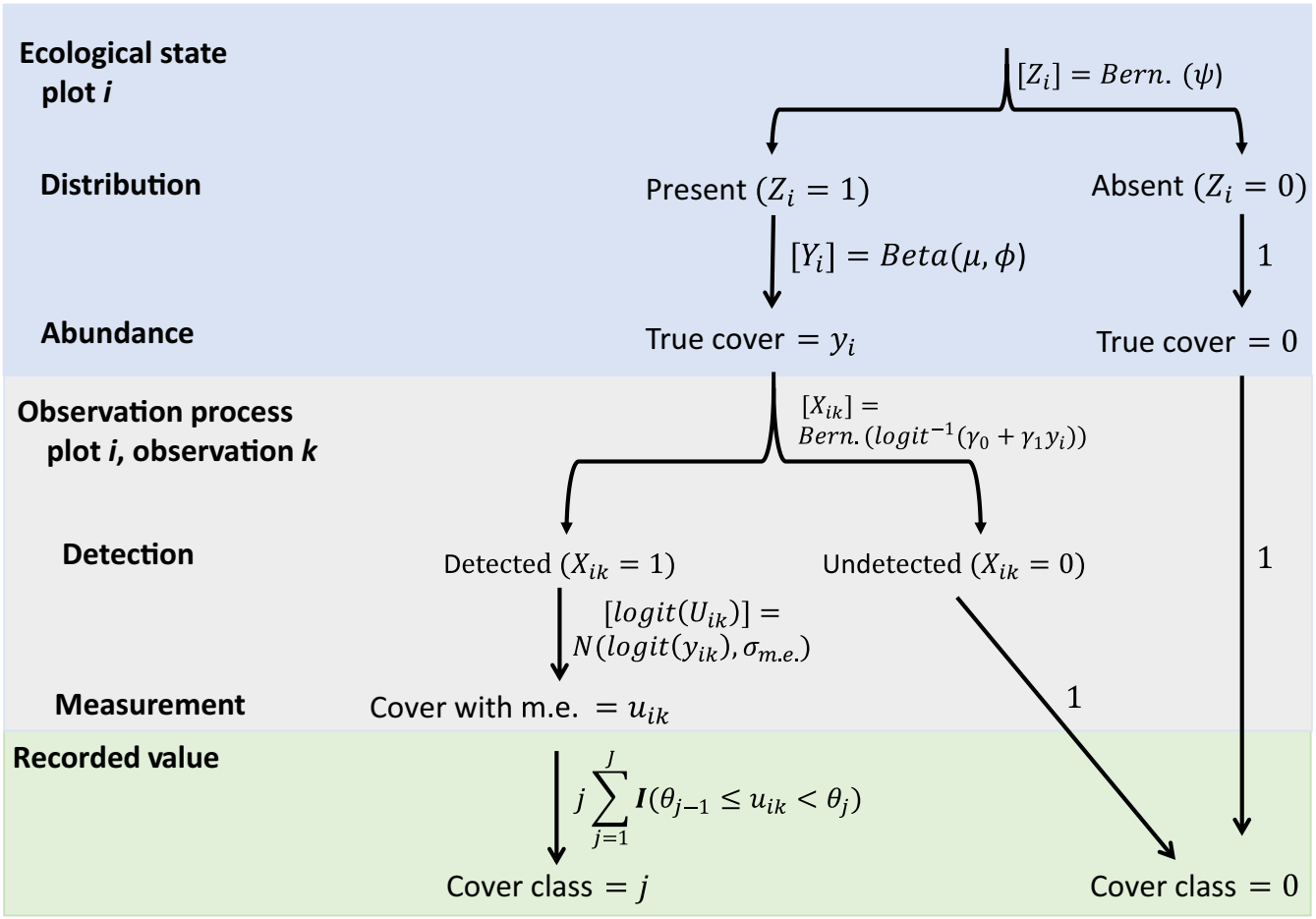


FIGURE 1 Diagram or conceptual model of ordinal zero-augmented beta model with both imperfect detection and measurement errors (OZABE). The right side depicts the zero-process that is contaminated by false zeros in which a species went undetected even though present (X_{ik} = 0 | Z_i = 1) and detection probabilities are dependent on latent true cover y_i. The left side depicts the abundance process that leads to the ordinal cover classes contaminated with measurement errors. The arrows and associated labels indicate the probability distributions for each variable, conditional on the previous variable states. A diagram for the OZAB model is only the top and bottom panel

or beta portion of the model is modified by using the formulation of a zero-augmented beta with errors (ZABE) model developed by Wright et al. (2017). In order to estimate the parameters in OZABE, multiple independent observations *k* must be available for a set of plots *i*, similar to N-mixture or occupancy model data requirements.

We provide a diagrammatic representation of the OZABE model in Figure 1. For simplicity assume only one species is recorded and a simple random sampling design. In the following, we introduce the model hierarchically. We specify that conditional on species presence in a plot, Z_i = 1 and y_i > 0, it is detected on observation *k* with probability p_i, that is

$$[X_{ik}|tY_i = y_i] = \text{Bernoulli}(p_i), \quad (4)$$

where

$$\text{logit}(p_i) = \gamma_0 + \gamma_1 y_i. \quad (5)$$

In Equation (4), X_{ik} represents an indicator variable for whether the species was detected (1) or not (0) in plot *i* on observation *k*. The probability of detection (p_i) is related to the underlying cover proportion (y_i) through a logistic regression in Equation (5) similar to the abundance–detectability relationship described for occupancy models (Royle & Nichols, 2003) and based on empirical evidence found in plant studies (Burg, Rixen, & Stöckli, & Wipf, 2015; Milberg et al., 2008; Vittoz et al., 2010). As percent cover increases, the probability of detection approaches 1 for γ_1 > 0 and for trace amounts detection approaches γ_0, which could depend on plant morphology or other species-specific characteristics.

After accounting for imperfect detection, the final portion of the model describes the observed cover class distribution and allows for measurement error in the classifications. For OZABE, the assumed underlying cover given detection that is discretized is contaminated with measurement error, denoted U_{ik} with [logit(U_{ik})|{X_{ij} = 1, Y_i = y_i}] = Normal(logit(y_{ik}), $\sigma_{m.e.}$) (see Figure 1). Again the recorded cover class (C_{ik}) for plot *i* and observation *k* is a multinomial random variable,

$$[C_{ik}|U_{ik} = u_{ik}] = \text{Multinomial}(1, \boldsymbol{\pi}_i), \quad (6)$$

where $\boldsymbol{\pi}_i$ is a *J*-length vector of probabilities that sum to 1 and correspond to each of the possible cover classes greater than zero (notice this differs slightly from the notation used in the OZAB model). For the OZABE model, the probabilities of observing each non-zero cover class are defined as

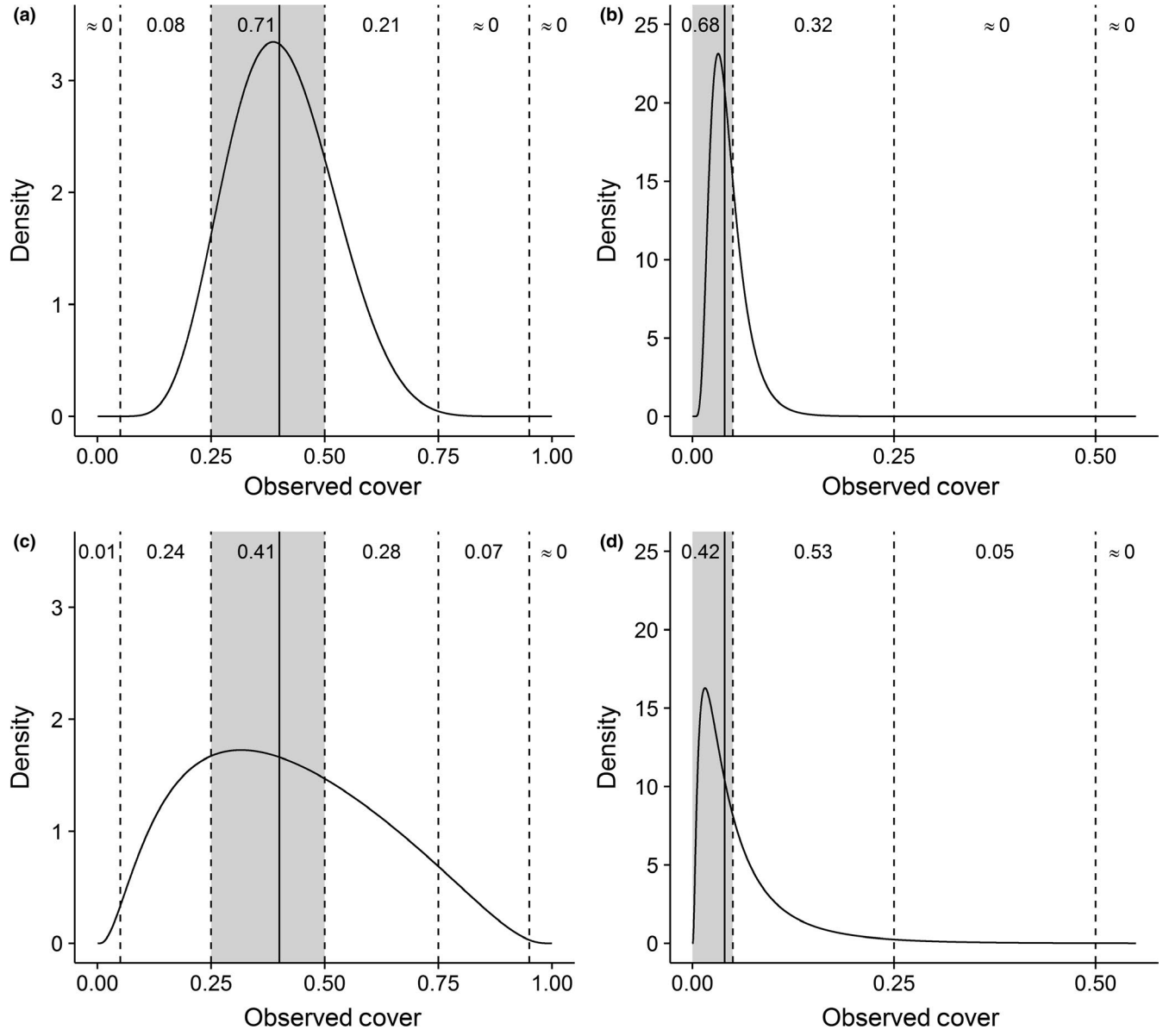


FIGURE 2 Probability density functions for continuous measurement errors when the underlying cover proportion is 0.4 (a, c) or 0.04 (b, d) and the standard deviation of the measurement model ($\sigma_{m.e.}$) is either 0.5 (a, b) or 1.0 (c, d). The vertical solid lines indicate the underlying cover values (y_i) and the vertical dashed lines correspond to the breaks between cover classes using the Daubenmire scale (0.05, 0.25, 0.5, 0.75, 0.95). For each class, the resulting classification probabilities are shown along the top of each plot with the shaded region indicating classification to the correct cover class. Note that displayed probabilities may not sum to one due to rounding and the restricted *x*-axis limits for b and d

$$\pi_{ij} = \begin{cases} F_{N_i}(\text{logit}(\theta_1)) & j = 1 \\ F_{N_i}(\text{logit}(\theta_j)) - F_{N_i}(\text{logit}(\theta_{j-1})) & 2 \le j \le (J-1) \\ 1 - F_{N_i}(\text{logit}(\theta_{j-1})) & j = J \end{cases}, \quad (7)$$

where F_{N_i} represents the Gaussian cumulative distribution function for $\text{logit}(U_{ik})$ with mean $\text{logit}(y_i)$ and standard deviation $\sigma_{m.e.}$ as used to specify measurement error in Wright et al. (2017). This parameterization results in greater variation among observations (U_{ik}) for a given plot *i* for true percent cover values close to 50% (y_i = 0.40 examples in Figure 2a,c). Individual observer effects could be incorporated in the mean of the logit-normal distribution, as follows $\text{logit}(y_{ik} + \alpha\text{Observer}_k)$, to account for systematic observer bias.

For OZABE, increasing measurement error ($\sigma_{m.e.}$) results in decreasing the probability of assigning the correct cover class and increasing the probabilities of assigning an incorrect cover class for a given cover proportion (Figure 2). For example, assume the true cover for a plot is 0.40 (Figure 2a,c), the probability of correctly assigning a cover class of 3 would be 0.71 with $\sigma_{m.e.}$ = 0.5 compared to 0.41 with $\sigma_{m.e.}$ = 1. The larger standard deviation results in an increase in an observer incorrectly assigning an adjacent class (e.g. cover class 2 or 4). Notice that assuming detection is perfect (p_i = 1)

and no measurement error ($\sigma_{m.e.}$ = 0) the OZABE model simplifies to the OZAB model. Under those assumptions, the species is always detected at plots it is present and the observed cover class corresponds to the underlying cover proportion exactly.

To fit models, we used a Bayesian approach in Stan (Carpenter et al., 2016) called from R (R Core Team, 2016) with the RSTAN (Stan Development Team, 2016) package (Appendix S1). We describe our prior specification in the subsequent sections; generally, we used weakly informative priors for all parameters. For each model, parameters were estimated with four Markov chains consisting of 1,000 iterations each after the first 1,000 iterations of each chain were discarded. For the empirical data, we assessed convergence using $\hat{R}$ statistics and traceplots. To make our OZABE and OZAB methods accessible and improved on by others, we provide our model code and empirical datasets at https://doi.org/10.5066/P99JX2ND.

2.3 | Empirical example with multiple species

We analysed multiple observer records of plant species in the Clarno unit of John Day Fossil Beds (hereafter, Clarno) collected as part of the Upper Columbia Basin Inventory and Monitoring Networks' sagebrush steppe vegetation monitoring program (Yeo et al., 2009). We briefly discuss the field methods, see Esposito, Shanahan, and Rodhouse (2016) for further details. Two hundred twenty areal plots (1 × 1 m) were randomly sampled following a generalized random tessellation stratified (GRTS) design (Stevens & Olsen, 2004). At each plot, observers visually estimated the plant cover class using the Daubenmire system for each species detected (Daubenmire 1959). Most plots were visited by a single observer, but two observers independently estimated plant cover in a portion (approximately 30%) of the sampled plots. The multiple observer records allowed us to explore observation error for the various species of interest for this monitoring program.

We included data on four species of conservation concern, two invasive annual grasses (cheatgrass *Bromus tectorum* and medusahead *Elymus caput-medusae*) and two native perennial bunchgrasses *Poa secunda* and *Pseudoroegneria spicata*, and bare ground to demonstrate a joint species model that allows for a zero-augmented beta latent variable and observation errors. The annual species actively invade steppe ecosystems following fire and other disturbances, and are very difficult to eradicate. When abundant, the native perennials, especially *Pseudoroegneria spicata*, can induce invasion resistance and ecological resilience to fire. Hence, modelling patterns of both occurrence (establishment) and percent cover (abundance) for these species over space and time are critical, and influence actions that include habitat management for endangered species (e.g. greater sage-grouse, *Centrocercus urophasianus*) and parks and protected-area restoration.

To estimate the impact of observer errors, we compared inferences from OZABE to OZAB with both models extended to incorporate multiple species. We modified the latent variable Y_i^* by allowing each plant species to have a different zero-augmented beta distribution; $[Y_{i,s}^*] = ZAB(\psi_s, \mu_s, \phi_s)$ with s denoting a particular species. Then the priors on (ψ_s, μ_s) were specified using the non-centresd parameterization as follows Carpenter (2016),

$$\mu_s = \text{logit}^{-1}(\mu + \alpha_s^{std} \times \sigma_\mu), \tag{8}$$

$$\psi_s = \text{logit}^{-1}(\delta + \rho_s^{std} \times \sigma_\psi), \tag{9}$$

where logit^{-1} denotes the inverse-logit transformation with hyper-priors $[\alpha_s^{std}], [\rho_s^{std}] = \text{Normal}(0, 1)$, $[\mu]$, $[\delta]$ =Cauchy(0, 2.5), and $[\sigma_\mu]$, $[\sigma_\psi]$ = half-Cauchy(0, 2.5). Alternatively, a centred parameterization could be used, for example $\mu_s = \text{logit}^{-1}(\alpha_s)$ where $\alpha_s \sim N(\mu, \sigma_\mu)$. For both OZAB and OZABE, we specified independent priors on ϕ_s ($[\phi_s]$ = half-Cauchy(0, 2.5)). Our parameterization is a simplified joint species model that assumes independence among species. Our joint species model is similar to those available in the BORAL package in R that specifies species-specific generalized linear models and allows a range of distributions (e.g. Tweedie, beta, gamma, etc, Hui, 2016), but does not accommodate the ordinal beta formulation we propose for plant cover classes.

Then to extend the OZABE observation process for a joint species model, we expand Equations 4 and 5 by allowing for species-specific baseline detectability with independent priors, $[\gamma_{0,s}]$ = Cauchy(0, 10). We assume a common relationship between detection and abundance ($\gamma_1 = \gamma_{1,s}$ for all species s) and a common measurement error standard deviation ($\sigma_{m.e.}$) across all species with priors $[\gamma_1]$ = Cauchy(0, 5) and $[\sigma_{m.e.}]$ = half-Cauchy(0, 5). These assumptions could be relaxed for other joint species applications by adding complexity into the OZABE observation process (Figure 1 middle panel), e.g. hyper-priors similar to Equations 8 and 9 and/or covariates to explain differences among species detectability. For our dataset, we explored fitting a model with species-specific relationships with abundance ($\gamma_{1,s}$), but the additional complexity led to posterior intervals that were uninformative.

2.4 | Empirical example with hierarchical sampling design

Multiple observer data were also collected in 2015 from Bighorn Canyon National Recreation Area (BICA). Fifty-two plots were located within each of three units (Unit 60, Unit 80, and Unit 90) following a GRTS sampling design within a unit (Esposito et al., 2016). For BICA Unit 60, 12 plots were surveyed by two observers and 13 plots with three observers and Units 80 and 90 had 50 and 51 plots, respectively, surveyed by double observers. We analysed the ordinal cover data for big sagebrush *Artemisia tridentata* collected across three sample frames or management units of interest in BICA.

We specify unit-specific parameters to account for the sampling design instead of species-specific parameters. The same type of model structure can be used to account for sampling designs with a hierarchical structure with primary units (e.g. a unit of interest) and secondary units (e.g. plots within an area). In this case, the subscript s in Equations 8 and 9 denotes each primary unit and i is still the 1-m^2 plot or secondary unit. We assume unit-specific mean proportion cover

(μ_s), dispersion or spatial aggregation (ϕ_s) and occurrence probability (ψ_s). Applying OZABE to this dataset containing only one species (big sagebrush), we assume that the detection process parameters (γ_0 and γ_1) did not vary among the three units. We use similar Cauchy and half-Cauchy priors as for the Clarno joint species model example.

2.5 | Simulation study

We explored potential trade-offs among recording data based on ordinal classes versus a continuous measure of percent cover when both are contaminated by observation errors using simulations. Further, we evaluated the performance (estimation error and interval coverage) of the OZAB and OZABE model under a variety of data generating parameter values. We assume one species was recorded and a simple random sampling design with 100 plots (N = 100) and 3 independent observations (K = 3) to each. The framework for simulating data is conveyed in Figure 1. Data were simulated following the ZABE model based on twelve combinations of ψ = {0.25, 0.7}, γ_0 = { − 1, 0, 1}, and $\sigma_{m.e.}$ = {0.5, 1.0}. The remaining parameters in the model were fixed across the scenarios with μ = 0.25, ϕ = 3, and γ_1 = 3. For each simulated observation with a detection X_{ik} = 1, we saved the continuous 'visually assessed' cover that included measurement error ($u_{ik} \neq y_i$) and the corresponding cover class based on the Daubenmire scale which could include classification error ($[c_{ik} = j|u_{ik}] \neq [c_{ik} = j|y_i]$). We used the ordinal and continuous observations in order to compare whether inferences were affected by partially observing percent cover within an interval.

We fit four models to each simulated dataset based on using the ordinal or continuous measurements and whether the observation errors were ignored or not. For continuous percent cover, we used the zero-augmented beta (ZAB) and zero-augmented beta with errors (ZABE) models proposed in (Wright et al., 2017). We fit the ZAB and OZAB models to the first observation (k = 1) at each simulated plot in the continuous and ordinal datasets respectively to reflect analyses that assume no observation errors. For comparison, we analysed all the simulated continuous and ordinal observations at each plot using the ZABE and OZABE models respectively. We simulated 250 datasets under every scenario described previously.

We assessed convergence of each model by checking $\hat{R}$ statistics and only summarized results from models with all $\hat{R}$ values less than 1.1. Additionally, we excluded any results from models that issued a 'divergent transition' warning from Stan which can indicate problems with the sampler. Converged models that did not issue warnings were summarized by plotting the average posterior means and average 95% posterior intervals (PIs) for the ecological parameters (ψ, μ, ϕ). We consider estimation error, the difference between the average posterior means and data generating parameter values, as a measure of bias for each model. For a model fit to a simulated dataset, the average 95% PIs consisted of the average 2.5% and 97.5% posterior quantiles for the parameter specified. We summarized results by determining the coverage properties of the 95% posterior interval (number of times the posterior interval contained the data generating value) under the different scenarios for all four models.

3 | RESULTS

3.1 | Joint species ordinal zero-augmented beta model with and without observation errors

For the five species we analysed, bare ground cover had the most discrepancy among observers (Table 1). Cheatgrass cover had the most measurement errors and the remaining three species had relatively little evidence of observer variation. For this particular dataset and survey protocol there was little difference between the models with and without observation errors (Figure 3), with the exception of bare ground. This finding is consistent with the general agreement among observers and high estimated probability of detections (Table 1). The estimated measurement error

TABLE 1 Summary of observer variation in Clarno and BICA during 2015. Clarno include four species and bareground. BICA data included only big sagebrush *Artemisia tridentata* and three units or areas of interest. A detection error occurred when one observer recorded a cover class of 1 or higher and the other observer recorded a zero (detect. error column). The cover class discrepancy column was the number of times observers recorded a different non-zero cover class. Agreement was simply the number of times both observers recorded the same cover class. The estimated probability of detection (Prob. Detection) given 1% proportion cover ($\hat{\gamma}_0 + \hat{\gamma}_1 \times .01$) was based on fitting the OZABE model to Clarno data. The BICA analysis assumed a constant probability of detection for sagebrush among the units

	Detect. error	Cover class discrep.	Agreement	Prob. detection
Clarno species				
Bareground	13	15	32	0.86 (0.79, 0.92)
Bromus Tectorum	2	23	35	0.96 (0.90, 0.99)
Elycap medusae	5	4	51	0.89 (0.77, 0.96)
Poa secunda	4	3	53	0.92 (0.82, 0.97)
Pseudoroegneria spicata	2	7	51	0.96 (0.85, 0.99)
BICA unit				
60	1	0	24	0.89 (0.76, 0.96)
80	6	6	38	–
90	0	3	48	–

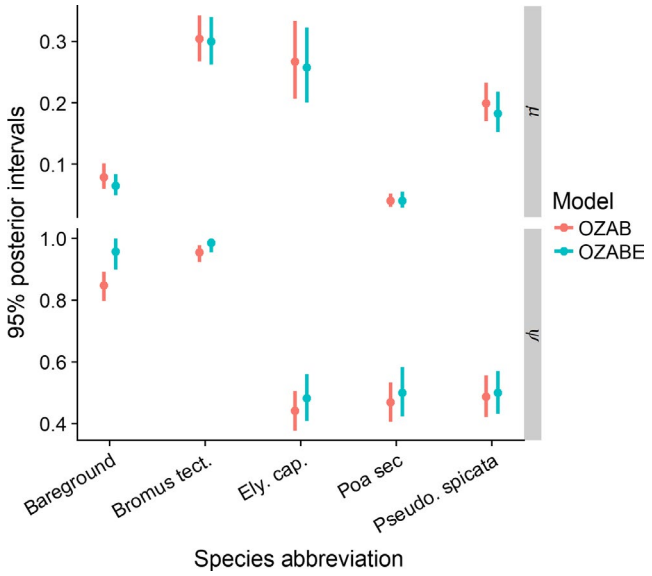


FIGURE 3 Posterior intervals for mean proportion cover (μ_s) and probability of occurrence (ψ_s) for OZAB and OZABE joint species or community model applied to data from 2015 collected in Clarno unit of John Day Fossil Beds National Monument. OZAB and OZABE allow for species-specific dispersion of percent cover (ϕ_s) and baseline probability of detection ($\gamma_{0,s}$). Posterior quantiles for all parameters provided in Tables S2.1 and S2.2

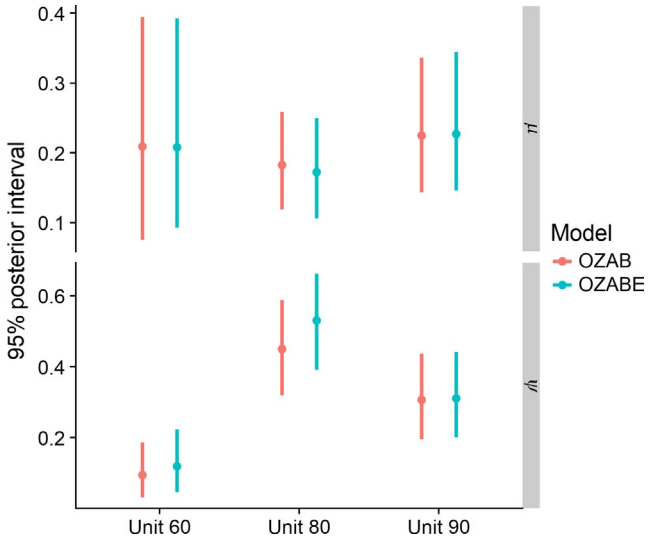


FIGURE 4 Sagebrush (*Artemisia tridentata*) posterior intervals for mean proportion cover (μ_s) and probability of occurrence (ψ_s) for OZAB and OZABE by unit. Three units (Unit 60, 80, and 90) were surveyed during 2015 at Big Horn Canyon National Monument with 52 1 − *m*x1 − *m* plots within each. OZAB and OZABE allow for varying dispersion (ϕ) and assume a common baseline detection (γ_0), relationship between detection and latent abundance (γ_1), measurement error ($\sigma_{m.e.}$). Posterior quantiles for all parameters provided in Tables S2.3 and S2.4

standard deviation was 0.64 (95% posterior interval 0.54, 0.76) pooled across all species, which suggests an observer was most likely to assign the correct cover class and to a lesser degree incorrectly assign an adjacent category (Figure 2a,b).

3.2 | Hierarchical sampling design ordinal zero-augmented beta model with and without error

The mean estimated percent cover for sagebrush across all three units was generally ~20% (Figure 4 top panel). There was evidence of a difference among units in terms of the prevalence or occurrence of sagebrush (Figure 4 bottom panel). For this particular dataset, there was even less evidence for observation errors and the results were similar between OZAB and OZABE. However, unit 80 had the highest degree of variation among observers and, as expected, the OZAB estimate for occurrence was less than and mean abundance was greater than the OZABE estimates. The estimated detection probability assuming 1% cover was 0.89 (95% posterior interval 0.76 to 0.96), which is quite high and increases as percent cover increases. The estimated measurement error variance for sagebrush was 0.51 (95% posterior interval 0.33 to 0.76), similar to the other species analysed.

3.3 | Effects of observation errors on parameter estimation with ordinal and continuous percent cover data

In the simulation study, the ordinal models were more difficult to fit and had divergent transitions or failed to converge more frequently than the continuous models (Table S2.5). These issues were more prevalent in the scenarios with lower occurrence probability (ψ = 0.25) and when detection probabilities (γ_0) were lower. These simulation results were consistent with empirical examples (results not shown). The model fits with these issues were excluded from the remaining summaries because posterior estimates were considered unreliable and, therefore, were not an accurate representation of model performance.

Overall, models that ignored errors (OZAB or ZAB), regardless of whether the ordinal or continuous responses were analysed, showed more estimation error and reduced coverage than their counterparts that included the observation process (OZABE or ZABE). The mean proportion cover, μ, was over-estimated on average using OZAB and ZAB models across all scenarios (Figure 5a). The positive bias was due to the relationship between cover and detectability (Wright et al., 2017), and the magnitudes of bias were similar for both models that ignore observation errors (OZAB and ZAB). On the other hand, ignoring the observation process led to under-estimating probability of occurrence (ψ) using OZAB and ZAB and the magnitude of estimation error was similar in the scenarios we examined (Figure 5b). For both of these parameters (μ, ψ) estimation error using OZAB and ZAB models was reduced as the probability of detection increased (γ_0 = {−1, 0, 1}).

Accounting for the observation process using either the OZABE or ZABE models resulted in no estimation error for both parameters for the most part (Figure 5). Fitting these models to the scenario with ψ = 0.25 and γ_0 = −1, however, resulted in the posterior means, on average, over-estimating probability of occurrence (ψ). We believe this bias is due to the overall data sparseness in this scenario as well as having to exclude many of these model fits due to divergent

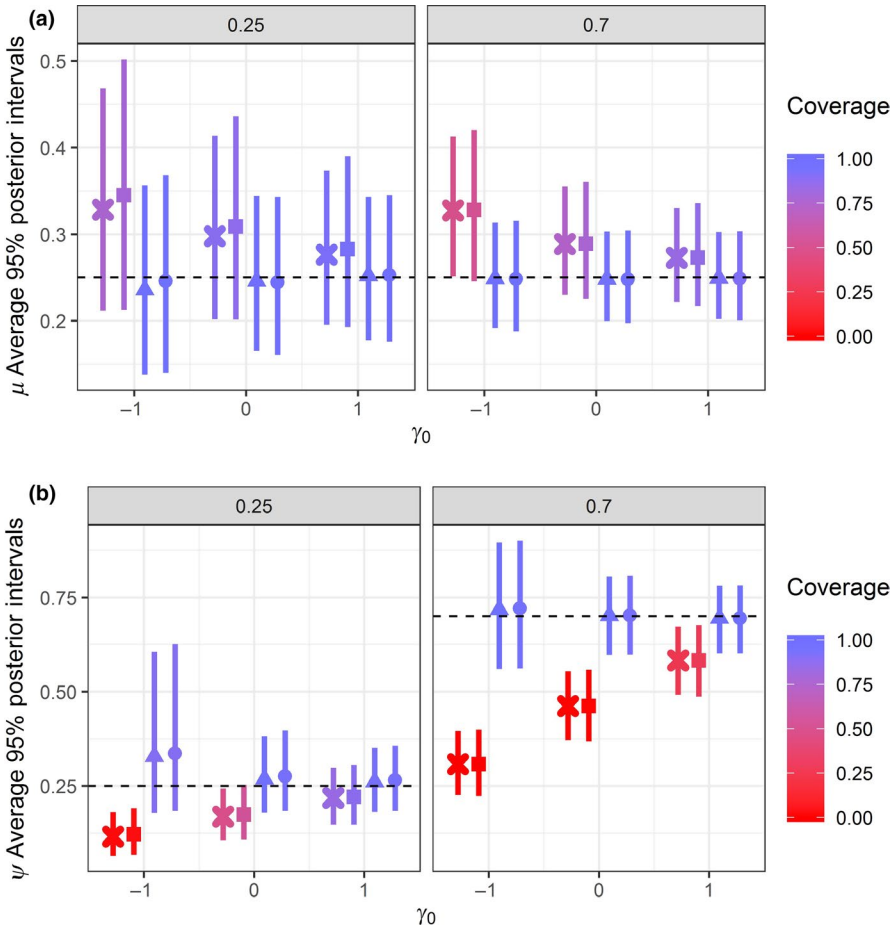


FIGURE 5 Average posterior intervals from simulated datasets under different parameter combinations for ψ = {0.25, 0.7} (correspond to columns) and levels of detection γ_0 = {−1, 0, 1} (*x*-axis reference). Point shapes represent the different models with X for ZAB, square for OZAB, triangle for ZABE, and circle for OZABE. All scenarios assumed 100 plots with 3 observers to each, mean proportion cover μ = 0.25, dispersion ϕ = 3, relationship between detection and abundance γ_1 = 3 and measurement error standard deviation $\sigma_{m.e.}$ = 0.5. The patterns were similar for $\sigma_{m.e.}$ = 1.0. The assumed data generating parameter values were similar to estimates from the empirical examples. Colours show the coverage rates for the true parameter values. (a) Average posterior intervals for mean proportion cover (μ); (b) average posterior intervals for probability of occurrence (ψ)

transitions or convergence issues. Again, when accounting for the observation process, the average posterior means and average 95% PIs were similar regardless of whether the ordinal (OZABE) or continuous (ZABE) responses were analysed.

For the final ecological parameter, precision (ϕ) or a measure of spatial aggregation $\left(\delta = \left(\frac{1}{1+\phi}\right)\right)$, inferences were not impacted by ignoring errors in the case of $\sigma_{m.e.}$ = 0.50 and all models had similar average 95% PIs (Fig. S2.1). Interval estimates for ϕ were, on average, narrower for the scenarios with a higher probability of occurrence (ψ = 0.7) compared to when occurrence was lower (ψ = 0.25). However, increasing the measurement error ($\sigma_{m.e.}$ = 1.0) led to wider intervals and the potential magnitude of estimation error was more noticeable (Fig. S2.1). In that, OZABE and ZABE over-estimated and ZAB and OZAB under-estimated ϕ, on average, which would correspond to the converse for estimating spatial aggregation.

4 | DISCUSSION

We found estimated detection probabilities for most species in our datasets were high (>0.85, Table 1) and measurement errors were moderate ($\sigma_{m.e.} \approx 0.50$). The measurement error estimate was consistent with the finding that most observers differed by only one cover class (Esposito et al., 2016). These estimates corresponded to the examples shown in Figure 2 (panels a and b), notice that the measurement error portion of OZABE results in an increased probability of an incorrectly assigned adjacent cover class value due to observation error. The magnitude of $\sigma_{m.e.}$ should increase if observers differed by more than one class. For plant species, the magnitude of observation errors likely depends on plot size, habitat type, species of interest, and other characteristics that vary across studies. The data explored here, for instance, were collected using quite small (1 × 1 m) plots that could in part explain the high detection probabilities. Additionally, observers were trained in plant identification and using cutouts with known percent cover prior to the field season for calibration (Yeo et al., 2009). However, because there was some evidence to suggest detectability was related to cover, if abundance increases or decreases in a study area over time, detectability will also vary and cannot necessarily be ignored when estimating temporal trends in plant surveys.

In simulations, high detection probabilities (corresponding to γ_0 = 1) led to small differences among OZABE and OZAB, in fact our empirical datasets had a minimum estimated value of $\hat{\gamma}_0 = 1.8$ for bare ground at Clarno. Our simulations showed that the OZABE model is most beneficial when the probability of detection is low. However, without directly comparing the OZABE and OZAB models, it is impossible to know the extent to which ignoring observation errors could impact ecological inferences. Ignoring observation errors in the analysis of ordinal plant cover data could result in biased

point estimates and reduced coverage for interval estimates as our simulations showed. Even though our empirical results showed little impact of observation errors, we caution that imperfect detection and measurement errors have been widely reported in other vegetation studies (reviewed in Morrison, 2016).

Interestingly, in our simulation study, ordinal measurements produced similar estimates for mean proportion cover and occurrence to those of continuous measurements with or without errors. These results suggest that using the Daubenmire classification scheme does not compromise statistical accuracy or precision in comparison to estimating continuous percent cover when using our models. In our simulations, the assumed variation in percent cover among observers resulted in discrepancies among assigned cover classes as well and only directly modelling the observation process resulted in unbiased estimates of cover. Sometimes the motivation for using an ordinal scale to estimate cover, in addition to saving time, is that it may lead to less observer variation compared to estimating a continuous percent cover for a plot. Our datasets cannot specifically address this claim because ordinal values were only recorded and our simulations assumed that the magnitude of variation in continuous cover among observers resulted in the potential for an observer to assign an incorrect cover class. However, our modelling framework can be leveraged by others to determine and, importantly, quantify sources of observation errors in plant surveys.

Although we did not explore different sampling efforts in our simulation study, the general guidance provided for occupancy models would apply for our models (MacKenzie et al., 2018). Specifically, species with lower detection probabilities ($p < .50$) require a larger number of independent observations ($K > 3$). For our empirical datasets, two observers was sufficient because of high detectability, but in simulated scenarios with low detection probability ($p \approx .27$) the occurrence probability posterior intervals were substantially wider (Figure 5b) and would likely decrease with more survey effort. An additional complication with ordinal data is that sparse representation of some cover classes can result in estimation instability. For example, the number of simulated datasets that resulted in unreliable OZAB or OZABE posterior estimates was substantially greater when occurrence probabilities were 0.25 (Table S2.5). A possible remedy is to consider collapsing categories to increase the number of observations in each cover class, but further work exploring optimal survey designs and alternative classification schemes would be beneficial.

Our hierarchical OZAB(E) framework can be modified to include correlations among species by specifying a multivariate normal prior with non-zero off-diagonal elements. In our examples, percent cover of all species within a plot was not constrained to equal 100% and not all species were recorded within a plot. Further study is warranted into a comparison of our joint species model formulation to other approaches (Damgaard, 2015; Schliep et al., 2016), particularly for datasets that record all species within a plot and have the constraint that percent cover equals 100%. The OZAB has additional flexibility to explore how explanatory variables may influence the precision by using a log-link (e.g. $\log(\phi) = X\beta$ Cribari-Neto & Zeileis, 2010), although we did not exploit this property for our datasets. A log-link could be specified instead of the logit-link on mean proportion cover ($\log(\mu)$) for studies interested in species–environment relationships. The interpretation of a species association with the environment when using a log-link is as a multiplicative change in mean proportion cover. Finally, for any analysis, model assessments should be considered to interrogate data-model agreement and can also inform the appropriate link functions to use (e.g. logit vs. log). Our proposed Bayesian framework allows for latent residual assessments as suggested for occupancy models (Wright, Irvine, & Higgs, 2019). Latent beta residuals defined as for beta regression (Espinheira, Ferrari, & Cribari-Neta, 2008) could be beneficial for identifying model inadequacies in the abundance component of OZAB(E).

Our proposed framework for analysing cover class data holds appeal for two main interpretational benefits compared to other ordinal models. One is the ability to separate distribution from abundance allowing for deeper ecological insight. For example, in another analysis we found evidence that cheatgrass was less likely to occur in steep north facing areas, whereas abundance was higher in flatter areas of John Day Fossil Beds. These statistical associations suggest potential ways for managers to prioritize areas in the monument for restoration activities to contain the current invasion or early detection efforts to prevent invasion (Rodhouse et al., 2014). The other benefit is by linking the cover classes to a latent beta distribution, we can interpret and visualize parameter point and interval estimates in terms of mean percent cover – the biological variable of interest. In the OZABE model, the observation process is linked to the underlying beta distributed cover values, therefore, after modelling the observation process, inferences are still in terms of mean percent cover. Even if different ordinal cover class systems were used, the results for the ecological abundance process (i.e. proportion cover parameters μ and ϕ) are still comparable across studies providing a unique opportunity for integration and synthesis among datasets, if appropriate. The flexibility of the OZABE framework to account for survey design features (observation errors, hierarchical sampling design, and multiple species) while maintaining a connection to the ecological process of interest provides an appealing alternative for plant cover class analyses.

ACKNOWLEDGEMENTS

We appreciate two anonymous reviewers and the Associate Editor for their constructive comments and insights. Christian Damgaard and Hans Van Calster provided valuable discussions on the statistical analyses of multivariate ordinal cover data. We thank Rob Daly for data management assistance, and Kristin Legg and Gordon Dicus for their continued support for advancing statistical methods for monitoring programs. Dr. K. Irvine's participation was supported by the US Geological Survey, National Park Service (IAA P12PG70586), and US Fish and Wildlife Service (IAA 450063656). Wilson Wright was supported by Montana State University and National Park Service agreement P17AC00623. Any use of trade, firm, or product names

is for descriptive purposes only and does not imply endorsement by the US Government.

AUTHORS' CONTRIBUTIONS

K.M.I conceived the original OZAB(E) approach, extended the OZABE code and analysed the empirical datasets, and led writing of the manuscript; W.J.W. authored original OZABE code and conducted simulation study and contributed revisions to the manuscript; T.J.R. collaborated on the original approach and ecological interpretation of the models and contributed critical revisions to the manuscript; E.K.S had substantial contribution to acquisition and interpretation of data and assisted with writing the manuscript. All authors gave final approval for publication.

DATA AVAILABILITY STATEMENT

All JAGS and Stan model code and datasets are archived at https://doi.org/10.5066/P99JX2ND. The data are publicly available at https://irma.nps.gov/DataStore/Reference/Profile/2263575.

ORCID

Kathryn M. Irvine https://orcid.org/0000-0002-6426-940X
Thomas J. Rodhouse https://orcid.org/0000-0001-5953-9113

REFERENCES

Agresti, A. (2010). *Analysis of ordinal categorical data*. Hoboken, NJ: John Wiley and Sons.

Burg, S., Rixen, C., Stöckli, V., & Wipf, S. (2015). Observation bias and its causes in botanical surveys on high-alpine summits. *Journal of Vegetation Science*, *26*(1), 191–200. https://doi.org/10.1111/jvs.12211

Carpenter, B. (2016). Hierarchical partial pooling for repeated binary trials. Technical report. Retrieved from https://mc-stan.org/users/documentation/case-studies/pool-binary-trials.html.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2016). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1):1–32.

Chen, G., Kéry, M., Plattner, M., Ma, K., & Gardner, B. (2013). Imperfect detection is the rule rather than the exception in plant distribution studies. *Journal of Ecology*, *101*(1), 183–191. https://doi.org/10.1111/1365-2745.12021

Chen, G., Kéry, M., Zhang, J., & Ma, K. (2009). Factors affecting detection probability in plant distribution studies. *Journal of Ecology*, *97*(6), 1383–1389. https://doi.org/10.1111/j.1365-2745.2009.01560.x

Clark, J. S., Nemergut, D., Seyednasrollah, B., Turner, P. J., & Zhang, S. (2017). Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious. *Ecologial Monographs*, *87*(1), 34–56. https://doi.org/10.1002/ecm.1241

Clarke, K. D., Lewis, M., Brandle, R., & Ostendorf, B. (2012). Non-detection errors in a survey of persistent, highly-detectable vegetation species. *Environmental Monitoring and Assessment*, *184*(2), 625–635. https://doi.org/10.1007/s10661-011-1991-0

Cribari-Neto, F., & Zeileis, A. (2010). Beta regression in R. Retrieved from https://cran.r-project.org/web/packages/betareg/vignettes/betareg.pdf.

Damgaard, C. F. (2013). Hierarchical and spatially aggregated plant cover data. *Ecological Informatics*, *18*, 35–39. https://doi.org/10.1016/j.ecoinf.2013.06.001

Damgaard, C. (2015). Modelling pin-point cover of complementary vegetation classes. *Ecological Informatics*, *30*, 179–184. https://doi.org/10.1016/j.ecoinf.2015.10.006

Damgaard, C. F., & Irvine, K. M. (2019). Using the beta distribution to analyze plant cover data. *Journal of Ecology*. https://doi.org/10.1111/1365-2745.13200.

Daubenmire, R. F. (1959). A canopy-coverage method of vegetational analysis. *Northwest Science*, *33*(1), 43–64.

Espinheira, P. L., Ferrari, S. L. P., & Cribari-Neta, F. (2008). On Beta regression residuals. *Journal of Applied Statistics*, *35*(4), 407–419. https://doi.org/10.1080/02664760701834931

Esposito, D. M., Shanahan, E., & Rodhouse, T. J. (2016). UCBN and GRYN sagebrush steppe vegetation monitoring; double observer study 2015: John Day Fossil Beds National Monument - Clarno Unit and City of Rocks National Reserve. Technical report,National Park Service, Fort Collins, Colorado: Natural Resources Data Series, NPS/UCBN/NRDS-2016/1052.

Ferrari, S. L. P., & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*., *31*(7), 799–815.

Gorrod, E. J., Bedward, M., Keith, D. A., & Ellis, M. V. (2013). Systematic underestimation resulting from measurement error in score-based ecological indices. *Biological Conservation*, *157*, 266–276. https://doi.org/10.1016/j.biocon.2012.09.002

Helm, D. J., & Mead, B. R. (2004). Reproducibility of vegetation cover estimates in south-central Alaska forests. *Journal of Vegetation Science*, *15*(1), 33–40. https://doi.org/10.1111/j.1654-1103.2004.tb02234.x

Herpigny, B., & Gosselin, F. (2015). Analyzing plant cover class data quantitatively: Customized zero-inflated cumulative beta distributions show promising results. *Ecological Informatics*, *26*, 18–26. https://doi.org/10.1016/j.ecoinf.2014.12.002

Hui, F. (2016). BORAL – Bayesian ordination and regression analysis of multivariate abundance data in R. *Methods in Ecology and Evolution*, *7*, 744–750. https://doi.org/10.1111/2041-210x.12514

Irvine, K. M., Rodhouse, T., & Keren, I. N. (2016). Extending ordinal regression with a latent zero-augmented beta distribution. *Journal of Agricultural, Biological and Environmental Statistics*, *21*(4), 619–640. https://doi.org/10.1007/s13253-016-0265-2

Kéry, M., & Royle, J. A.. (2016). *Applied hierarchical modeling in ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and static models*. San Diego, CA: Academic Press.

Killourhy, C. C., Crane, D., & Stehman, S. V. (2016). Precision and accuracy of visual estimates of aquatic habitat. *Freshwater Science*, *35*(3), 1062–1072. https://doi.org/10.1086/687844

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2018). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence* (2nd ed.). San Diego, CA: Academic Press.

Milberg, P., Bergstedt, J., Fridman, J., Odell, G., & Westerberg, L. (2008). Observer bias and random variation in vegetation monitoring data. *Journal of Vegetation Science*, *19*(5), 633–644. https://doi.org/10.3170/2008-8-18423

Morrison, L. W. (2016). Observer error in vegetation surveys: A review. *Journal of Plant Ecology*, *9*(4), 367–379. https://doi.org/10.1093/jpe/rtv077

Ospina, R., & Ferrari, S. L. P. (2010). Inflated beta distributions. *Statistical Papers*, *51*, 111–126.

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 (DSC 2003)* (pp. 1–10). Vienna, Austria. Retrieved from https://www.r-project.org/conferences/DSC-2003/Proceedings/

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rodhouse, T., Irvine, K. M., Sheley, R., Smith, B., Hoh, S., Esposito, D., & Mata-Gonzalez, R. (2014). Predicting foundation bunchgrass species abundances: Model-assisted decision-making in protected-area sagebrush steppe. *Ecosphere*, *5*(9), 108. https://doi.org/10.1890/ES14-00169.1

Royle, J. A., & Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, *84*(3), 777–790. https://doi.org/10.1890/0012-9658(2003)084[0777:eafrpa]2.0.co;2

Schliep, E. M., Lany, N. K., Zarnetske, P. L., Schaeffer, R. N., Orians, C. M., Orwig, D. A., & Preisser, E. L. (2016). Joint species distribution modelling for spatio-temporal occurrence and ordinal abundance data. *Global Ecology and Biogeography*, *27*, 142–155. https://doi.org/10.1111/geb.12666

Stan Development Team. (2016). RSTAN: the R interface to Stan. Retrieved from http://mc-stan.org.

Stevens, D. L. Jr, & Olsen, A. R. (2004). Spatially balanced sampling of natural resources. *Journal of American Statistical Association*, *99*(465), 262–278. https://doi.org/10.1198/016214504000000250

Tercek, M., Jean, C., Daley, R., & Legg, K. (2013). Greater yellowstone network upland vegetation monitoring protocol: Narrative version 1.0. Natural Resource Report NPS/GRYN/NRR 2013/623. National Park Service, Fort Collins, Colorado.

Vittoz, P., Bayfield, N., Brooker, R., Elston, D. A., Duff, E. I., Theurillat, J.-P., & Guisan, A. (2010). Reproducibility of species lists, visual cover estimates and frequency methods for recording high-mountain vegetation. *Journal of Vegetation Science*, *21*(6), 1035–1047. https://doi.org/10.1111/j.1654-1103.2010.01216.x

Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. (2015). So many variables: Joint modeling in community ecology. *Trends in Ecology and Evolution*, *30*(12), 766–779. https://doi.org/10.1016/j.tree.2015.09.007

Warton, D. I., Foster, S. D., & De'ath, G., Stoklosa, J., & Piers, K. D. (2015). Model-based thinking for community ecology. *Plant Ecology*, *216*, 669–682. https://doi.org/10.1007/s11258-014-0366-3

Wright, W. J., Irvine, K. M., & Higgs, M. D. (2019). Identifying occupancy model inadequacies: Can residuals separately assess detection and presence? *Ecology*, *100*(6), e02703. https://doi.org/10.1002/ecy.2703

Wright, W. J., Irvine, K. M., Warren, J. M., & Barnett, J. K. (2017). Statistical design and analysis for plant cover studies with multiple sources of observation errors. *Methods in Ecology and Evolution*, *8*, 1832–1841. https://doi.org/10.1111/2041-210x.12825

Yeo, J., Rodhouse, T., Dicus, G., Irvine, K., & Garrett, L. (2009). Upper Columbia Basin Network Sagebrush Steppe Vegetation Monitoring Protocol: Narrative and Standard Operating Procedures (Appendix to Narrative) Version 1.0. Natural Resource Report NPS/UCBN/NRR 2009/142 Published Report-662446. National Park Service, Fort Collins, Colorado.

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: Irvine KM, Wright WJ, Shanahan EK, Rodhouse TJ. Cohesive framework for modelling plant cover class data. *Methods Ecol Evol*. 2019;10:1749–1760. https://doi.org/10.1111/2041-210X.13262