

RESEARCH ARTICLE

Propagating observation errors to enable scalable and rigorous enumeration of plant population abundance with aerial imagery

Andrii Zaiats¹  | T. Trevor Caughlin¹  | Jennyffer Cruz¹  | David S. Pilliod²  |
Megan E. Cattau³  | Rongsong Liu⁴ | Richard Rachman¹  | Maisha Maliha⁵  |
Donna Delparte⁶  | John D. J. Clare^{1,7} 

¹Department of Biological Sciences, Boise State University, Boise, Idaho, USA; ²U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Boise, Idaho, USA; ³Human-Environment Systems, Boise State University, Boise, Idaho, USA; ⁴Department of Mathematics and Statistics, University of Wyoming, Laramie, Wyoming, USA; ⁵Department of Computer Science, Boise State University, Boise, Idaho, USA; ⁶Department of Geosciences, Idaho State University, Pocatello, Idaho, USA and ⁷Department of Forest and Wildlife Ecology, University of Wisconsin–Madison, Madison, Wisconsin, USA

Correspondence

John D. J. Clare

Email: clare.john@peregrinefund.org

Funding information

National Aeronautics and Space Administration, Grant/Award Number: #80NSSC21K1638; Directorate for Biological Sciences, Grant/Award Number: BIO-2207158; National Science Foundation, Grant/Award Number: OIA-1757324

Handling Editor: Hooman Latifi

Abstract

1. Estimating and monitoring plant population size is fundamental for ecological research, as well as conservation and restoration programs. High-resolution imagery has potential to facilitate such estimation and monitoring. However, remotely sensed estimates typically have higher uncertainty than field measurements, risking biased inference on population status.
2. We present a model that accounts for false negative (missed plants) and false positive (misclassified or double-counted plants) error in counts from high-resolution imagery via integration with ground data. We apply it to estimate the abundance of a foundational shrub species in post-wildfire landscapes in the western United States. In these landscapes, plant recruitment is crucial for ecological recovery but locally patchy, motivating the use of spatially extensive measurements from unoccupied aerial systems (UAS). Integrating >16 ha of UAS imagery with >700 georeferenced field plots, we fit our model to generate insights into the prevalence and drivers of observation errors associated with classification algorithms used to distinguish individual plants, relationships between abundance and landscape context, and to generate spatially explicit maps of shrub abundance.
3. Raw counts of plant abundance in high-resolution imagery resulted in substantial false negative and false positive observation errors. The probability of detecting (p) adult plants (≥ 0.25 m tall) varied between sites within $0.52 < \hat{p}_{\text{adult}} < 0.82$, whereas the detection of smaller plants (< 0.25 m) was lower, $0.03 < \hat{p}_{\text{small}} < 0.3$. On average, we estimate that 19% of all detected plants were false positive errors, which varied spatially in relation to topographic predictors. Abundance

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

declined toward the interior of previous wildfires and was positively associated with terrain roughness.

4. Our study demonstrates that integrated models accounting for imperfect detection improve estimates of plant population abundance derived from inherently imperfect UAS imagery. We believe such models will further improve inference on plant population dynamics—relevant to restoration, wildlife habitat and related objectives—and echo previous calls for remote sensing applications to better differentiate between ecological and observational processes.

KEYWORDS

abundance, aerial surveys, imperfect detection, misclassification, population monitoring, remote sensing, UAS

1 | INTRODUCTION

Monitoring plant species distribution and abundance is critical for both fundamental research and adaptive land management (McCord & Pilliod, 2021). For example, the abundance of invasive plant species can foreshadow leading edge invasion of neighbouring areas (Ibáñez et al., 2009). Similarly, changes in the density of a plant species following restoration can inform restoration effectiveness or adaptive management decision-making (Caughlin et al., 2019). While field-based counts are fundamental data for estimating species distributions, population trends and demographic parameters (Gurevitch et al., 2016), collecting these data is logistically difficult and resource intensive at large spatial extents relevant to many ecological problems.

Recent advances in remotely sensed data, particularly high-resolution imagery where individual plants can be detected, enables scalable estimates of species abundance (Rominger & Meyer, 2019; Weinstein et al., 2020; Young et al., 2022). Individual tree crowns have been delineated in both high-resolution satellite imagery (Brandt et al., 2020) and high-resolution aerial imagery (Young et al., 2022). These advances present exciting opportunities in ecology and conservation: for example, recent research has revealed that trees outside forests contribute substantially to national-level carbon storage (Liu et al., 2023). Mapping individual plants using remote sensing can provide new perspectives on long-standing questions like the role of negative density dependence on tree population dynamics (Kellner & Hubbell, 2018), and facilitate efforts to monitor endangered populations in fragile or inaccessible landscapes (Rominger & Meyer, 2019).

Delineating individuals to quantify abundance from high-resolution imagery comes with risk of observation error from atmospheric distortion, crown overlap, sensor limitations and other factors (Davis et al., 2022). Counts of individual plants in high-resolution imagery likely exhibit both false negative error (an individual is not detected within the image) and false positive error (a non-target object is mistaken as a target individual, or the crown of a single individual is mistakenly classified as multiple individuals). These errors are well-acknowledged within wildlife surveys

performed in-situ or using remote sensing, and several statistical models have been described to account for these errors when using varied data types to estimate different ecological quantities (Chambert et al., 2015; Clare et al., 2021; Doser et al., 2021; Kéry & Royle, 2020; Miller et al., 2011). Although observational errors can bias studies focusing on plant occurrence or abundance (e.g. Chen et al., 2013; Louthan & Doak, 2018; Perret et al., 2023), these errors are rarely addressed within studies using imagery to study plants.

Models that disentangle false positive and false negative error processes from an ecological process of interest typically require some form of data replication (Chambert et al., 2015; Clement et al., 2022; Conn et al., 2014; Miller et al., 2011). A type of replicated data that is relatively common in maps of plant abundance and could be used to inform observation error is auxiliary 'ground-validated' data (georeferenced ground points indicating an individual's location). Integrating relatively accurate but spatially limited ground-validated counts with remotely sensed counts that are less accurate but highly scalable provides a means to address multiple types of error and borrow from the strengths of both approaches (Miller et al., 2019). Here, we develop two models that leverage ground-validated counts within an integrated framework to estimate plant population abundance with appropriate uncertainty when remotely sensed count data are subject to false negative and false positive errors.

We demonstrate our model by estimating the abundance of a plant species with spatially variable recruitment. In our study region, novel wildfire regimes threaten many plant and animal species, including our focal species, big sagebrush (*Artemisia tridentata*; hereafter, sagebrush). The abundance of sagebrush recruits is an important indicator of post-fire recovery (Germino et al., 2018; Schlaepfer et al., 2014). However, sagebrush recruitment after fire is difficult to quantify in the field because it is spatially variable and sensitive to fine-grained environmental variation (Arkle et al., 2022; O'Connor et al., 2020; Ziegenhagen & Miller, 2009), motivating our use of remotely sensed data to detect recruitment hotspots. Our objectives for this application were to estimate total sagebrush abundance across our field sites, clarify key drivers of abundance (e.g. topography and distance to fire edges) and present initial drivers of observation errors in plant focused UAS

surveys. More broadly, our work aims to establish the importance of hierarchical models that distinguish ecological and observation processes when estimating abundance of plant species from high-resolution imagery.

2 | MATERIALS AND METHODS

2.1 | Problem overview

We first envision a situation where only error-prone count data from imagery are available. Let N_i denote the abundance of a species across a collective of H sampling units: a set of discrete and equally sized areal units (cells) within an image (or multiple images), where i indexes specific units ($i=1, 2, \dots, H$). Assume $N_i \sim \text{Poisson}(\lambda^{TP})$ and that the latent abundance states or their expected value λ^{TP} are the focus of interest. The Poisson distribution is a natural choice for quantities like abundance, but this distributional assumption is not strict. In each unit, a count c_i^* is recorded. In many remote sensing applications, it is common to assume that $c_i^* = N_i$. However, some individuals may be missed, others may be counted multiple times and the count may include misclassified species or other phenomena. In these circumstances, the count is a mixture of true detections c_i^{TP} and false detections c_i^{FP} , where $c_i^* = c_i^{TP} + c_i^{FP}$ and c_i^{TP} may be smaller than N_i . For example, following Royle (2004), let $c_i^{TP} \sim \text{Binomial}(N_i, p)$, where p is the probability of detecting an individual organism. Let c_i^{FP} arise from another process, for example, $c_i^{FP} \sim \text{Poisson}(\lambda^{FP})$.

This situation in which entities are counted and true and false detections are difficult to distinguish is common to aerial surveys (Conn et al., 2014). In this simple case, the marginal distribution of the observed data c_i^* is not $\text{Poisson}(\lambda^{TP})$ but $\text{Poisson}(\lambda^{TP}p + \lambda^{FP})$. The challenge is that not only are N_i , p , λ^{TP} and λ^{FP} unknown but also c_i^{TP} and c_i^{FP} . Hence, a unique solution to the values of p , λ^{TP} and λ^{FP} under the assumptions above requires additional information (Chambert et al., 2015). We focus on two designs (*Scenario 1* and *2* below) that use ground-validated observations to harmonize and correct estimates (sensu Chambert et al., 2015; Doser et al., 2021).

2.2 | Auxiliary data

In both scenarios, a subset of the H sampling units (J) are subject to further ground sampling, and the complementary subset K only is sampled using a less accurate method (e.g. aerial survey). We index individual units as $j=1, 2, \dots, J$ ($j \in H$), $k=1, 2, \dots, K$ ($k \in H$), where $H = J \cup K$ and $J \cap K = \emptyset$.

2.2.1 | Auxiliary data: Scenario 1

Under *Scenario 1*, we assume that target individuals within sampling units J can be perfectly enumerated by ground sampling. Thus,

N_j is an observed variable. By spatially resolving the correctly enumerated ground data (mapped individuals within a subset of units) with detections made using high resolution imagery (e.g. mapped crowns), the total count of image-based detections in a unit (c_j^*) can be subdivided into individuals correctly detected via aerial observation (c_j^{TP}), and observations introduced by misclassification or over-segmentation (double-counting) (c_j^{FP}).

The state process is shared across J and K , such that $N_i \sim \text{Poisson}(\lambda^{TP})$. To be more explicit about this information sharing, we can equivalently subdivide H into J and K (where K denotes sub-units without aerial sampling):

$$[N_j | \lambda^{TP}] \sim \text{Poisson}(\lambda^{TP}) \quad (1)$$

$$[N_k | \lambda^{TP}] \sim \text{Poisson}(\lambda^{TP}) \quad (2)$$

Within the ground-validated units J , the likelihood for the aerial data follows as:

$$[c_j^{TP} | N_j^{TP}, p] \sim \text{Binomial}(N_j, p) \quad (3)$$

$$[c_j^{FP} | \lambda_j^{FP}] \sim \text{Poisson}(\lambda_j^{FP}) \quad (4)$$

In the unvalidated sampling units K , c_k^{TP} and c_k^{FP} remain latent, and only their sum c_k^* is observed. Conditioning on a latent N_k^{TP} sampled during Markov chain Monte Carlo simulation, the likelihood for c_k^* is the convolution of the likelihoods for c_k^{TP} and c_k^{FP} that integrates (sums) over their potential combinations:

$$[c_k^* | N_k^{TP}, p, \lambda^{FP}] = \sum_{c_k^{FP}=0}^{c_k^*} [c_k^{TP} | N_k^{TP}, p] [c_k^{FP} = c_k^* - c_k^{TP} | \lambda^{FP}] \quad (5)$$

Although ground-sampled data is often treated as 'ground-truthed' data, previous assessments suggest that observers on the ground also fail to detect plants (Chen et al., 2013; Perret et al., 2023). In this case, operating under *Scenario 1* assumptions will result in biased estimates, although if the ground-validated counts exhibit less observation error than the aerial counts, the estimates may still be better than treating c_i^* as the truth. Hence, further extension may be useful.

2.2.2 | Auxiliary data: Scenario 2

In the second scenario, we assume the auxiliary ground data across sampling units J may be subject to false negative error but not false positive error—a person may miss an individual plant on the ground but will not falsely introduce one. Field practices like flagging counted plants to reduce risk of repeat counting may help ensure this assumption is met. If the probability of detecting a plant on the ground $p_{(g)}$ and in high-resolution imagery $p_{(a)}$ are the same, then a single ground

replicate is sufficient to correct for aerial observation errors. This seems unlikely in practice, and to estimate distinct $p_{(g)}$ and $p_{(a)}$ parameters, the ground-validated counting and mapping procedure must be replicated L times using multiple observers or visits. In this scenario, neither N_j nor aerial false positives c_j^{FP} are directly observed. Instead, there L independent efforts to map individuals, where each replicate is indexed as l ($l = 1, 2, \dots, L$) and the count for each replicate l arises as $c_{j,l}^g \sim \text{Binomial}(N_j, p_{(g)})$. We subsequently assume $L=2$ for simplicity. In turn, the aerial count c_j^* remains a convolution of true positives $c_j^{a,TP} \sim \text{Binomial}(N_j, p_{(a)})$ and false positives $c_j^{a,FP} \sim \text{Poisson}(\lambda^{FP})$. Assuming the mapped ground counts and aerial detections can be spatially resolved across replicates such that one can derive the total number of distinct individuals detected on the ground across all replicate surveys (n_j), the total count of individuals can be subdivided into 7 different classes (assuming $L=2$): (1) detected on the first ground survey only, (2) detected on the second ground survey only, (3) detected on both first and second ground surveys, (4) detected on first ground survey and via the air, (5) detected on the second ground survey and via the air, (6) detected on both ground surveys and aerially and (7) detected aerially only. The likelihoods for N_j and n_j are:

$$[N_j | \lambda^{TP}] \sim \text{Poisson}(\lambda^{TP}) \quad (6)$$

$$[n_j | N_j, p_{(g)}] \sim \text{Binomial}(N_j, p_{(g)}^*) \quad (7)$$

$$p_{(g)}^* = p_{(g)}^2 + 2(p_{(g)}(1 - p_{(g)})) \quad (8)$$

We re-arrange intersections of $c_{j,l}^g$ and c_j^* into a count-matrix $\mathbf{y}$ with J rows and 7 columns that correspond to the combinations of aerial and ground-validated detections when $L = 2$. Counts $y_{j,1:6}$ are definitively true positives that we model as

$$[y_{j,1:6} | n_j, \boldsymbol{\pi}] \sim \text{Multinomial}(n_j, \boldsymbol{\pi}) \quad (9)$$

where

$$\pi_1 = \frac{p_{(g)}(1 - p_{(g)})(1 - p_{(a)})}{p_{(g)}^*}$$

$$\pi_2 = \frac{(1 - p_{(g)})p_{(g)}(1 - p_{(a)})}{p_{(g)}^*}$$

$$\pi_3 = \frac{p_{(g)}p_{(g)}(1 - p_{(a)})}{p_{(g)}^*}$$

$$\pi_4 = \frac{p_{(g)}(1 - p_{(g)})p_{(a)}}{p_{(g)}^*}$$

$$\pi_5 = \frac{(1 - p_{(g)})p_{(g)}p_{(a)}}{p_{(g)}^*}$$

$$\pi_6 = \frac{p_{(g)}p_{(g)}p_{(a)}}{p_{(g)}^*}$$

Division by $p_{(g)}^*$ is needed for $\boldsymbol{\pi}$ to sum to 1, as y_j is conditional on the n_j individuals detected on the ground.

This leaves $y_{j,7}$, the count of individuals detected in aerial imagery only that includes true ($y_{j,7}^{TP}$) positive observations that eluded ground-validated detection $[y_{j,7}^{TP} | N_j, p_{(a)}, p_{(g)}] \sim \text{Binomial}(N_j, p_{(a)}(1 - p_{(g)}))^2$ and false positive observations $[y_{j,7}^{FP} | \lambda^{FP}] \sim \text{Poisson}(\lambda^{FP})$. We sum over the latent composition sensu Equation 5 as

$$[y_{j,7} | N_j, p_{(g)}, p_{(a)}, \lambda_j^{FP}] = \sum_{y_{j,7}^{TP}=0}^{y_{j,7}} [y_{j,7}^{TP} | N_j, p_{(g)}, p_{(a)}] [(y_{j,7}^{FP} = y_{j,7} - y_{j,7}^{TP}) | \lambda^{FP}] \quad (10)$$

Finally, the likelihood in the sampling units K without ground-sampling follows Equations 2 and 5:

$$[N_k | \lambda^{TP}] \sim \text{Poisson}(\lambda^{TP}) \quad (11)$$

$$[c_k^* | N_k, p_{(a)}, \lambda_k^{FP}] = \sum_{c_k^{TP}=0}^{c_k^*} [c_k^{TP} | N_k, p_{(a)}] [(c_k^{FP} = c_k^* - c_k^{TP}) | \lambda^{FP}] \quad (12)$$

Equations 6–9 describe a conditional multinomial N-mixture model (Kéry & Royle, 2015, p. 334), where the ground-validated surveys arise from a ‘independent double-observer’ protocol. Other approaches like a ‘removal’ protocol—where the first survey flags and maps all plants detected in the unit, and the second flags and maps only remaining unflagged plants in the unit—might be more efficient in the field and only require a few modifications to the equations above. Both reframe the issue as a capture-recapture problem (see Section 4).

2.3 | Simulation study

We performed a simulation study to determine that models with auxiliary data under *Scenario 1* and *Scenario 2* could recover generating parameters and latent abundance states. To briefly explore sensitivity to the amount of auxiliary data, we ran simulations with varying proportions of ground-sampled data, from 3% to 13% coverage (50–500 cells/units in a 1600 unit lattice). Further details and results are provided in Appendix S1.

2.4 | Application to sagebrush steppe

We apply the model to data collected from 10 non-contiguous landscapes in southwestern Idaho, USA. Each site included an aerial UAS survey and ground-mapped individual plants obtained in a supporting field campaign (no permissions were needed for field work: see Appendix S2 for additional details on data collection, Figure S10 for a visual overview of our approach, and Marie et al., 2023 for details on image processing and georeferencing). We discretized high-resolution imagery into 5 m grid cells ($H=6605$). Approximately 11% were randomly validated with exhaustive ground sampling of individual plants ($J=729$) using a survey-grade RTK GPS unit (Topcon HiPer V, Topcon

Positioning Systems Inc., Livermore, CA, USA) with $\sim 0.02\text{m}$ accuracy (Rayburn et al., 2011). This ground-sampling process took about 20h per replicate landscape on average. The remaining cells ($K=5876$) contained only aerial counts. We processed the aerial data using an object-based image analysis (OBIA) that resulted in a map of sagebrush crowns. Our OBIA used an algorithm for tree segmentation applied to the canopy height model (Silva et al., 2016) implemented in the lidR R package (Roussel et al., 2020), and we refer readers to Appendix S2 for additional details. We next overlaid the crown maps from OBIA with field-mapped sagebrush points (Figure 1b), which allowed us to spatially match aerial and ground detections and aggregate the aerial detections across J cells (Figure 1c) as the counts of true detections (c_j^{TP}), missed detections (c_j^{FN}) and misclassified or double-counted plants (c_j^{FP}). The unvalidated cells contained only the total count of plants detected from air ($c_k^* = c_k^{TP} + c_k^{FP}$).

We note two important caveats about our application. As a pre-processing step (Young et al., 2022), we used random forest classification to 'thin' likely non-sagebrush crowns initially identified via the OBIA using additional spectral data. This step is not necessary to implement the statistical model, and exploration suggests using the crowns initially identified strictly during OBIA process as input data without thinning results in very similar

inference about plant abundance. However, inference about error parameters is sensitive to this step, which likely reduces false positive errors within the initial pool of observations and introduces additional false negative error, and we emphasize that our estimates of false positive and false negative error parameters reflect OBIA data *after* this pre-processing. We later discuss extensions to more directly leverage machine learning scores within the statistical model. A second caveat is that the pre-processing classification model was trained using the ground-validation data used within the statistical model. Thus, the application uses certain data twice, which we do not recommend but which was unavoidable here because there was not external data to use for classifier training. However, we believe the steps used here create a more valid comparison between our approach and current approaches where users might use ground-sampling to inform a classification model and treat the count of predicted target individuals as the truth (see further discussion in Appendix S2).

We expected the abundance of juvenile and mature plants to exhibit different abundance associations with topography and wildfire legacy, and early analysis focusing on field-labelled plants only suggested the probability of detection was relatively stable for plants with a photogrammetrically estimated height $>0.25\text{m}$. Therefore,

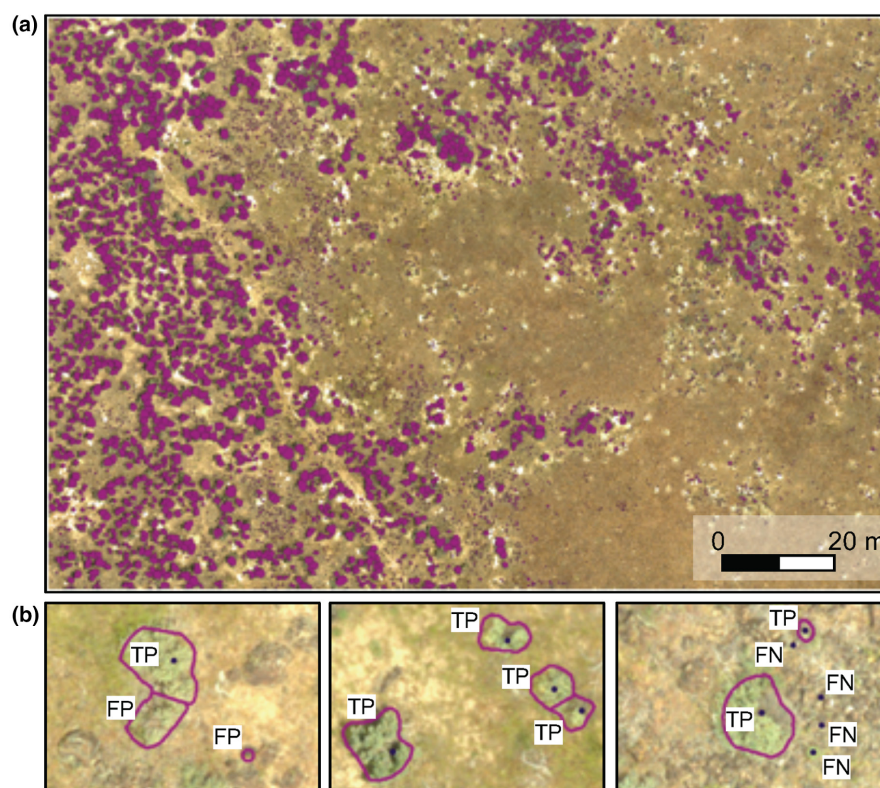


FIGURE 1 Sampling design (a, b) used in the statistical model for imperfect detection. Purple segments represent plant crowns. Starting from the top, (a) an unmanned aircraft systems (UAS) image capturing a landscape partially burnt in a wildfire collected as part of the study, with crowns identified as big sagebrush (*Artemisia tridentata*) shaded in purple. Magnified sections of the UAS orthomosaic (b), showing the automatically delineated shrub crowns (outlined in purple) overlaid on top of the field GPS points that mark ground-mapped sagebrush (black points). The labels indicate whether a field identified plant was detected in high-resolution imagery (true positive, TP), a field identified plant was not detected in high-resolution imagery (false negative, FN), or if a field identified plant was either double-counted or if a non-target entity was misclassified as big sagebrush in the high-resolution imagery (false positive, FP).

we applied the model separately to two distinct classes of sagebrush, 'juvenile' (<0.25 m height) and 'adult' (≥0.25 m height) individuals. For each plant mapped in the field, we used a tape measure to record whether the plant was above or below the 0.25 m threshold. We applied the same 0.25 m threshold to detected plants in canopy height models derived from aerial imagery, assuming the plant height corresponded to the highest canopy pixel within a delineated crown.

We hypothesized three covariates would influence the probability of detecting individuals across the replicate landscapes: (1) maximum wind recorded during the UAS flight (Wind_max), (2) average vegetation height (Height_avg) and (3) the proportion of area segmented into objects (Area_segment). The last metric, the ratio between the total area of individual objects to the area of the site, reflects the OBIA parameter choices determining the process of canopy delineation based on canopy height. Smaller polygons (i.e. conservative segmentation parameters) are less likely to overlap a ground-validated GPS point compared to relatively large polygons that would result from less conservative segmentation parameters. Our detection model with these covariates is as follows, with subscript s indicating replicate landscapes:

$$[c_j^{TP} | N_j^{TP}, p_{s[j]}] \sim \text{Binomial}(N_j^{TP}, p_{s[j]}) \quad (13)$$

$$\text{logit}(p_s) = \psi_0 + \psi_1 \text{Wind_max}_s + \psi_2 \text{Height_avg}_s + \psi_3 \text{Area_segment}_s \quad (14)$$

We hypothesized that surface topography would primarily influence variation in false positive errors. Topographic properties of the terrain, for example terrain roughness or sun exposure, may translate to variable performance of structure-from-motion algorithm on UAS data. We obtained topographic predictors based on digital terrain models (DTMs), including topographic roughness index (TRI), topographic position index (TPI) and heat load index (HLI) derived from the calculated aspect of each cell (Hijmans et al., 2022). We scaled the topographic predictors by centring and dividing them by two standard deviations (SD). We modelled the FP errors as

$$\log(\lambda_i^{FP}) = \beta_0^{FP} + \beta_1^{FP} \text{TRI}_i + \beta_2^{FP} \text{TPI}_i + \beta_3^{FP} \text{HLI}_i \quad (15)$$

Third, we were interested in quantifying the effect of surface topography and of the distance from wildfire edge on the true abundance of sagebrush (Equation 14). We hypothesized that the topographic predictors in Equation 13 also influenced the cell-level abundance of sagebrush, by creating microsites that alter establishment (Condon & Weisberg, 2016; Germino et al., 2018). We also hypothesized that sagebrush abundance declined with distance to unburned patches, due to seed limitation (Appelstein et al., 2022). To quantify this effect, we used distance from wildfire edge as a predictor (Equation 14, parameters θ_1 and θ_2), assuming an exponential decline in sagebrush abundance further from wildfire edge. We assigned a zero value for the distance-to-edge variable to all unburnt cells. To account for residual spatial autocorrelation in sagebrush abundance, we included an intrinsic conditional autoregressive

effect ϕ_i at the cell level (Morris et al., 2019). Finally, we assumed $N_i^{TP} \sim \text{NB}(\lambda_i^{TP}, \kappa)$, where κ is a dispersion parameter and

$$\log(\lambda_i^{TP}) = \mathbf{X}_i \boldsymbol{\beta}^{TP} + \theta_1 \exp(-\text{dist}_i^2 * \theta_2) + \phi_i; i \geq \{j, k\} \quad (16)$$

We use the same model specification for both small and large plants.

We processed the data and model outputs in R version 4.2.2 (R Core Team, 2021). We fit models using Markov-Chain Monte Carlo simulation in 'NIMBLE' package in R version 4.1.1 (de Valpine et al., 2017) employing regularizing priors: $\boldsymbol{\beta}, \boldsymbol{\psi} \sim N(\mathbf{0}, \mathbf{0.25I})$ and ecologically informed priors on the distance-dependent wildfire edge effect, $\theta_1 \sim \text{Half-normal}(0, 1)$ and $\theta_2 \sim \text{Gamma}(3, 0.58)$ based on previous estimates of sagebrush seed dispersal (Appelstein et al., 2022). Each simulation consisted of 8 chains of 140,000 iterations, where 20,000 were discarded as burn-in, and the remainder were thinned by 120. We used standard diagnostics and visually inspected trace-plots to assess mixing and convergence (Brooks & Gelman, 1998). We used 'tidyverse', 'terra', 'sf' and 'ggspatial' packages for data processing and visualization (Dunnington, 2021; Hijmans et al., 2022; Pebesma, 2018; Wickham et al., 2019) and 'MCMCvis' for model post-processing (Youngflesh, 2018).

3 | RESULTS

Simulation confirmed the proposed models were identifiable and estimator performance generally improved with increased ground sampling. Further details are in the Supporting Information (Appendix S1 and Figures S1–S6 therein).

3.1 | Observation errors: False negatives

Detection parameters varied among sites and between the two size classes (Figures 2 and 3). For 'adults', the average probability of detection across all sites was 0.63 (95% CI: 0.46, 0.84), with site-level detection probabilities (p_s) ranging from 0.47 (95% CI: 0.44, 0.50) to 0.82 (95% CI: 0.75, 0.88). 'Juvenile' detection probabilities were much lower, with an average of 0.13 (95% CI: 0.02, 0.32) and ranged across sites from 0.03 (95% CI: 0.02, 0.03) to 0.30 (95% CI: 0.26, 0.35). Juvenile and adult detection probabilities increased with greater overall plant canopy height, and adult detection probability was positively associated with segmentation area and negatively associated with wind speed (Figure 2).

3.2 | Observation errors: False positives

Posterior simulations predicted an average of 0.04 (95% CI: 0.004, 0.37) and 0.06 (95% CI: 0.02, 0.45) false positive adult and juvenile false positive detections per m⁻², respectively. Overall, topographic covariates each affected the rates of FP detections (Figure 3), although we interpret results cautiously because post-hoc assessment suggests mediocre fit for this process (Figure S7).

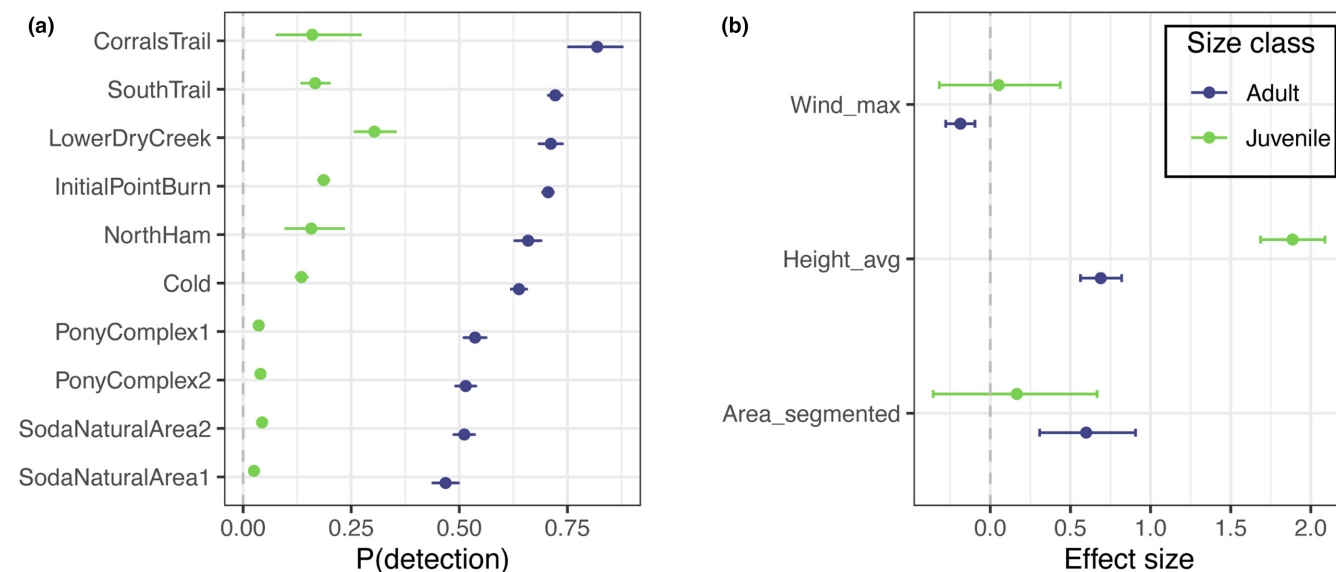


FIGURE 2 The estimated detection probabilities of sagebrush from high-resolution unmanned aircraft systems imagery across 10 sites in SW Idaho, USA (a) and factors explaining detection variability across sites (b). Points indicate mean parameter estimates, and the error bars correspond to 95% credibility intervals. Colour of points indicates size class: Blue points are adults and green points are juveniles. The effect size (right x-axis) is shown on the logit scale relative to the mean detection by size class, indicated by zero.

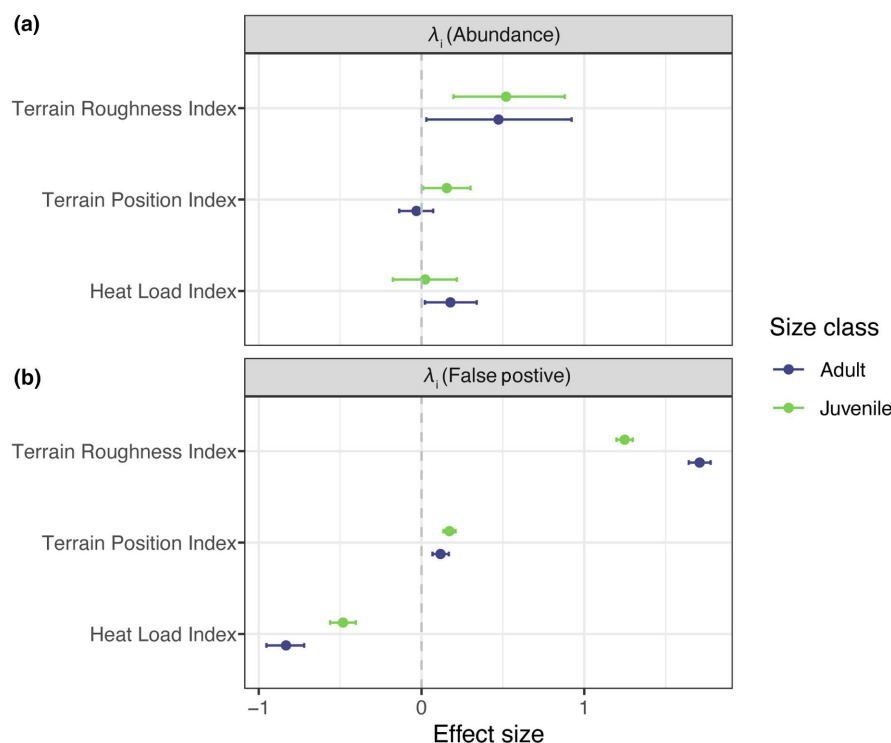


FIGURE 3 The effect of surface topography on the abundance of sagebrush (a) and false positive detections (b) from unmanned aircraft systems imagery across 10 sites in SW Idaho, USA. The variables (y-axis) summarize the effect of TRI (topographic roughness index), TPI (topographic position index), and HLI (heat load index) derived from Digital Terrain Models (DTM) at 5 m resolution.

3.3 | Abundance effects

Mean predicted abundance across cells was 0.32 plants m^{-2} (95% CI: <0.01, 1.76) for adult and 1.67 (95% CI: <0.01, 7.96). Terrain roughness was positively associated with both juvenile and adult abundance and the rate of false positive errors for each size class. Topographic position index was positively associated with true

abundance for the juvenile size class but had an uncertain effect on true abundance of adult plants. Although heat load index was negatively associated with FP detections for both juvenile and adult plants, it was positively associated with abundance of adult plants.

The effect of distance to wildfire edge manifested over a larger range with a shallower decline for the juvenile size class but had higher uncertainty (Figure 4). Adult plants were effectively absent

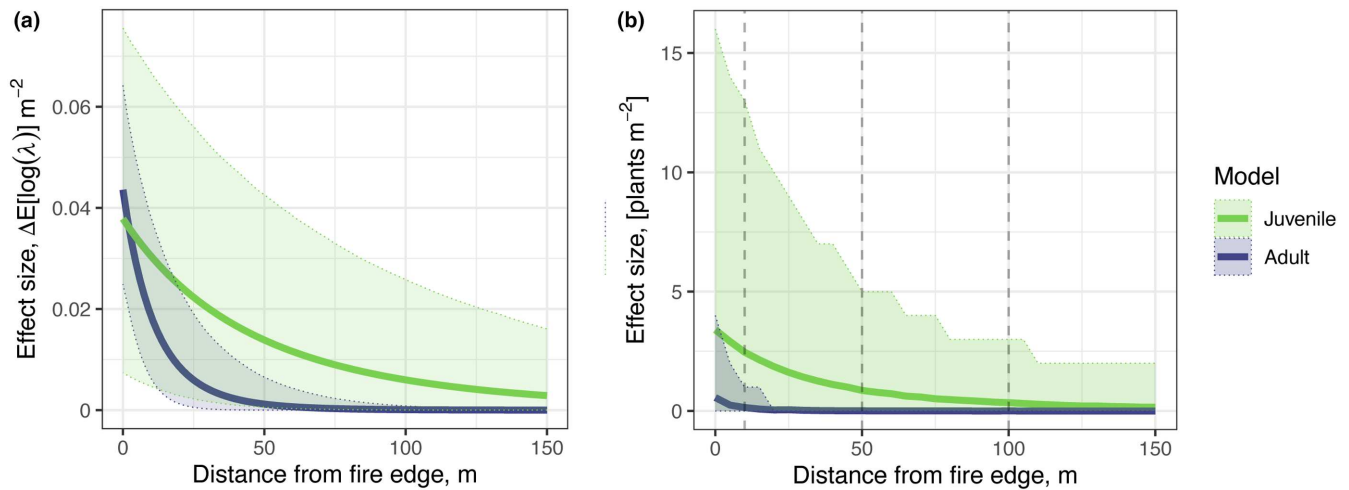


FIGURE 4 The spatial effect of a wildfire edge on the abundance of sagebrush in partially burnt areas across 10 sites in SW Idaho, USA. The y-axis indicates the marginal effect of unburnt vegetation relative to the corresponding means for the two size classes. The left panel (a) shows the predicted effect of unburnt vegetation on the average expected count. The right panel (b) shows the same effect using posterior prediction to simulate counts. Thick lines depict means and shadowed regions indicate the central 95% CI of the predicted effect.

100m within the wildfire burn scar, while the expected density for juvenile plants was 0.35 m^{-2} at the same distance.

3.4 | Goodness of fit

Posterior predictions suggested agreement between observed and predicted false positive errors (mean absolute error estimates of 0.07 and 0.14 false positive detections m^{-2} for the adult and juvenile classes, respectively; Figure S7), although overall fit and the coverage of predictive intervals suggested room for improvement. In contrast, posterior predictions of plant abundance within ground-validated cells and well as the number of detected features in cells without ground-validation exhibited broader uncertainty, but the uncertainty appeared better calibrated and predictive fit appeared stronger (Figures S8 and S9). While we suggest that estimates of false positive error parameters should be interpreted cautiously (the effect uncertainty presented is likely overconfident), we believe inference about drivers of plant abundance and abundance predictions is more robust.

3.5 | Abundance estimates

Overall, juvenile plants were more abundant than adults, with estimates of $>150,000$ juvenile and $>30,000$ adult plants in burnt areas averaged across our 10 sites (Table S1, Figure 5). The variation in average density between sites was also substantially greater for juveniles than for adults (Table S2). Compared to model-based estimates of abundance, raw OBIA counts were generally negatively biased (Table S4). Across all sites and size-classes, $<50\%$ of the OBIA counts were within the 95% CI of the predicted site-level abundance, with counts at 4/10 sites for adult and 7/10 for juvenile size class falling

beneath the lower 95% CI. As previous posterior predictive results suggest abundance estimates fit reasonably well and uncertainty intervals were reasonably well-calibrated, it appears that raw counts of OBIA classifications (even with pre-processing) provided poor estimates of sagebrush abundance.

4 | DISCUSSION

Our purpose here was to: (1) develop an approach to estimate plant abundance in high-resolution imagery accounting for both false negative and false positive observation errors, (2) gain insights into potential drivers of spatial variation in sagebrush abundance and (3) initially assess the magnitude of these observation errors within similar applications and potential factors moderating their magnitude. Scalable and spatially contiguous estimates of plant species abundance derived from high-resolution imagery have direct application for landscape-level restoration and conservation efforts (Arkle et al., 2022) and provide a starting point for spatial models of plant population dynamics (Kellner & Hubbell, 2018). However, harnessing imagery to achieve these objectives will require grappling with observation errors. In our application, counts from an object-based image analysis that were pre-processed to reduce false positive errors failed to detect more than 1/3 of plants observed on the ground, and roughly 19% of plants detected via aerial imagery in a typical cell were still false positives. Our approach provides a means to account for biases associated with these errors and account for detection uncertainty when estimating abundance and its environmental associations. Although our approach requires auxiliary ground data, this is commonly collected in UAS or other remote sensing campaigns. Indeed, the general approach is transferable to any type of remotely sensed data where distinguishing individual plant crowns is possible. This includes aerial lidar (Barber et al., 2022) and high-resolution

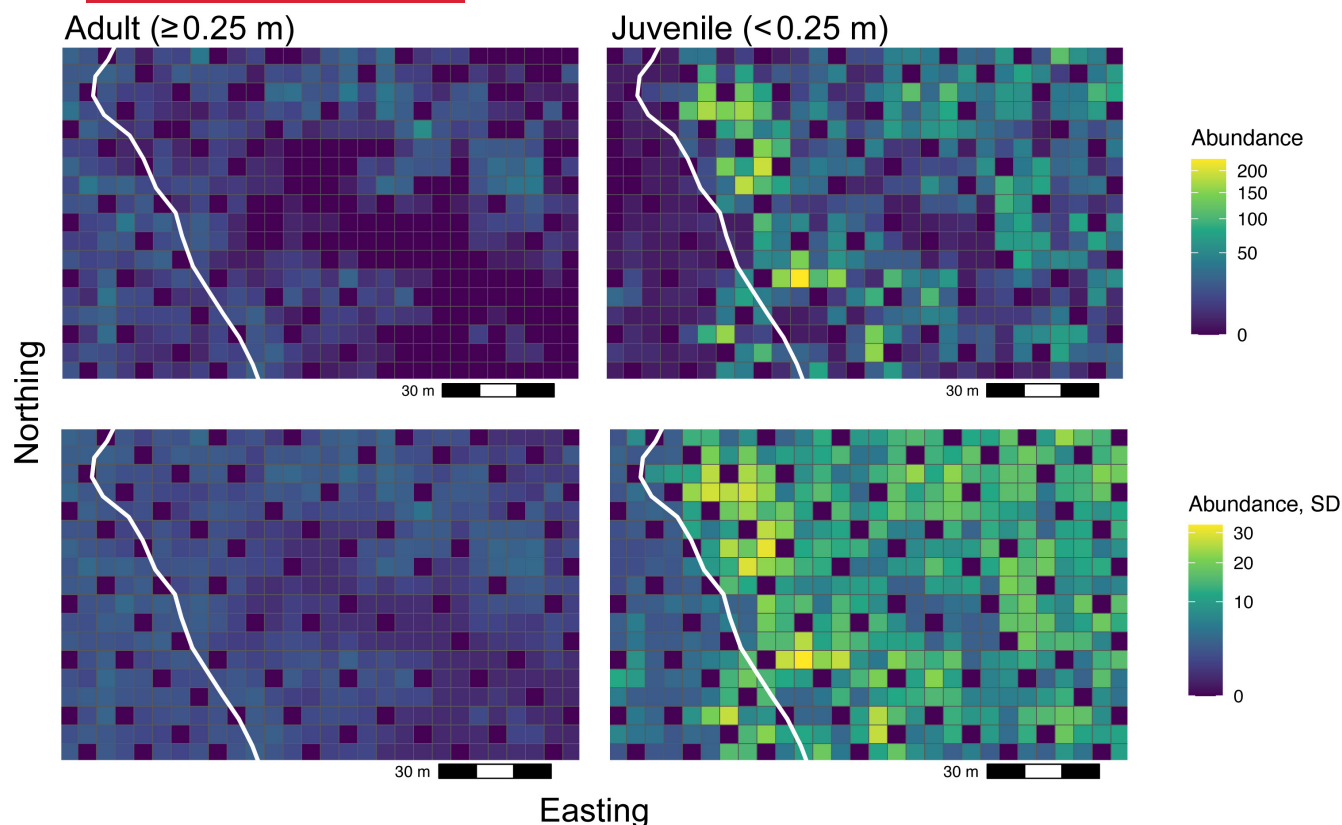


FIGURE 5 Predictive maps of sagebrush abundance within a single landscape (see aerial image in [Figure 1](#)). Top row shows model predictions of true abundance. The bottom row shows the uncertainty (the standard deviation of the posterior distribution). Cells with $SD=0$ represent cells where field-validated counts occurred (under the assumption that all plants were detected). The north–south white line indicates the boundary of a wildfire that occurred 26 years before the data collection east of the line, with adults more abundant in the unburnt section (west) and recruits more abundant near the edge of the burn (east).

satellite data (Liu et al., 2023), although application to different data-sources will likely require identifying different covariates for observation errors (e.g. changing the terms in [Equation 13–16](#)).

4.1 | Drivers of variation in abundance and observation errors

Our approach reduces the frailty associated with high-resolution imagery efforts while retaining its advantages. Sparse but ecologically important areas of recruitment in post-fire landscapes are difficult to reliably delineate without the data volume and spatial coverage provided by remote sensing (Ziegenhagen & Miller, 2009). We found pronounced spatial variability in the abundance of smaller (likely younger) plants: >90% of ground-sampled plots contained zero recruits but some had as many as $17.4 \text{ recruits m}^{-2}$. Proximity to unburnt locations where adult seed sources remain prevalent was a useful predictor of sagebrush recruit abundance. Associations between undisturbed vegetation and recruitment are both crucial for forecasts of natural regeneration (Barber et al., 2022) and difficult to determine with patchy in-situ sampling (Applestein et al., 2022), and contiguous ('wall to wall') coverage provided by aerial imagery data can help resolve

these issues. Spatial patterns of juvenile and adult plant abundance here are consistent with an expanding wave of natural regeneration, with juvenile plants more abundant deeper within the burn footprint.

Understanding abundance drivers is often of equal interest to estimating population size. Here, we used a structure-from-motion algorithm to estimate fine-scale differences in relative topography, topographic roughness, topographic position and heat load. Of these, the topographic roughness index had the strongest effect on both adult and juvenile abundance. Microtopographic variation captured by this index may represent differences in soil moisture, crucial for shrub demography in semi-arid ecosystems (Germino et al., 2018). In contrast, topographic position index, which represents whether a point is higher or lower than the surrounding landscape, had a positive effect on juvenile abundance, but an uncertain effect on adult abundance. Heat load index showed the reverse pattern: a positive effect for adults, but not juveniles. Repeat UAS flights to quantify individual plant growth and survival (e.g. Olsoy et al., 2024) may advance our understanding of stage-structured differences in abundance.

Several factors appeared to influence variation in observation errors. False-negative errors were more pronounced when vegetation was shorter on average, imagery was collected under windy

flight conditions and when more conservative segmentation parameters were employed within the OBIA. False-positive errors were impacted by environmental variation, with topographic roughness having the strongest (and positive) estimated effect. As our data processing sought to limit misclassification, it is not surprising that plant abundance was also greater in areas with rougher topography as it suggests most false positive errors arose from double counting. Continued application will improve understanding of the factors that affect detection error associated with plants in UAS imagery and as it has for other remote sensing applications (e.g. Hofmeester et al., 2019). While our application assumed that false-negative error was predominantly related to site-level factors and false-positive error was related to finer-grained environmental variation, we expect each process is impacted by sampling and environmental considerations. For example, it seems likely—given relatively poor predictive fit to false positive counts—that factors we assumed impacted false negative error only, such as segmentation parameters, play an important role in the false positive process, too. A practical benefit is that our model makes optimizing the segmentation parameters somewhat less important: rather than needing to tune these to make the immediate OBIA output as accurate as possible, one can use the statistical model to account for the (inevitable) errors that arise during segmentation.

4.2 | Contextualizing and improving the model

Our approach modifies previous models for binary data with false negative and positive error and a binary state variable at a set of replicate units (Chambert et al., 2015; Miller et al., 2011) or binary data and varying state variables (Clare et al., 2021) to address similar issues with count data and a count state variable. We assume ground-validation at the level of spatial units (i.e. the unit is exhaustively searched for individuals). Validation could instead focus on a sub-set of putative individuals within units, in which case N_i is never observed, but can still be inferred by considering further processes (e.g. equation 7 in Doser et al., 2021). Consistent with these aforementioned works, we lump misclassification and double-counting errors into an omnibus false positive error. It may be beneficial to differentiate these (Clement et al., 2022; Conn et al., 2014; Spiers et al., 2022; Wright et al., 2020), although this may require ground sampling to partially enumerate and map non-target objects (rocks, other plant species) and may increase field-time. Such data were not available for our application, although we believe more false positives in our application arose from double counting errors than misclassification due to data pre-processing decisions (see Appendix S2).

The two models we present differ in that *Scenario 1* assumes the ground data lacks any observation error and *Scenario 2* assumes the ground data may exhibit false-negative error. Although our application follows *Scenario 1* due to the available data, ground surveys may also be imperfect (Chen et al., 2013; Perret et al., 2023). Thus, while we expect that expect ground-based counts will typically be more accurate than counts from high-resolution imagery—and

that *Scenario 1* will improve estimates relative to using raw counts from high-resolution imagery—we recommend that future applications employ *Scenario 2*. Both models assume individuals detected on the ground and via the air can be reconciled. The assumption is more important for *Scenario 2*, but we believe it is reasonably met when locations of individuals detected on the ground are mapped with high accuracy (e.g. the survey-grade GPS used here has <3 cm horizontal error on average), and the high-resolution imagery is georeferenced with high precision. If individual reconciliation proves challenging, there are number of other ways to collect auxiliary data that might either relax the need to reconcile individuals (independent replicated counts following Royle, 2004) or leverage similar spatial location and other individual characteristics (size, spectral characteristics) to probabilistically reconcile individuals (Augustine et al., 2020).

Because *Scenario 2* invokes a capture-recapture design for ground data, there are further opportunities for improvement beyond accounting for false negative error in the field. Our results indicate that aerial false negative error is associated with individual size. Rather than split detected objects into size groups, capture-recapture approaches employing data augmentation or similar techniques allow variation in detection to be associated with individual-level predictors (Royle, 2009). Similarly, double-counting due to segmentation error also seems more likely for larger individuals, and the rate with which individual plants are 'correctly' detected (but also double counted) could be posed as increasing with size (extending Clement et al., 2022). Many users may wish to use machine learning scores and additional spectral information to help distinguish between target and non-target objects as we did here, even if misclassification and double-counting are not distinguished. Rather than use pre-processing to alter the input to a subsequent statistical model, one could 'couple' the two by treating machine learning scores as individual level covariates that influence the probability that a detected object is a target object or non-target object (Kéry & Royle, 2020; Rhinehart et al., 2022). This may further reduce bias and better account for uncertainty associated with imperfect machine learning scores used for pre-processing.

Although our application focuses on two closed populations with different stage/size attributes assumed to be independent, extensions focusing on open populations or with explicit dependencies between stage-structured abundance in space or time are relatively straightforward (Clare et al., 2016; Dail & Madsen, 2011; Hostetler & Chandler, 2015; Zipkin et al., 2014). Such extensions might improve insights into the processes structuring population patterns, population change and limitations to natural regeneration (Paniw et al., 2023).

4.3 | Summary

Spatially explicit and contiguous estimates of plant abundance are critical for landscape-level understanding of population variability and practical management (Gurevitch et al., 2016; Shriver

et al., 2019; Young et al., 2022). However, our results demonstrate that studies leveraging remotely sensed data and segmentation algorithms to achieve this goal must contend with substantial observation error that compromises population inference even if using post-hoc classification algorithms to limit errors. Issues arising from observation errors have been long recognized in wildlife ecology and more recently both in plant ecology (Chen et al., 2013) and remote sensing applications focused on land surface attributes (Veran et al., 2012). High-resolution imagery has opened new frontiers into studying plant distributions and demography, and we urge investigators pursuing these questions to be conscientious of such errors. We hope the framework here provides a useful starting point.

AUTHOR CONTRIBUTIONS

Andrii Zaiats, Trevor T. Caughlin, Jennyffer Cruz, David S. Pilliod, Megan E. Catau, Richard Rachman, Maisha Maliha, Donna Delparte and John D. J. Clare conceived of the ideas and designed methodology. Andrii Zaiats collected the data; Andrii Zaiats, Richard Rachman, & John D. J. Clare analysed the data; Andrii Zaiats, Trevor T. Caughlin and John D. J. Clare led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

We thank Anna Roser, Peter Olsoy and Valorie Marie for their help in UAS data collection, processing and management. We also thank Ryan Wickersham, Cana Foncannon, Amy Johnson, Ian Lambrecht, Cayden Whipkey, Lena Griffith and Shyam Batchu for their help with fieldwork. This publication was made possible by the NASA FINESST Program #80NSSC21K1638 and the National Science Foundation under award numbers OIA-1757324 and BIO-2207158. Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the US Government.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14421>.

DATA AVAILABILITY STATEMENT

Data and code associated with the study are available at <https://doi.org/10.5281/zenodo.13528399> (Clare & Zaiats, 2024).

ORCID

Andrii Zaiats  <https://orcid.org/0000-0001-8978-4152>
T. Trevor Caughlin  <https://orcid.org/0000-0001-6752-2055>
Jennyffer Cruz  <https://orcid.org/0000-0002-5321-8017>
David S. Pilliod  <https://orcid.org/0000-0003-4207-3518>

Megan E. Cattau  <https://orcid.org/0000-0003-2164-3809>
Richard Rachman  <https://orcid.org/0009-0009-8463-0483>
Maisha Maliha  <https://orcid.org/0009-0004-9635-1389>
Donna Delparte  <https://orcid.org/0000-0002-9107-5117>
John D. J. Clare  <https://orcid.org/0000-0003-2583-8742>

REFERENCES

- Applestein, C., Caughlin, T. T., & Germino, M. J. (2022). Post-fire seed dispersal of a wind-dispersed shrub declined with distance to seed source, yet had high levels of unexplained variation. *AoB Plants*, 14(6), plac045. <https://doi.org/10.1093/aobpla/plac045>
- Arkle, R. S., Pilliod, D. S., Germino, M. J., Jeffries, M. I., & Welty, J. L. (2022). Reestablishing a foundational species: Limitations on post-wildfire sagebrush seedling establishment. *Ecosphere*, 13(8), e4195. <https://doi.org/10.1002/ecs2.4195>
- Augustine, B. C., Royle, J. A., Linden, D. W., & Fuller, A. K. (2020). Spatial proximity moderates genotype uncertainty in genetic tagging studies. *Proceedings of the National Academy of Sciences of the United States of America*, 117(30), 17903–17912.
- Barber, C., Graves, S. J., Hall, J. S., Zuidema, P. A., Brandt, J., Bohlman, S. A., Asner, G. P., Bailón, M., & Caughlin, T. T. (2022). Species-level tree crown maps improve predictions of tree recruit abundance in a tropical landscape. *Ecological Applications*, 32(5), e2585. <https://doi.org/10.1002/eap.2585>
- Brandt, M., Tucker, C. J., Kariryaa, A., Rasmussen, K., Abel, C., Small, J., Chave, J., Rasmussen, L. V., Hiernaux, P., Diouf, A. A., Kergoat, L., Mertz, O., Igel, C., Gieseke, F., Schöning, J., Li, S., Melocik, K., Meyer, J., Sinno, S., ... Fensholt, R. (2020). An unexpectedly large count of trees in the West African Sahara and Sahel. *Nature*, 587(7832), Article 7832. <https://doi.org/10.1038/s41586-020-2824-5>
- Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434–455. <https://doi.org/10.1080/10618600.1998.10474787>
- Caughlin, T. T., Damschen, E. I., Haddad, N. M., Levey, D. J., Warneke, C., & Brudvig, L. A. (2019). Landscape heterogeneity is key to forecasting outcomes of plant reintroduction. *Ecological Applications*, 29(2), e01850. <https://doi.org/10.1002/eap.1850>
- Chambert, T., Miller, D. A. W., & Nichols, J. D. (2015). Modeling false positive detections in species occurrence data under different study designs. *Ecology*, 96(2), 332–339. <https://doi.org/10.1890/14-1507.1>
- 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>
- Clare, J., & Zaiats, A. (2024). uas_spp_adundance2: Zenodo (v0.1.1). *Zenodo* <https://doi.org/10.5281/zenodo.13528399>
- Clare, J. D. J., Linden, D. W., Anderson, E. M., & MacFarland, D. M. (2016). Do the antipredator strategies of shared prey mediate intraguild predation and mesocarnivore suppression? *Ecology and Evolution*, 6(12), 3884–3897.
- Clare, J. D. J., Townsend, P. A., & Zuckerberg, B. (2021). Generalized model-based solutions to false-positive error from species detection/nondetection data. *Ecology*, 102(2), e03241.
- Clement, M. J., Royle, J. A., & Mixan, R. J. (2022). Estimating occupancy from autonomous recording unit data in the presence of misclassifications and detection heterogeneity. *Methods in Ecology and Evolution*, 13(8), 1719–1729. <https://doi.org/10.1111/2041-210X.13895>
- Condon, L. A., & Weisberg, P. J. (2016). Topographic context of the burn edge influences postfire recruitment of arid land shrubs. *Rangeland Ecology & Management*, 69(2), 129–133. <https://doi.org/10.1016/j.rama.2015.09.002>

- Conn, P. B., Ver Hoef, J. M., McClintock, B. T., Moreland, E. E., London, J. M., Cameron, M. F., Dahle, S. P., & Boveng, P. L. (2014). Estimating multispecies abundance using automated detection systems: Ice-associated seals in the Bering Sea. *Methods in Ecology and Evolution*, 5(12), 1280–1293. <https://doi.org/10.1111/2041-210X.12127>
- Dail, D., & Madsen, L. (2011). Models for estimating abundance from repeated counts of an open metapopulation. *Biometrics*, 67(2), 577–587. <https://doi.org/10.1111/j.1541-0420.2010.01465.x>
- Davis, K. L., Silverman, E. D., Sussman, A. L., Wilson, R. R., & Zipkin, E. F. (2022). Errors in aerial survey count data: Identifying pitfalls and solutions. *Ecology and Evolution*, 12(3), e8733. <https://doi.org/10.1002/ece3.8733>
- de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Lang, D. T., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, 26(2), 403–413. <https://doi.org/10.1080/10618600.2016.1172487>
- Doser, J. W., Finley, A. O., Weed, A. S., & Zipkin, E. F. (2021). Integrating automated acoustic vocalization data and point count surveys for estimation of bird abundance. *Methods in Ecology and Evolution*, 12(6), 1040–1049.
- Dunnington, D. (2021). ggspatial: Spatial Data Framework for ggplot2. R package version 1.1. 5 (2021).
- Germino, M. J., Barnard, D. M., Davidson, B. E., Arkle, R. S., Pilliod, D. S., Fisk, M. R., & Applestein, C. (2018). Thresholds and hotspots for shrub restoration following a heterogeneous megafire. *Landscape Ecology*, 33(7), 1177–1194.
- Gurevitch, J., Fox, G. A., Fowler, N. L., & Graham, C. H. (2016). Landscape demography: Population change and its drivers across spatial scales. *The Quarterly Review of Biology*, 91(4), 459–485. <https://doi.org/10.1086/689560>
- Hijmans, R. J., Bivand, R., Forner, K., Ooms, J., Pebesma, E., & Sumner, M. D. (2022). Package 'terra'. Maintainer, Vienna, Austria.
- Hofmeester, T. R., Cromsigt, J. P., Odden, J., Andrén, H., Kindberg, J., & Linnell, J. D. (2019). Framing pictures: A conceptual framework to identify and correct for biases in detection probability of camera traps enabling multi-species comparison. *Ecology and Evolution*, 9(4), 2320–2336. <https://doi.org/10.1002/ece3.4878>
- Hostetler, J. A., & Chandler, R. B. (2015). Improved state-space models for inference about spatial and temporal variation in abundance from count data. *Ecology*, 96(6), 1713–1723. <https://doi.org/10.1890/14-1487.1>
- Ibáñez, I., Jr, S., John, A., Allen, J. M., Treanor, S. A., & Wilson, A. (2009). Identifying hotspots for plant invasions and forecasting focal points of further spread. *Journal of Applied Ecology*, 46(6), 1219–1228. <https://doi.org/10.1111/j.1365-2664.2009.01736.x>
- Kellner, J. R., & Hubbell, S. P. (2018). Density-dependent adult recruitment in a low-density tropical tree. *Proceedings of the National Academy of Sciences of the United States of America*, 115(44), 11268–11273. <https://doi.org/10.1073/pnas.1800353115>
- Kéry, M., & Royle, J. A. (2015). *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS: Volume 1*. New York: Academic Press.
- Kéry, M., & Royle, J. A. (2020). *Applied hierarchical modeling in ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 2: Dynamic and advanced models*. Academic Press.
- Liu, S., Brandt, M., Nord-Larsen, T., Chave, J., Reiner, F., Lang, N., Tong, X., Ciais, P., Igel, C., Pascual, A., Guerra-Hernandez, J., Li, S., Mugabowindekwe, M., Saatchi, S., Yue, Y., Chen, Z., & Fensholt, R. (2023). The overlooked contribution of trees outside forests to tree cover and woody biomass across Europe. *Science Advances*, 9(37), eadh4097. <https://doi.org/10.1126/sciadv.adh4097>
- Louthan, A., & Doak, D. (2018). Measurement error of state variables creates substantial bias in results of demographic population models. *Ecology*, 99(10), 2308–2317. <https://doi.org/10.1002/ecy.2455>
- Marie, V., Zaiats, A., Wickersham, R., & Caughlin, T. T. (2023). *Open drone map: Structure-from-motion workflow* [Application/pdf]. 1.4 MB. <https://doi.org/10.7923/92HF-GP09>
- McCord, S. E., & Pilliod, D. S. (2021). Adaptive monitoring in support of adaptive management in rangelands. *Rangelands*, 44, 1–7. <https://doi.org/10.1016/j.rala.2021.07.003>
- 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(7), 1422–1428. <https://doi.org/10.1890/10-1396.1>
- Miller, D. A. W., Pacifici, K., Sanderlin, J. S., & Reich, B. J. (2019). The recent past and promising future for data integration methods to estimate species' distributions. *Methods in Ecology and Evolution*, 10(1), 22–37. <https://doi.org/10.1111/2041-210X.13110>
- Morris, M., Wheeler-Martin, K., Simpson, D., Mooney, S. J., Gelman, A., & DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. *Spatial and Spatio-temporal Epidemiology*, 31, 100301. <https://doi.org/10.1016/j.sste.2019.100301>
- O'Connor, R. C., Germino, M., Barnard, D. M., Andrews, C. M., Bradford, J. B., Pilliod, D. S., Arkle, R. S., & Shriver, R. K. (2020). Small-scale water deficits after wildfires create long-lasting ecological impacts. *Environmental Research Letters*, 15, 044001. <https://doi.org/10.1088/1748-9326/ab79e4>
- Olsoy, P. J., Zaiats, A., Delparte, D. M., Germino, M. J., Richardson, B. A., Roser, A. V., Forbey, J. S., Cattau, M. E., & Caughlin, T. T. (2024). Demography with drones: Detecting growth and survival of shrubs with unoccupied aerial systems. *Restoration Ecology*, 32, e14106. <https://doi.org/10.1111/rec.14106>
- Paniw, M., García-Callejas, D., Lloret, F., Bassar, R. D., Travis, J., & Godoy, O. (2023). Pathways to global-change effects on biodiversity: New opportunities for dynamically forecasting demography and species interactions. *Proceedings of the Royal Society B: Biological Sciences*, 290(1993), 20221494. <https://doi.org/10.1098/rspb.2022.1494>
- Pebesma, E. J. (2018). Simple features for R: Standardized support for spatial vector data. *The R Journal*, 10(1), 439.
- Perret, J., Besnard, A., Charpentier, A., & Papuga, G. (2023). Plants stand still but hide: Imperfect and heterogeneous detection is the rule when counting plants. *Journal of Ecology*, 111(7), 1483–1496. <https://doi.org/10.1111/1365-2745.14110>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rayburn, A. P., Schiffrs, K., & Schupp, E. W. (2011). Use of precise spatial data for describing spatial patterns and plant interactions in a diverse Great Basin shrub community. *Plant Ecology*, 212(4), 585–594. <https://doi.org/10.1007/s11258-010-9848-0>
- Rhinehart, T. A., Turek, D., & Kitzes, J. (2022). A continuous-score occupancy model that incorporates uncertain machine learning output from autonomous biodiversity surveys. *Methods in Ecology and Evolution*, 13(8), 1778–1789. <https://doi.org/10.1111/2041-210X.13905>
- Rominger, K., & Meyer, S. E. (2019). Application of UAV-based methodology for census of an endangered plant species in a fragile habitat. *Remote Sensing*, 11(6), 719. <https://doi.org/10.3390/rs11060719>
- Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J.-F., de Boissieu, F., & Achim, A. (2020). lidR: An R package for analysis of Airborne Laser Scanning (ALS) data. *Remote Sensing of Environment*, 251, 112061. <https://doi.org/10.1016/j.rse.2020.112061>
- Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60(1), 108–115.
- Royle, J. A. (2009). Analysis of capture-recapture models with individual covariates using data augmentation. *Biometrics*, 65(1), 267–274.

- Schlaepfer, D. R., Lauenroth, W. K., & Bradford, J. B. (2014). Natural regeneration processes in big sagebrush (*Artemisia tridentata*). *Rangeland Ecology & Management*, 67(4), 344–357. <https://doi.org/10.2111/REM-D-13-00079.1>
- Shriver, R. K., Andrews, C. M., Arkle, R. S., Barnard, D. M., Duniway, M. C., Germino, M. J., Pilliod, D. S., Pyke, D. A., Welty, J. L., & Bradford, J. B. (2019). Transient population dynamics impede restoration and may promote ecosystem transformation after disturbance. *Ecology Letters*, 22(9), 1357–1366. <https://doi.org/10.1111/ele.13291>
- Silva, C. A., Hudak, A. T., Vierling, L. A., Loudermilk, E. L., O'Brien, J. J., Hiers, J. K., Jack, S. B., Gonzalez-Benecke, C., Lee, H., Falkowski, M. J., & Khosravipour, A. (2016). Imputation of individual longleaf pine (*Pinus palustris* Mill.) tree attributes from field and LiDAR data. *Canadian Journal of Remote Sensing*, 42(5), 554–573. <https://doi.org/10.1080/07038992.2016.1196582>
- Spiers, A. I., Royle, J. A., Torrens, C. L., & Joseph, M. B. (2022). Estimating species misclassification with occupancy dynamics and encounter rates: A semi-supervised, individual-level approach. *Methods in Ecology and Evolution*, 13(7), 1528–1539. <https://doi.org/10.1111/2041-210X.13858>
- Veran, S., Kleiner, K. J., Choquet, R., Callzo, J. A., & Nichols, J. D. (2012). Modeling habitat dynamics accounting for possible misclassification. *Landscape Ecology*, 27, 943–956.
- Weinstein, B. G., Marconi, S., Aubry-Kientz, M., Vincent, G., Senyondo, H., & White, E. P. (2020). DeepForest: A python package for RGB deep learning tree crown delineation. *Methods in Ecology and Evolution*, 11(12), 1743–1751.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686.
- Wright, W. J., Irvine, K. M., Almborg, E. S., & Litt, A. R. (2020). Modelling misclassification in multi-species acoustic data when estimating occupancy and relative activity. *Methods in Ecology and Evolution*, 11(1), 71–81. <https://doi.org/10.1111/2041-210X.13315>
- Young, D. J. N., Koontz, M. J., & Weeks, J. M. (2022). Optimizing aerial imagery collection and processing parameters for drone-based individual tree mapping in structurally complex conifer forests. *Methods in Ecology and Evolution*, 13(7), 1447–1463.
- Youngflesh, C. (2018). MCMCvis: Tools to visualize, manipulate, and summarize MCMC output. *Journal of Open Source Software*, 3(24), 640.
- Ziegenhagen, L. L., & Miller, R. F. (2009). Postfire recovery of two shrubs in the interiors of large burns in the intermountain west, USA. *Western North American Naturalist*, 69(2), 195–205. <https://doi.org/10.3398/064.069.0208>
- Zipkin, E. F., Thorson, J. T., See, K., Lynch, H. J., Grant, E. H. C., Kanno, Y., Chandler, R. B., Letcher, B. H., & Royle, J. A. (2014). Modeling structured population dynamics using data from unmarked individuals. *Ecology*, 95(1), 22–29. <https://doi.org/10.1890/13-1131.1>

SUPPORTING INFORMATION

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

Appendix S1. Further description of and results from simulation study.

Appendix S2. Additional field sampling details and results.

Figure S1. True (simulated) values of shrub abundance summed over 1600 contiguous sites (x-axis), and estimates (posterior mode) and 95% CI following the model of scenario 1.

Figure S2. Relative bias in relation to the number of verified sampling units following scenario 1 (note, x-axis is jittered).

Figure S3. Root mean-squared error in relation to the number of verified sampling units for scenario 1 (note, x-axis is jittered).

Figure S4. True (simulated) values of shrub abundance summed over 1600 contiguous sites (x-axis), and estimates (posterior mode) and 95% CI following the model of scenario 2.

Figure S5. Relative bias in relation to the number of verified sampling units following scenario 2 (note, x-axis is jittered). Line and shading is a LOESS smoother with SE to guide visualization of the key trends.

Figure S6. Root mean-squared error in relation to the number of verified sampling units for scenario 2 (note, x-axis is jittered). Line and shading is a LOESS smoother with SE to guide visualization.

Figure S7. The observed counts of false-positive observations *Artemisia tridentata* in validated cells plotted against the posterior predictive distributions of false-positives.

Figure S8. The observed abundances in validated cells for *Artemisia tridentata* plotted against the posterior predictive distributions.

Figure S9. The UAS-based counts of *Artemisia tridentata* in the unvalidated cells plotted against the corresponding posterior predictive distributions.

Figure S10. A diagram showing data processing and analysis workflow.

Table S1. Abundance estimates for sagebrush plants across size classes and burn status.

Table S2. Range of site-average density for sagebrush plants across size classes.

Table S3. Parameter inputs for the derivation of canopy height models (CHM) from digital surface models (DSM).

Table S4. Parameter input for individual plant detection and image segmentation.

Table S5. The confusion matrix from the object based image analysis (OBIA).

Table S6. Predicted counts of *Artemisia tridentata* across 10 sites based on the classified imagery and models that account for observation errors.

How to cite this article: Zaiats, A., Caughlin, T. T., Cruz, J., Pilliod, D. S., Cattau, M. E., Liu, R., Rachman, R., Maliha, M., Delparte, D., & Clare, J. D. J. (2024). Propagating observation errors to enable scalable and rigorous enumeration of plant population abundance with aerial imagery. *Methods in Ecology and Evolution*, 15, 2074–2086. <https://doi.org/10.1111/2041-210X.14421>