

## RESEARCH ARTICLE

## Integrating multiple sign types to improve occupancy estimation for inconspicuous species

Mia R. Goldman<sup>1,2,3</sup>  | Matthew Shinderman<sup>4</sup>  | Mackenzie R. Jeffress<sup>5</sup>  | Thomas J. Rodhouse<sup>4,6</sup>  | Kevin T. Shoemaker<sup>1</sup> <sup>1</sup>Department of Natural Resources and Environmental Sciences, University of Nevada – Reno, Reno, Nevada, USA<sup>2</sup>School of Biology, University of St Andrews, St Andrews, UK<sup>3</sup>Department of Biology and Center for Research into Environmental and Ecological Modeling, University of St Andrews, St Andrews, UK<sup>4</sup>Department of Natural Resources and Sustainability, Oregon State University-Cascades, Bend, Oregon, USA<sup>5</sup>Nevada Department of Wildlife, Elko, Nevada, USA<sup>6</sup>National Park Service, Bend, Oregon, USA

## Correspondence

Mia R. Goldman, Department of Natural Resources and Environmental Sciences, University of Nevada – Reno, Reno, Nevada, USA.

Email: [miargoldman@gmail.com](mailto:miargoldman@gmail.com)

## Abstract

Standard occupancy models enable unbiased estimation of occupancy by accounting for observation errors such as missed detections (false negatives) and, less commonly, incorrect detections (false positives). Occupancy models are fitted to data from repeated site visits in which surveyors record evidence of species presence. Use of indirect sign (e.g., scat, tracks) as evidence of presence can vastly improve survey efficiency for inconspicuous species but can also introduce additional sources of error. We developed a “multi-sign” occupancy approach to model the detection process separately for unique sign types and used this method to improve estimates of occupancy dynamics for an inconspicuous species, the American pika (*Ochotona princeps*). We investigated how estimates of pika occupancy and environmental drivers differed under four increasingly realistic representations of the observation process: (1) perfect detection (commonly assumed for modeling pika occupancy), (2) standard occupancy model (single observation process without possibility of false detection), (3) multi-sign with no false detections (non-false positive model), and (4) multi-sign with false detections (full model). For the multi-sign occupancy models, we modeled the detection of each sign type (fresh scat, fresh haypiles, pika calls, and pika sightings) separately as a function of climatic and environmental covariates. Estimates of occupancy processes and inferences about environmental drivers were sensitive to different detection models. Simplified representations of the detection processes generally resulted in higher occupancy estimates and higher turnover rates than the full multi-sign model. Environmental drivers also varied in their influence on occupancy models, where (e.g.) forb cover was estimated to more strongly influence occupancy in the full multi-sign model than the simpler models. As has been reported previously in other contexts, unmodeled heterogeneity in the observation process can lead to biases in occupancy processes and uncertainty in the relationships between occupancy and environmental covariates. Overall, our multi-sign approach to dynamic occupancy modeling, which accounts for spatio-temporal variation in reliability among sign types, has strong potential to generate more realistic estimates of occupancy dynamics for inconspicuous species.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## KEYWORDS

American pika, detection probability, false negative, false positive, indirect sign, multi-sign, occupancy model, population dynamics

## TAXONOMY CLASSIFICATION

Applied ecology, Population ecology

## 1 | INTRODUCTION

Dynamic occupancy models have emerged as one of the most powerful tools to investigate species responses to global change at regional and range-wide scales (MacKenzie, 2002; Nichols et al., 2008; Royle, 2006). These models use binary detection surveys to assess occupancy rates, changes, and drivers of occupancy. Because occupancy models rely on noninvasive survey designs that can be replicated over large areas at relatively low cost, this tool has gained popularity as research questions in conservation biology and ecology have increased in their geographic scope of inference. Occupancy modeling relies on direct (aural or visual) or indirect (e.g., scat, tracks) sign to determine the presence or absence of a species at each survey location. These models are used to advance important conservation and management goals across a wide range of taxa and geographic scales including investigating metapopulation dynamics of potentially climate-sensitive species (Hanski & Gilpin, 1991; Moilanen et al., 1998), determining habitat use of a range of imperiled tropical avian species (Ruiz-Gutierrez et al., 2010), and understanding the breeding status of potentially threatened amphibians in the Greater Yellowstone Area (Gould et al., 2019).

Most occupancy surveys are prone to detection errors, including failing to detect a species when it is present (false negative) or, perhaps less commonly, falsely detecting a species when it is in fact not present (false positive). If not correctly accounted for, detection errors can result in incorrect conclusions regarding occupancy processes (Berigan et al., 2019; MacKenzie et al., 2003; McClintock et al., 2010). Standard occupancy models are able to produce unbiased estimates of occupancy by explicitly modeling the observation process and treating occupancy status as an unobserved or partially observed latent process within a hierarchical modeling framework (Royle & Dorazio, 2008). To effectively tease apart the detection and occupancy processes, these models rely on surveyors conducting repeated visits to a site (one or more resurveys) within a time horizon short enough to assume no change in occupancy status (MacKenzie, 2002). In this framework, both the detection process and the occupancy process can be modeled separately as functions of distinct sets of covariates, thereby allowing researchers to control for detection errors when estimating the drivers of occupancy (MacKenzie et al., 2003).

In occupancy modeling, many analyses lack sufficient data to inform robust estimates of false-positive errors. As a result, it remains common practice to ignore or make unrealistic simplifying assumptions about false-positive error rates when modeling species occurrence probabilities. However, this type of detection error can occur if surveyors misidentify a species via visual or auditory cues, or

misidentify indirect sign (scat, tracks, and other species-specific sign) as indicators of presence. False-positive detection errors are known to occur in standard wildlife surveys (McClintock et al., 2010). For example, use of presence-absence data for 10 anuran species across the Northeastern United States revealed that when false-positive error was not accounted for, occupancy estimates were biased by up to 70% for four of the 10 species studied (Ruiz-Gutierrez et al., 2016). Other studies have documented similar biases resulting from failure to account for false-positive errors (Clement et al., 2014; Royle & Link, 2006). As such, it is critical to model these potential detection errors.

For inconspicuous species, occupancy models tend to rely on indirect signs of presence. For example, Karanth et al. (2011) assessed the distribution of a population of tigers in India using indirect sign such as scat and tracks to inform occupancy models. It is increasingly common for researchers to infer occupancy status on the basis of tracks, scat, hair/fur middens, and even eDNA (Da Silva Neto et al., 2020; Nichols et al., 2008; Schmidt et al., 2013). In addition to these methods being low cost and noninvasive, it is important to use indirect sign where species may be missed if relying on direct sign alone. However, potential for quick decay of sign can add difficulty to assigning an occupied or unoccupied status at a given site and easily leads to higher risk of detection errors. On the other hand, long-term preservation of sign can lead surveyors to mistakenly classify formerly occupied sites as currently occupied. Furthermore, misidentification of indirect sign by surveyors may result in either false-positive or false-negative errors. Finally, additional biases may be induced if environmental drivers differentially influence the reliability of sign types (e.g., precipitation or temperatures may affect scat decay rates), or if geographic regions differ in prevalence of reliable sign types. A “multi-sign” occupancy modeling approach accounting for different detection rates and responses to environmental drivers across sign types may help reduce these sources of bias. Because both direct and indirect sign type detection can vary by region, accounting for each observation process separately as a function of a set of climatic and environmental covariates may allow for a more realistic representation of occupancy.

Nichols et al. (2008) was among the first to discuss the use of multiple sign types for fitting occupancy models. More recently, several studies have harnessed multiple detection methods to enable robust estimation of false detection rates (Guillera-Aroita et al., 2017; Miller et al., 2011). Other studies, such as Clement et al. (2014), have also allowed for false detections, but have not estimated unique false detection rates across multiple sign types, nor have they investigated how environmental drivers differentially affect detection or false detection rates for each unique sign

type. Furthermore, Chambert et al. (2015) and others (e.g., Kéry & Royle, 2021) have reinforced the importance of accounting for false detections and provided a framework and lexicon for building hierarchical false-positive models. Our multi-sign approach expands upon previous studies by modeling the detection process (both true and false detection probabilities) for each sign type as independent functions of unique covariate sets (e.g., environmental drivers). Our approach enables us to evaluate the shifting reliability of each sign type across space and time, and thereby help to develop more informed and robust survey protocols.

While this approach can be applied to any species for which surveys rely on multiple sign types as evidence of occupancy, we test this approach using American pika (*Ochotona princeps*), an inconspicuous species thought to be vulnerable to climate change across portions of its range due to potential upslope range shifts, subsequent vegetation availability, and sensitivity to extreme temperature changes (Beever et al., 2010; Erb et al., 2011, 2014; Jeffress et al., 2013). As such, occupancy models are critical to monitoring these potentially vulnerable populations and assessing the role of environmental and climatic factors as drivers of shifting occupancy patterns and range margins. Pika are an excellent test case for a multi-sign occupancy modeling approach, as surveys for American pika typically monitor multiple indirect sign types—notably, scat and haypiles (piles of vegetation thought to serve as a key winter food resource; Dearing, 1997; Millar & Zwickel, 1972) in addition to direct observations of pika and auditory detections of the unique call produced by this species (Conner, 1985). These sign types differ in regional prevalence—for example, direct observations are common in some areas and rare in others—and are also likely to vary in reliability. Since scat and haypiles can persist in the environment for over 400 years (Castillo et al., 2016; Nichols, 2010), surveyors must be able to discriminate between fresh and old observations of indirect sign types in order to use these observations to infer current occupancy status. Fresh haypiles can be difficult to accurately distinguish from old haypiles, especially in lower elevations (Shinderman, 2015) where a few freshly browsed sprigs on an old pika haypile may be deposited by other sympatric herbivore species such as marmots or woodrats, raising the potential for false detections. Although previous evidence (Nichols, 2010) indicates fresh scat is easily distinguishable from old scat using color and consistency, anecdotal evidence suggests fresh scat decays quickly in the sun and especially at lower elevation sites and that determining the apparent age of scat may be regionally and even patch dependent (M. Jeffress and M. Shinderman, Personal communication, November 2021). Because scat is typically the most abundant sign detected in pika surveys, accurately assessing the age of scat detections and being able to distinguish fresh from old is crucial for generating precise and unbiased occupancy estimates.

In this study, we model occupancy dynamics for American pika using multiple direct and indirect indicators of pika presence (fresh scat, fresh haypiles, calls, and sightings) collected from 2010 to 2021 at five national parks in the Pacific Northwest. Furthermore, we investigate how estimates of pika occupancy trends and environmental

drivers differ under four increasingly realistic representations of the pika observation process: (1) perfect detection (a common assumption for modeling pika occupancy), (2) standard occupancy model (single observation process with no possibility of false detection), (3) a multi-sign occupancy model with no false positive detections (non-false positive model), and (4) a multi-sign occupancy model with false positive detections (full model). For the multi-sign occupancy models, we modeled each observation process separately as a function of climatic and environmental covariates including substrate complexity, season, survey period, and vegetation cover. In addition, we modeled each occupancy process (initial patch occupancy, colonization, and extinction) separately as a function of covariates including temperature, precipitation, forb, rock, and shrub cover. Our primary objective was to model occupancy using both false-positive and false-negative detection probabilities for individual sign type. We also predicted that our “multi-sign” approach for modeling observation error would alter estimates of occupancy processes (e.g., colonization, extinction) and its environmental drivers relative to more standard occupancy models in complex and unexpected ways. We also sought to model occupancy dynamics of pika across our study region as a function of climatic and environmental covariates and to assess how these estimates change under our four different treatments of observation error. We expected that accounting for false-positive detection errors across sign types would reduce estimates of mean occupancy because failure to account for false detections typically results in upward bias in occupancy estimates (Royle, 2006). We anticipate that the results of this study will have broad implications for other systems in which multiple sign types are used to assess regional occupancy dynamics for other inconspicuous species.

## 2 | METHODS

### 2.1 | Study species

The American pika is a small lagomorph inhabiting patchy talus slopes across diverse ecosystems of the western US. American pika do not hibernate and are generalist herbivores that often store vegetation caches called “haypiles” for winter food (Smith, 1974). Food caches serve as both winter diet and insulation in talus and are considered the major source of food for pika during winter (Dearing, 1997; Millar & Zwickel, 1972). American pika have long been considered to exhibit classical metapopulation dynamics (Hanski & Gilpin, 1991; Kreuzer & Huntly, 2003; Moilanen et al., 1998; Peacock & Smith, 1997). In 2010, American pika were considered for listing under threatened or endangered status by the United States Fish and Wildlife Service (Wolf et al., 2007), but were subsequently denied listing in 2010 due to both insufficient evidence of decline across known populations and a lack of survey data across much of their range. As such, considerable attention has been given to addressing how pika respond to climatic changes (Beever et al., 2003, 2011; Jeffress et al., 2013; Schwalm et al., 2016; Wolf et al., 2007) across their range. This

species is a highly sensitive thermoregulator and may adjust their activity levels based on heat and cold exposure (MacArthur & Wang, 1973; Smith, 1974). While they are more commonly found above 2500m, American pika metapopulations are also present in habitat below 2500m (Ray et al., 2016; Rodhouse et al., 2010; Shinderman, 2015; Simpson, 2009). A wide range of factors, operating at different spatial scales, may promote or inhibit the persistence of pika populations including climate, substrate complexity, and available vegetation.

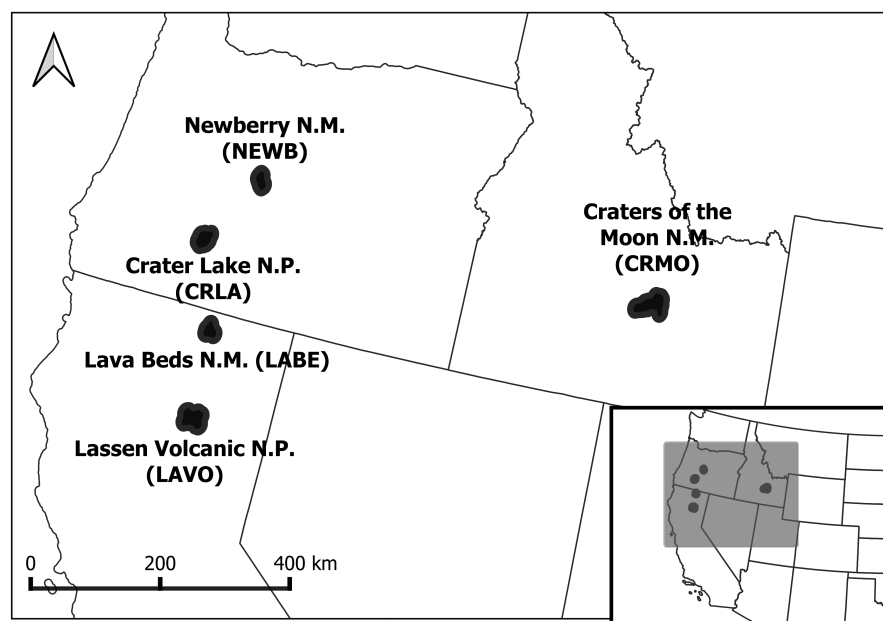
## 2.2 | Occupancy survey and study area

We modeled American pika occupancy dynamics using two related datasets: The first is a 2010–2014 monitoring dataset established by the US National Park Service and the second a 2017–2021 dataset established by the Oregon State University Cascades HERS Lab with support from the National Park Service. Surveys occurred between 2010 and 2021 across five park units within the states of Oregon, Idaho, and California in the Pacific Northwest, including: Crater Lake National Park (CRLA), Craters of the Moon National Monument and Preserve (CRMO), Lassen Volcanic National Park (LAVO), Lava Beds National Monument (LABE), and Newberry National Monument (NEWB; Figure 1). Survey plots were designed to represent a typical pika territory size (Smith & Weston, 1990) and comprised a 12-m radius circular surveyed area (following Jeffress et al. (2010); also used by Rodhouse et al. (2010) and Ray et al. (2012)). An average of 100 plots (hereafter referred to as “sites”) per park were surveyed each year. Resurveys were conducted for 30% of the sites within 1 month of original surveys each year to assess detection probability (Rodhouse et al., 2010). Surveyors spent up to 30min searching a site and recorded up to 10 sign observations per survey (with 10 being considered a reasonable cap based on the density of sign typically observed within the sites and the time it takes to complete the

survey). Observations comprised fresh or old scat (FS/OS), fresh or old haypiles (FH/OH), pika calls (PC), and pika sightings (PS). In practice, there were very few observations in any park with 10 or more instances of sign at a site within a single survey. Furthermore, given the amount of time spent in sites observing sign, it is very unlikely that any haypiles present would not have been detected if they were visible. The most likely scenario in those cases is that there either were no haypiles or they were not detectable (deep below the surface). For sites with multiple surveyors, independent blind surveys or double observer surveys that communicated results together for a combined “up to 30 min” were conducted. With two or more surveyors, surveys were not considered complete until both observers searched all large crevices and beneath large rocks. Surveyors ranged in experience level from newly trained to experienced. Only sign that observers felt they could confidently identify as pika were recorded. If a pika call was heard near the boundary of the site and not clearly within the site, it was determined to be outside the site. Observations of old sign were not included in our analyses.

In addition to collecting records of pika sign, surveyors also classified each survey site according to three levels of rock (talus or lava) complexity: low, intermediate, and high (Figure 2). Lastly, surveyors used Daubenmire's (1959) method to estimate site-level percent cover of rock, bare ground (dirt, mineral soil, and litter), forbs (non-graminoid flowering herbaceous plants), grasses (graminoids [grasses and sedges]), shrubs (woody plants), and trees (Jeffress et al., 2010). A complete description of survey methods can be found in Jeffress et al. (2010) and Rodhouse et al. (2010).

To enable identifiability of false detection rates (we assume identifiability issues would arise if unambiguous detections were not available; Guillerá-Aroita et al. (2017)), we classified sign observations as ambiguous (potentially misidentified) and unambiguous (whereby a false detection was virtually impossible; Rodhouse et al., 2018). We used written notes recorded by surveyors in 2019–2021 surveys (during which note-taking was consistent across all



**FIGURE 1** US Pacific Northwest map identifying 2010–2021 American pika survey areas within Newberry National Monument, Crater Lake National Park, Lava Beds National Monument, Lassen National Volcanic Park and Craters of the Moon National Monument.



**FIGURE 2** Example of low, intermediate, and high site-level complexity in Lava Beds National Monument (1), Lassen Volcanic National Park (2), and Crater Lake National Park (3; Photos: OSU-Cascades HERS Lab Unpublished Protocol Revision 3 Update, June 2019).

**TABLE 1** Occupancy and observation process parameters.

| Occupancy process covariates   |                                     |                            |
|--------------------------------|-------------------------------------|----------------------------|
| Short name                     | Full name                           | Source                     |
| forb                           | Percent forb cover                  | NPS database               |
| rock                           | Percent rock cover                  | NPS database               |
| shrub                          | Percent shrub cover                 | NPS database               |
| cplx <sup>a</sup>              | Rock/substrate complexity           | NPS database               |
| tmaxmean                       | Mean daily max summer temp (°C)     | gridMET (Abatzoglou, 2013) |
| precip                         | Sum daily precip over water yr (mm) | gridMET (Abatzoglou, 2013) |
| Observation process covariates |                                     |                            |
| Survey period                  | pre-2017 or post-2017 survey        | NPS database               |
| ssn                            | season (summer or fall)             | NPS database               |

Note: NPS database data provided by National Park Service, all other covariate data taken from publicly sourced data sources.

<sup>a</sup>Missing complexity data for Lava Beds National Monument (LBE).

surveyors) to classify each sign observation as ambiguous or unambiguous. All fresh scat observations indicating the presence of fresh mucous holding scat together (*stuck*, *pyramid*, *tower*, *perched*, *stacked*, *teepee*, *cluster*) were designated as unambiguous, as such observations are necessarily fresh and are uniquely characteristic of pika scat. Fresh scat detections containing none of these keywords were considered ambiguous (some possibility of false detection i.e., mistaken for scat deposited by pika in a prior year before a site became extirpated). Fresh haypiles were also treated as an ambiguous sign type given the difficulty of knowing whether the vegetation was browse from another species (marmot or woodrat) and uncertainty around the quantity of browse that is to be considered a haypile. All pika sightings were treated as unambiguous. Pika have a distinctive size and shape relative to other small mammals in the study areas, and observers only recorded a pika sighting for high-confidence

observations; if an observer saw a small mammal in the site but could not confirm it was a pika, the observation was not recorded as a pika sighting. In addition, we designated as unambiguous all observations of fresh scat in which one or more fresh haypiles were observed during the same survey (M. Jeffress and M. Shinderman, Personal communication, November 2021).

## 2.3 | Environmental and climatic covariates

Environmental covariates provided by the NPS that were considered in the model included survey season, site-level cover class, and substrate complexity. A fourth covariate, “survey period” was used to distinguish between the 2010–2014 survey period and 2017–2021 period. These periods denote differences in maximum number of surveyors of a given survey. In the pre-2017 period, surveys were conducted by one observer whereas in the post-2017 period 2–4 observers conducted surveys; however, preliminary models suggested little to no effect of observer experience on detection outcomes. Thus, observer experience was not included in the final model. Climatic covariates included in the model, maximum mean summer temperature, and annual sum of daily precipitation for the previous water year were obtained from publicly available online sources (Table 1). Substrate complexity was centered to zero and all other non-binary covariates were scaled for comparison across covariates.

## 2.4 | Statistical analysis

### 2.4.1 | Selection of environmental covariates

Prior to fitting our dynamic occupancy model (see occupancy process model section), we determined which vegetation and climatic variables to use for modeling site-level extinction and colonization processes using an initial model selection procedure (reverse step-wise AIC; Akaike, 1974). We fitted a logistic regression model for the extinction process that included all vegetation cover and climatic variables, and which was fitted using data for all site-year combinations that were occupied in the previous time-step (i.e.,

observations exposed to potential extinction). For this preliminary modeling step, we assumed perfect detection and no false-positive errors (i.e., surveys in which fresh sign was observed were presumed occupied). Similarly, we fitted a model for the colonization process that included all vegetation cover and climatic variables, using data for all site-year combinations that were unoccupied in the previous time-step (i.e., exposed to potential colonization). For each of these models, we then performed reverse step-wise selection with AIC using the step function in R v4.0.1 (R Core Team, 2020) and incorporated the variables remaining in each model to fit the integrated Bayesian occupancy model (Appendix S2, Table 4). We chose covariates for the observation model based on a combination of previous studies demonstrating specific influences on pika occupancy (Erb et al., 2014; Huntly et al., 1986; Jeffress et al., 2013; Rodhouse et al., 2010) and ecological relevance (Appendix S2, Table 5). All quantitative covariates included in the Bayesian occupancy model were <75% correlated.

## 2.4.2 | Occupancy process model

Our process model followed a standard dynamic occupancy modeling framework with three distinct processes: (1) initial occupancy, (2) colonization of previously unoccupied patches, and (3) extinction of previously occupied patches (MacKenzie, 2002; Royle, 2004). Initial occupancy status  $z_{i1}$  (where  $i$  is site) was modeled as a Bernoulli random variable, in which the probability of initial occupancy  $\Lambda_1$  was a logit-linear function of substrate complexity which was shown to influence pika occupancy in (Rodhouse et al., 2010). In subsequent years, the true occupancy status,  $z_{it}$  was modeled as a Bernoulli random variable with a conditional probability based upon prior-year occupancy status  $z_{i,t-1}$ , using:

$$z_{it} \sim \begin{cases} \text{Bernoulli}(1 - \varepsilon_{it}), & \text{if } z_{i,t-1} = 1 \\ \text{Bernoulli}(\gamma_{it}), & \text{if } z_{i,t-1} = 0, \end{cases} \quad (1)$$

where  $\gamma_{it}$  is the probability of colonization in site  $i$  from year  $t - 1$  to year  $t$ , and  $\varepsilon_{it}$  is the probability of extinction in site  $i$  from year  $t - 1$  to year  $t$ . Probabilities of extinction and colonization were also modeled as logit-linear functions of environmental covariates (Table 1). Random intercept terms on colonization and extinction were added to allow for variation among parks and year. A random slope term on initial occupancy was added to allow for among-park variation in the effect of substrate complexity (fully derived equations can be found in Appendix S1).

## 2.4.3 | A “multi-sign” model of the observation process

Each direct and indirect pika sign-type: fresh scat (FS), fresh haypile (FH), pika call (PC), and pika sighting (PS), hereafter, “sign,” was

treated as an independent detection process that reflects the underlying occupancy status of each site. For each sign type, observations were summarized by survey. We treated all survey-level observations as binary (one or more observations of a sign type in a given survey was classified as 1, otherwise 0). We also constructed and fit multi-sign models that treated FS and FH as count data (total number of observations recorded in each survey), but goodness-of-fit tests indicated poor fit and therefore we did not interpret these models further. Probability of detection for each sign type was modeled as a mixture process, allowing for true detections and false-positive detections, following Royle and Link (2006). For each sign type, we modeled the survey-specific probability of correct detection  $p_{\text{sign}}$  and false detection  $p_{\text{sign}}^{\varepsilon}$  as a logit-linear function of covariates:

$$\text{logit}(p_{\text{sign}}) = \mathbf{X}\beta_{\text{sign}}, \quad (2)$$

$$\begin{cases} \text{logit}(p_{\text{sign}}^{\varepsilon}) = \mathbf{X}_{\text{pr}}\beta_{\text{pr}}, & \text{if unambiguous} = 0 \\ p_{\text{sign}}^{\varepsilon} = 0, & \text{if unambiguous} = 1, \end{cases} \quad (3)$$

where  $\mathbf{X}\beta_{\text{sign}}$  stands for a logistic regression of  $p_{\text{sign}}$  and  $p_{\text{sign}}^{\varepsilon}$  for each site, year, and survey as a function of multiple covariates  $\mathbf{X}$  (with regression coefficient vector  $\beta_{\text{sign}}$ ). Thus, if an observation is deemed “unambiguous,” the site is presumed occupied. We modeled the unconditional probability of detection for each sign type ( $pd_{\text{sign}}$ —representing the probability of either false or correct detection of a given sign type) as conditional on the current occupancy status following:

$$pd_{\text{sign}} \sim \begin{cases} p_{\text{sign}}, & \text{if } z = 1 \\ p_{\text{sign}}^{\varepsilon}, & \text{if } z = 0, \end{cases} \quad (4)$$

where the sign-specific probability of detection  $pd_{\text{sign}}$  for each site, year, and survey is equivalent to the probability of correct detection  $p_{\text{sign}}$  if a site is occupied and the probability of false detection  $p_{\text{sign}}^{\varepsilon}$  if a site is unoccupied.

## 2.4.4 | Prior model

We used vague normal priors to define coefficients for all logit-linear predictors for both occupancy and detection processes. We used vague gamma priors to define variance parameters (e.g., hyperpriors on random effect terms) and uniform distributions to initialize all probability parameters (e.g., initial occupancy probability, baseline colonization, and extinction probabilities; Appendix S1).

## 2.4.5 | Posterior predictive check

Model goodness-of-fit was based on posterior predictive checks using a Bayesian posterior predictive  $p$ -value calculating the sum of squared error for fresh scat, fresh haypile, and pika call detections where values close to 0.5 indicate good model fit (Gelman, 2014).

This allowed us to examine the degree to which the dispersion of the observed data matched the dispersion of simulated data. Bayesian  $p$ -values close to 0 or 1 indicate poor model fit and demonstrate weak support for interpretation of parameter estimates.

## 2.4.6 | Bayesian model implementation

We ran models in a Bayesian framework using Markov Chain Monte Carlo (Hooten & Hobbs, 2015). We used the software JAGS (Plummer, 2003) which was called from R using the “jagsUI” wrapper (Kellner, 2021) for “rjags” (Plummer, 2021). Samples were drawn from posterior distributions using three MCMC chains, 100,000 iterations thinned by a factor of 10, and the first 50,000 iterations were discarded as a burn-in period. Chains were checked for convergence, and if  $R$ -hat was  $> 1.1$  for any parameter, the model was rerun (Brooks & Gelman, 1998). We checked model convergence using Gelman-Rubin diagnostics and by examining trace plots of the MCMC chains. All models were run in R software. Our MCMC results indicated convergence ( $R$ -hat  $< 1.1$ ) for all model parameters and for all three alternative representations of detection probability. Bayesian  $p$ -values indicated adequate model fit for all three detection models and unique sign types ( $.42 < p < .50$ ; Kéry, 2010).

## 2.4.7 | Comparing alternative observation models

To test the extent to which our most realistic and sophisticated model influenced the observation process relative to more conventional models, we ran three additional models with increasingly simple representations of the observation process: (1) non-false-positive model: a “multi-sign” detection model with no potential for false detections (all direct or indirect pika observations imply that a site was currently occupied in a given year), (2) standard occupancy model: a model with a single observation process and no possibility of false detection, and (3) perfect detection model: a model in which all pika detection events were assumed to accurately reflect the underlying occupancy status and all surveys yielding no pika detections were assumed to reflect pika absence. In general, high detectability has been assumed for pika and thus detection error has generally been ignored in this modeling framework (Jeffress et al., 2013; Ray et al., 2016; Rodhouse et al., 2010). We computed site-occupancy rates for each park and year by dividing the estimated number of occupied sites in each park (determined by monitoring the binary site-occupancy status for each MCMC iteration) by the total number of monitored sites at each park. We then summarized occupancy rates for each park and year as a posterior mean (point estimate) and a 90% credible interval. We then ran weighted linear regressions of mean park-level occupancy as a function of year (one regression model per park), using an inverse-variance weighting technique (weights equal to the inverse of posterior variance of each site-occupancy estimate), and interpreted significant trends at  $\alpha = 0.1$ .

## 3 | RESULTS

Overall, our survey dataset included a total of 5126 surveys of 771 unique sites across 5 parks and 12 years (Table 2).

Of these surveys, at least 30% were resurveys conducted within 2 weeks of the original survey to ensure no occupancy turnover. Mean maximum summer temperature (June–August) at our study sites ranged between 16.44 and 30.17°C. The sum of annual precipitation ranged from 101 to 5130 mm (Appendix S2, Figure 5).

### 3.1 | Pika occupancy dynamics

The colonization rate was 5.8%, and extinction rate was 6.2%. Across our study period, we detected two significant park-level trends: occupancy increased from 19% to 40% at LABE and decreased from 40% to 22% in NEWB. There was no detectable trend in occupancy across all parks and occupancy ranged from 32 to 40% (Figure 3; Appendix S2, Table 6 and 7).

Rock complexity was positively correlated with initial occupancy ( $\alpha_{\text{complexity}} = 1.325$ , 95% CRI: 0.773 to 1.944,  $\omega^+ = 1.0$ ; where  $\omega^-$  or where  $\omega^+$  indicate weight of evidence above or below zero and point estimates represent the posterior mean). Percent forb cover was negatively associated with the probability of extinction ( $\alpha_{\text{forb}} = -2.700$ , 95% CRI: -4.458 to -1.265,  $\omega^- = 0.988$ ) and weakly associated with probability of colonization ( $\alpha_{\text{forb}} = 0.033$ , 95% CRI: -0.303 to 0.304,  $\omega^+ = 0.877$ ). Shrub cover ( $\alpha_{\text{shrub}} = 0.255$ , 95% CRI: 0.044 to 0.470,  $\omega^+ = 0.989$ ) was positively correlated with colonization (Appendix S1, Table 7).

Correlation in among-year process variation (colonization and extinction processes) between parks varied from 0.63 to -0.43 (Appendix S2, Table 8). The highest positive correlations (0.63 for colonization, 0.61 for extinction) were observed between the two parks with highest average temperatures and lowest average precipitation (LABE and CRMO). Other parks with positively correlated annual colonization rates included CRMO and CRLA (0.564), CRLA and LAVO (0.561), and CRLA and NEWB (0.575; Appendix S2, Table 8).

### 3.2 | Observation process

Total detection probability across sign types was 82%. Overall sign detection rates, individual sign type detection rate, and associated false-positive error rates varied among parks. The average probability of correct detection in CRLA was the highest at 92.4% and lowest in NEWB at 77.1% (Table 3).

Rate of fresh scat detections ( $p_{\text{fs}} = 0.801$ , 95% CRI: 0.764 to 0.801) was substantially higher than fresh haypiles ( $p_{\text{th}} = 0.288$ , 95% CRI: 0.248 to 0.331). Mean detection rate for pika calls ( $p_{\text{pc}} = 0.412$ , 95% CRI: 0.375 to 0.451) was substantially higher than pika sightings ( $p_{\text{ps}} = 0.134$ , 95% CRI: 0.110 to 0.160). Survey period was negatively associated with number of fresh scat detections ( $\beta_{\text{FS}_{\text{survey period}}} = -0.684$ , 95% CRI: -0.927 to -0.442,  $\omega^- = 1.0$ ). Seasonality (summer vs. fall) affected

TABLE 2 Park surveys by year.

|                              | 2010 | 2011 | 2012 | 2013 | 2014 | 2017 | 2018 | 2019 | 2020 | 2021 |
|------------------------------|------|------|------|------|------|------|------|------|------|------|
| Original surveys across park |      |      |      |      |      |      |      |      |      |      |
| CRLA <sup>a</sup>            | 85   | 108  | 141  | 101  | 102  | 0    | 95   | 100  | 0    | 96   |
| CRMO <sup>b</sup>            | 56   | 135  | 146  | 103  | 103  | 103  | 99   | 96   | 100  | 100  |
| LABE <sup>c</sup>            | 101  | 101  | 100  | 100  | 100  | 99   | 99   | 100  | 89   | 50   |
| LAVO <sup>d</sup>            | 76   | 113  | 75   | 103  | 103  | 101  | 100  | 98   | 49   | 0    |
| NEWB <sup>e</sup>            | 0    | 0    | 0    | 0    | 0    | 100  | 102  | 102  | 0    | 100  |
| Resurveys across parks       |      |      |      |      |      |      |      |      |      |      |
| CRLA                         | 50   | 30   | 30   | 41   | 30   | 0    | 29   | 29   | 0    | 0    |
| CRMO                         | 50   | 35   | 30   | 20   | 29   | 16   | 30   | 29   | 30   | 30   |
| LABE                         | 50   | 0    | 30   | 14   | 50   | 27   | 31   | 30   | 0    | 18   |
| LAVO                         | 41   | 30   | 0    | 21   | 20   | 27   | 29   | 29   | 0    | 0    |
| NEWB                         | 0    | 0    | 0    | 0    | 0    | 30   | 31   | 31   | 0    | 30   |

<sup>a</sup>Crater Lake National Park.<sup>b</sup>Craters of the Moon National Monument and Preserve.<sup>c</sup>Lava Beds National Monument.<sup>d</sup>Lassen Volcanic National Park.<sup>e</sup>Newberry National Monument: surveys began in 2017.

the probability of detecting pika calls ( $\beta_{PC_{season}} = -0.544$ , 95% CRI:  $-0.749$  to  $-0.344$ ,  $\omega^- = 1.0$ ) and sightings ( $\beta_{PS_{season}} = -0.584$ , 95% CRI:  $-0.906$  to  $-0.273$ ,  $\omega^- = 1.0$ ). Sites with higher substrate complexity tended to have fewer detectable fresh haypile detections ( $\beta_{FH_{complexity}} = -0.262$ , 95% CRI:  $-0.494$  to  $-0.025$ ,  $\omega^- = 1.0$ ), pika calls ( $\beta_{PC_{complexity}} = -0.273$ , 95% CRI:  $-0.476$  to  $-0.072$ ,  $\omega^- = 1.0$ ), and pika sightings ( $\beta_{PS_{complexity}} = -0.306$ , 95% CRI:  $-0.583$  to  $-0.030$ ,  $\omega^- = 0.985$ ). Lastly, higher forb cover was associated with a higher probability of detecting haypiles ( $\beta_{FH_{forb}} = 0.336$ , 95% CRI:  $0.239$  to  $0.438$ ,  $\omega^+ = 1.0$ ; Appendix S2, Figures 8–11; Table 7).

The average overall false-positive rate was 8.6%. The average false-positive rate was highest in NEWB (0.117) and lowest in CRMO (0.028; Table 3). The highest false-positive error rates were associated with pika call ( $p_{PC}^E = 0.042$ , 95% CRI: 0.031 to 0.055) followed by fresh scat ( $p_{FS}^E = 0.037$ , 95% CRI: 0.020 to 0.055). The mean false-positive rate for fresh haypiles was extremely low ( $p_{FH}^E = 0.020$ , 95% CRI: 0.012 to 0.031). False-positive scat detections tended to be higher during the 2017–2021 surveys versus the 2010–2014 survey period ( $\beta_{FS_{season}}^E = -1.139$ , 95% CRI:  $-2.077$  to  $-0.358$ ,  $\omega^- = 1.0$ ). False pika call detections were more likely in summer than in fall ( $\beta_{PC_{season}}^E = -1.069$ , 95% CRI:  $-1.678$  to  $-0.529$ ,  $\omega^- = 1.0$ ). Higher substrate complexity was associated with fewer false pika call detections ( $\beta_{PC_{complexity}}^E = -0.228$ , 95% CRI:  $-0.640$  to  $0.193$ ,  $\omega^+ = 0.974$ ). Lastly, false-positive fresh haypile detections were more common in areas with higher forb cover ( $\beta_{FH_{forb}}^E = 0.333$ , 95% CRI:  $0.143$  to  $0.503$ ,  $\omega^+ = 1.0$ ; Appendix S2, Table 7).

### 3.3 | Comparing alternative observation models

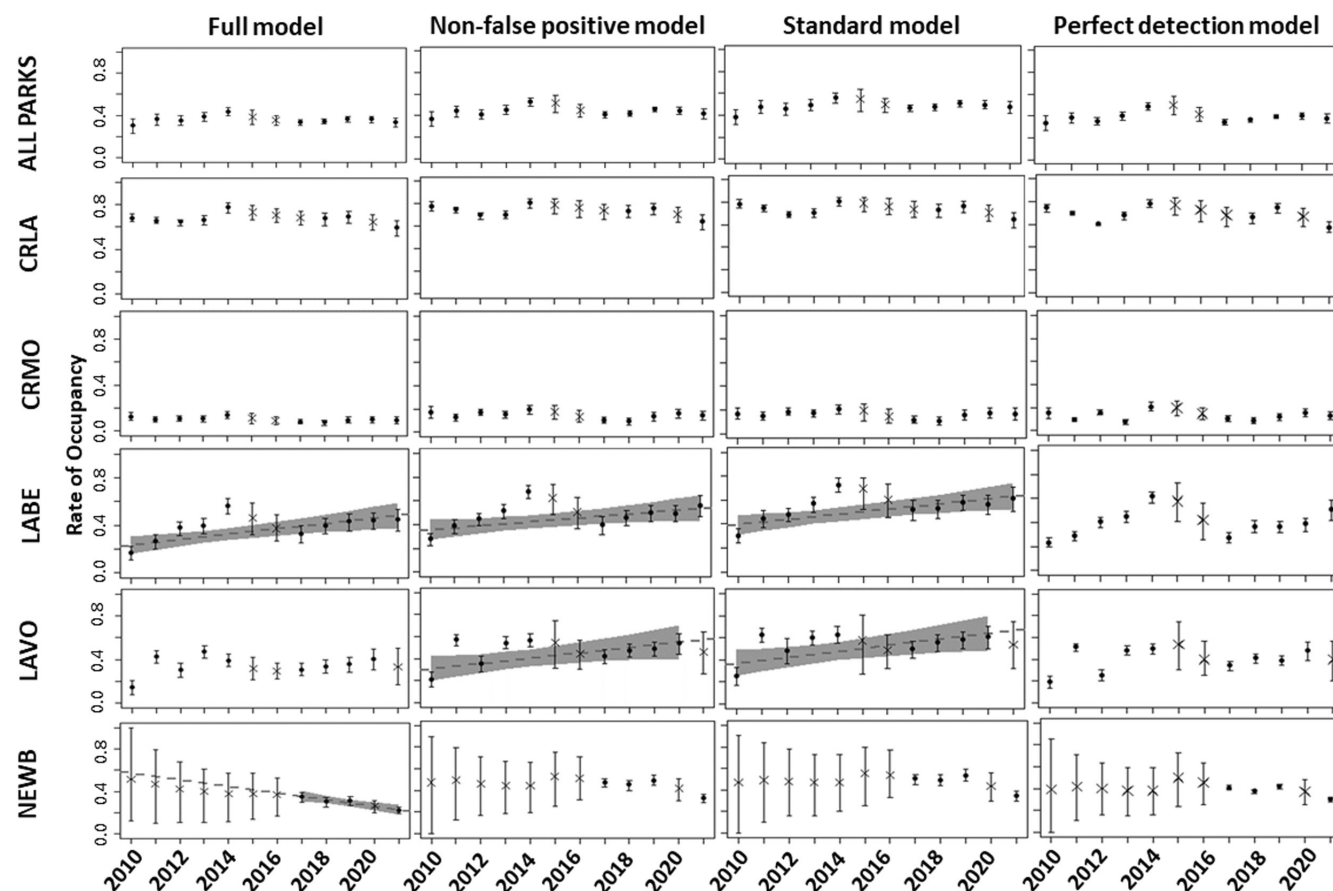
Temporal (year to year) variance across parks for mean rate of extinction tended to be greatest in the full model ( $\mu_{\sigma}^E = 0.65$ ) and lowest

in the perfect detection model ( $\mu_{\sigma}^E = 0.40$ ), whereas variance across park for colonization was highest in the non-false-positive model ( $\mu_{\sigma}^V = 0.65$ ) and lowest in the perfect detection model ( $\mu_{\sigma}^V = 0.38$ ; Appendix S2). Variance across park for initial occupancy was similar in the full ( $\mu_{\sigma}^A = 1.13$ ) and perfect detection model ( $\mu_{\sigma}^A = 1.14$ ) and lowest in the non-false-positive model ( $\mu_{\sigma}^A = 0.91$ ). The estimate for initial occupancy in the full model was 28.4%, compared with 37.0% in the non-false positive model, 39.5% in the standard detection model and 33.2% in the perfect detection model. Mean colonization rate was lowest in the full model (8%) and highest in the perfect detection model (16%). Mean extinction rate was lowest in the full model (6.2%) and highest in the perfect detection model (26.4%; Appendix S2, Table 7, 11, 13, 14).

Percent forb cover in the non-false-positive model indicated an increased negative effect on extinction ( $\alpha_{\epsilon_{forb}} = -1.948$ , CRI:  $-3.098$  to  $-1.065$ ,  $\omega^- = 1.0$ ) and (shrub cover; Figure 6 and summarized in Appendix S2, Table 14). A notable difference in the perfect detection model included producing the weakest negative effects of forb cover on extinction ( $\alpha_{\epsilon_{forb}} = -0.420$ , CRI:  $-0.679$  to  $-0.185$ ,  $\omega^- = 1.0$ ; Appendix S2, Table 14). Compared with the full model, the standard deviation of fraction of occupied sites across all parks was higher ( $\sigma^2 = 0.0527$ ) than the full model ( $\sigma^2 = 0.0351$ ).

## 4 | DISCUSSION

The processes regulating the reliability and detectability of direct and indirect sign types are likely to vary across space and time even when identical protocols are applied to target organisms. When occupancy status is determined on the basis of multiple sign types, these differences could result in biased estimates of occupancy and its environmental drivers. A multi-sign occupancy modeling



**FIGURE 3** Comparison of posterior site-occupancy rates (filled circles represent posterior means, error bars represent 90% credible intervals) through time for all parks across all models. Shaded regions indicate post hoc linear regression models with a 90% C.I. (using inverse-variance weighting to account for uncertainty of point estimates). Only those regression relationships significant at  $\alpha=0.1$  are depicted. Points represented with an "X" indicate years that the park was not surveyed.

**TABLE 3** Detection rates<sup>a</sup> across park and sign type fresh scat (FS), fresh haypiles (FH), pika calls (PC), and pika sightings (PS): full model.

| Park | True detection rate |       |       |       |       | False-positive detection rate |       |       |       |
|------|---------------------|-------|-------|-------|-------|-------------------------------|-------|-------|-------|
|      | Overall             | FS    | FH    | PC    | PS    | Overall                       | FS    | FH    | PC    |
| CRLA | 0.924               | 0.806 | 0.309 | 0.599 | 0.194 | 0.110                         | 0.017 | 0.020 | 0.076 |
| CRMO | 0.785               | 0.722 | 0.366 | 0.091 | 0.016 | 0.028                         | 0.007 | 0.017 | 0.004 |
| LABE | 0.764               | 0.719 | 0.150 | 0.014 | 0.004 | 0.065                         | 0.045 | 0.015 | 0.005 |
| LAVO | 0.859               | 0.626 | 0.404 | 0.345 | 0.060 | 0.108                         | 0.018 | 0.049 | 0.043 |
| NEWB | 0.771               | 0.540 | 0.144 | 0.301 | 0.106 | 0.117                         | 0.016 | 0.016 | 0.086 |
| AVG  | 0.820               | 0.683 | 0.275 | 0.270 | 0.076 | 0.086                         | 0.020 | 0.023 | 0.043 |

<sup>a</sup>Using the raw data of sign type observations for each site/year combination, the probability of either true or false detection was determined conditional on the site being occupied or unoccupied (per the posterior distribution for the binary occupancy variable at each unique site and year). Probabilities were computed as weighted averages across surveys.

approach like the one we present in this study allows unique detection processes to be modeled separately in order to more effectively model occupancy dynamics (MacKenzie et al., 2003). We found that distinct sign types exhibited substantial differences in overall reliability, false positive rates, and influence of environmental covariates.

Although occupancy models involving surveys of direct and indirect sign types typically ignore this variation—instead, treating all sign types equally in a pooled observation process—this practice

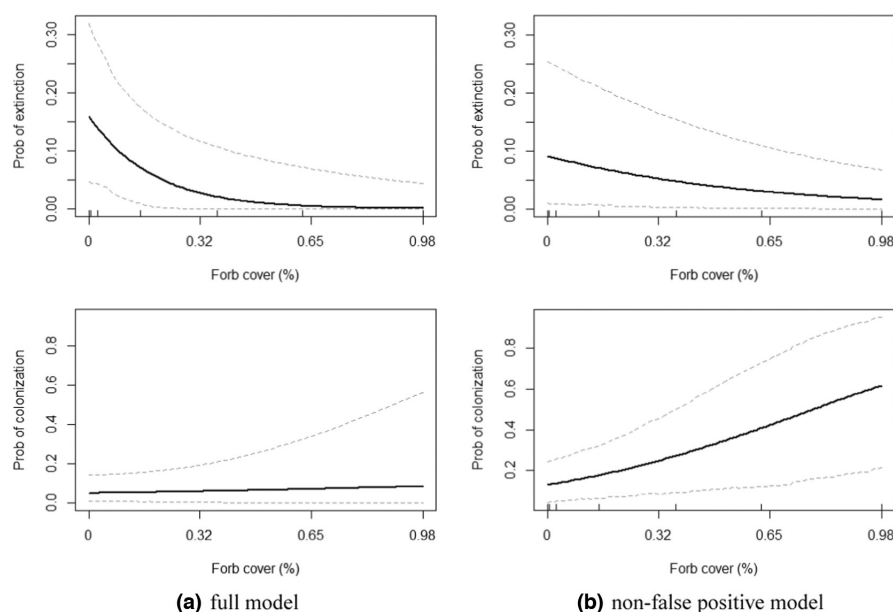
may introduce previously unexplored sources of bias in occupancy estimation if different sign types vary in prevalence across sites and regions. We found that overall survey reliability varied substantially among the parks included in our study despite survey protocols being standardized across all parks, with mean true detection rates varying from 77.1% to 92.4% and mean false detection rates varying from 2.8% to 11.7% for different parks. This variation in survey reliability was presumably driven by differences in the reliability and

prevalence of different sign types in each park (e.g., haypiles; see below). For example, detection rates of pika calls by park varied from 1.4% to 59.9%, and pika sightings varied from 0.4% to 19.4%. By modeling each distinct sign type separately, our “multi-sign” occupancy models are able to accommodate substantial regional differences in survey reliability while producing unbiased estimates of occupancy processes. For instance, in the American pika system, haypiling may be particularly sensitive to environmental gradients. If not modeled explicitly this could cloud estimates of occupancy trends.

While false detections are ignored under most standard occupancy modeling frameworks (MacKenzie, 2002), researchers are increasingly allowing for false-positive errors within occupancy modeling frameworks and acknowledging the important biases that can result from ignoring false-positive errors (Berigan et al., 2019; Chambert et al., 2015; Clement et al., 2014; Miller et al., 2013; Royle & Link, 2006). Our study echoes these findings, suggesting that false detection rates can have strong effects on key inferences about occupancy and occupancy dynamics, including mean occupancy rates, the magnitude and significance of occupancy trends, and the environmental drivers of occupancy. Estimates of several key environmental drivers were also sensitive to changes in how the observation process was represented. For example, the strength of the estimated effect of substrate complexity on initial occupancy probability was much stronger (nearly double in magnitude) in the full model versus the perfect detection model. Finally, the estimated effect of forb cover on patch-level extinction risk was also far stronger in the full model versus the perfect detection model; whereas the full model suggests that forb cover  $\geq$  ca. 65% implies a nearly 100% probability of site persistence, the perfect detection model implies that forb cover barely influences the probability of extinction (Figure 4). Overall, these results suggest that estimates of occupancy processes are sensitive to how sign types vary in reliability across geographic and environmental space, and failure to account for this variation may lead to biased or incomplete estimates of key

occupancy processes. For example, forb cover more strongly influenced occupancy in the full model than in the non-false positive, standard, and perfect detection models (Appendix S2, Table 7–14). Multi-sign occupancy models should yield a more realistic understanding of occupancy dynamics in cases where detection processes differ in reliability and response to environmental gradients and vary in prevalence across study sites or regions.

Most previous occupancy models for pika and other inconspicuous species have ignored false detections based on the assumption of low false detection rates. In our model, misidentification of pika calls was surprisingly prevalent, averaging 4.3% and varying substantially among our study sites (0.5% to 8.6%). When ignoring the potential for false detections, mean occupancy rates increased by 9% across the study region relative to the model that accommodated false detection errors, most likely reflecting the upward biases induced by treating false detections as true. In contrast to our assumption that false detections would be driven primarily by indirect sign observations, one of the largest sources of false positive errors in our study was direct observation of pika calls. Coupled with the knowledge that pika calls were relatively uncommon relative to other sign types (27%, compared to fresh scat at 68.3%), this suggests that future survey protocols may wish to focus on other sign types as primary indicators of occupancy. On the other hand, the most likely explanation for the high false detection rate for pika calls is that surveyors correctly detected a pika but that the call emerged from outside the survey site; pika calls are unlikely to be confused with calls made by any other species. Therefore, ignoring false detections of pika calls may not pose a severe issue for occupancy estimation if researchers are more interested in estimating occupancy beyond the level of the standard survey site. Among the indirect sign types, fresh haypiles were associated with the highest rate of false detections (2.3%). These false detections are likely the result of mistaking old haypiles for fresh, and such errors are likely to induce upward biases in estimates of regional occupancy if not modeled explicitly. Rigorous



**FIGURE 4** Effects plots comparing the relationships between occupancy processes (colonization and extinction probabilities) and forb cover for (a, left) a multi-sign occupancy model with false detections and (b, right) an otherwise-equivalent model with false detection rates assumed to be zero. Dashed lines indicate 95% Bayesian credible interval.

simulation studies are needed to determine thresholds for both true and false detection rates beyond which these errors can be safely ignored. In the meantime, we suggest that researchers should incorporate both false-negative and false-positive detection errors in occupancy models unless strong evidence suggests that one or both of these errors can be ignored.

At the site level, detection probabilities for this species are generally thought to be high ( $\geq 0.9$ ) across their range (Beever et al., 2008, 2010; Ray et al., 2012), leading many pika researchers to disregard detection errors as inconsequential sources of bias for estimating pika occupancy and metapopulation dynamics (Erb et al., 2011; Jeffress et al., 2013; Ray et al., 2016; Rodhouse et al., 2010; Shinderman, 2015). This assumption has been justified on the basis of previous findings that pika are detectable at a rate of  $\geq 90\%$  across their range (Beever et al., 2008, 2011; Jeffress et al., 2013; Ray et al., 2016; Rodhouse et al., 2010; Wilkening & Ray, 2016). Rodhouse et al., 2010 is the only one of these studies to actually estimate detection probability with replicate survey data. More recent research has suggested that the assumption of perfect detection may lead to biased results in occupancy and metapopulation models for this species due to imperfect detection as well as occasional false detections (Rodhouse et al., 2018). While our results confirm that the detection rate is high (82% true detection rate) for pika surveys, our results also suggest that failing to account for imperfect detection for pika in our study region can lead to biased estimates of occupancy trends and covariate effects and can result in substantially larger inter-annual variance relative to models that account for imperfect detection. Given our results, we caution against ignoring observation errors and suggest that it is critical to model the observation process when study goals involve estimating occupancy dynamics and its environmental drivers, even for species with detection rates approaching or exceeding 90%.

In our “multi-sign” occupancy modeling framework, we assumed that all unique sign types represented independent indicators of the underlying occupancy status. This implies that the probability of failing to detect all unique sign types during a single survey (probability of false negative) could be computed as the product of the probability of failing to detect each individual sign type, and that the probability of a false detection could be computed as the inverse of the product of correctly failing to observe each individual sign type. However, interactions among sign types are possible, and even likely; for example, two or more unique sign types may be extremely reliable indicators of occupancy when they co-occur, whereas neither is reliable separately. In our study system, surveys in which both fresh scat and haypiles co-occur were considered to be an unambiguous indicator of occupancy—which we modeled by disallowing false detections for such surveys. However, future multi-sign occupancy frameworks should consider modeling interactions among sign types explicitly. Such models may more realistically describe how information from surveys with multiple distinct sign types can be used to most precisely describe the underlying observation process.

Failure to allow for detection errors that can and do occur in real-world ecological surveys, such as false detections and use of

sign types that vary in reliability, can lead to biased or incomplete estimates of key occupancy processes. Building on other occupancy models that handle false detections (Clement et al., 2014; Rodhouse et al., 2018), and multiple unique detection methods (Guillera-Arroita et al., 2017; Kéry & Royle, 2021; Miller et al., 2011), we present a flexible approach for modeling detection probability for multiple sign types, whereby detection probability can respond to spatio-temporal environmental gradients uniquely for each sign type. Unlike previous studies, our method also allows any sign type to be present in both ambiguous and unambiguous forms, enabling additional power to estimate false-positive detection rates (Guillera-Arroita et al., 2017). In theory, knowledge of the degree of ambiguity of different sign types could be used to improve survey efficiency by (e.g.) resurveying sites only until an unambiguous sign is observed (ceasing surveys at that point), allowing more effort to be put into other sites. However, in the context of monitoring pika populations, this might not save that much time, as multiple sign types, when present, can typically be observed and recorded quickly.

We anticipate that our method has broad applicability for other systems in which multiple sign types are used to assess regional occupancy dynamics, especially for inconspicuous species. As hierarchical models (often fitted in a Bayesian framework) become increasingly accessible to ecologists, incorporation of more realistic and complex models of the observation process, such as multi-sign occupancy models, have become relatively straightforward to implement. We believe researchers can build upon our framework to develop more precise detection models that ultimately improve our ability to estimate regional occupancy and to develop more effective regional monitoring protocols. Ultimately, incorporation of multi-sign detection probabilities in occupancy models has great potential to positively impact management and conservation decisions for inconspicuous species.

## AUTHOR CONTRIBUTIONS

**Mia R. Goldman:** Conceptualization (equal); formal analysis (lead); visualization (lead); writing – original draft (lead); writing – review and editing (equal). **Matthew Shinderman:** Conceptualization (supporting); data curation (lead); investigation (supporting); methodology (supporting); resources (supporting); writing – review and editing (supporting). **Mackenzie R. Jeffress:** Conceptualization (supporting); data curation (supporting); methodology (supporting); writing – review and editing (supporting). **Thomas J. Rodhouse:** Conceptualization (supporting); data curation (supporting); methodology (supporting); resources (supporting); writing – review and editing (supporting). **Kevin T. Shoemaker:** Conceptualization (equal); formal analysis (supporting); investigation (supporting); methodology (supporting); supervision (lead); validation (supporting); visualization (supporting); writing – original draft (supporting); writing – review and editing (supporting).

## ACKNOWLEDGMENTS

We thank the Oregon State University Cascades HERS Lab for data collection and management, and for all their support throughout this

project. We also thank the National Park Service for data collection. M.R.G. was funded in part by a graduate research assistantship and the Robert E. Dickenson scholarship through the University of Nevada Reno. Funding for this project was provided by National Park Service Upper Columbia Basin Network Inventory and Monitoring Program through a Cooperative Agreement (Task Agreement P18AC01062).

## DATA AVAILABILITY STATEMENT

All data were collected by the OSU-Cascades HERS Lab with support from the National Park Service under Cooperative Agreement P18AC01062. Use of these data should include an attribution of source to the OSU Cascades Human and Ecosystem Resilience and Sustainability Lab. Dryad repository: <https://doi.org/10.5061/dryad.ttdz08m2n>.

## ORCID

Mia R. Goldman  <https://orcid.org/0000-0003-3356-0859>

Matthew Shinderman  <https://orcid.org/0000-0003-1522-1976>

Mackenzie R. Jeffress  <https://orcid.org/0000-0001-5090-7530>

Thomas J. Rodhouse  <https://orcid.org/0000-0001-5953-9113>

Kevin T. Shoemaker  <https://orcid.org/0000-0002-3789-3856>

## REFERENCES

- Abatzoglou, J. T. (2013). Development of gridded surface meteorological data for ecological applications and modelling. *International Journal of Climatology*, 33(1), 121–131. <https://doi.org/10.1002/joc.3413>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Beever, E. A., Brussard, P. F., & Berger, J. (2003). Patterns of apparent extirpation among isolated populations of Pikas (*Ochotona Princeps*) in the Great Basin. *Journal of Mammalogy*, 84(1), 37–54.
- Beever, E. A., Chris, R. A., Mote, P. W., & Wilkening, J. L. (2010). Testing alternative models of climate-mediated extirpations. *Ecological Applications*, 20(1), 164–178.
- Beever, E. A., Ray, C., Wilkening, J. L., Brussard, P. F., & Mote, P. W. (2011). Contemporary climate change alters the pace and drivers of extinction. *Global Change Biology*, 17(6), 2054–2070.
- Beever, E. A., Tausch, R. J., & Thogmartin, W. E. (2008). Multi-scale responses of vegetation to removal of horse grazing from Great Basin (USA) mountain ranges. *Ecology*, 196(2), 163–184.
- Berigan, W. J., Jones, G. M., Whitmore, S. A., Gutiérrez, R. J., & Peery, M. Z. (2019). Cryptic wide-ranging movements lead to upwardly biased occupancy in a territorial species. *Journal of Applied Ecology*, 56(2), 470–480. <https://doi.org/10.1111/1365-2664.13265>
- Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434–455.
- Castillo, J. A., Epps, C. W., Jeffress, M. R., Ray, C., Rodhouse, T. J., & Schwalm, D. (2016). Replicated landscape genetic and network analyses reveal wide variation in functional connectivity for American pikas. *Ecological Applications*, 26(6), 1660–1676.
- Chambert, T., Miller, D. A., & Nichols, J. D. (2015). Modeling false positive detections in species occurrence data under different study designs. *Ecology*, 96, 332–339. <https://doi.org/10.1890/14-1507.1>
- Clement, M. J., Rodhouse, T. J., Ormsbee, P. C., Szewczak, J. M., & Nichols, J. D. (2014). Accounting for false-positive acoustic detections of bats using occupancy models. *Journal of Applied Ecology*, 51(5), 1460–1467. <https://doi.org/10.1111/1365-2664.12303>
- Conner, D. A. (1985). Analysis of the vocal repertoire of adult pikas: Ecological and evolutionary perspectives. *Animal Behaviour*, 33(1), 11–134. [https://doi.org/10.1016/S0003-3472\(85\)80126-3](https://doi.org/10.1016/S0003-3472(85)80126-3)
- Da Silva Neto, J. G., Sutton, W. B., Spear, S. F., Freaque, M. J., Kéry, M., & Schmidt, B. R. (2020). Integrating species distribution and occupancy modeling to study hellbender (*Cryptobranchus alleganiensis*) occurrence based on eDNA surveys. *Biological Conservation*, 251, 108787. <https://doi.org/10.1016/j.biocon.2020.108787>
- Daubenmire, R. F. (1959). Canopy coverage method of vegetation analysis. *Northwest Science*, 33, 39–64.
- Dearing, M. D. (1997). The function of haypiles of pikas (*Ochotona princeps*). *Journal of Mammalogy*, 78(4), 1156–1163. <https://academic.oup.com/jmammal/article/78/4/1156/871854>
- Erb, L. P., Ray, C., & Guralnick, R. (2011). On the generality of a climate-mediated shift in the distribution of the American pika (*Ochotona princeps*). *Ecology*, 92(9), 1730–1735. <https://doi.org/10.1890/11-0175.1>
- Erb, L. P., Ray, C., & Guralnick, R. (2014). Determinants of pika population density vs. occupancy in the southern Rocky Mountains. *Ecological Applications*, 24(3), 429–435. <https://doi.org/10.1890/13-1072.1>
- Gelman, A. (2014). *Bayesian data analysis* (3rd ed.). Taylor & Francis.
- Gould, W. R., Ray, A. M., Bailey, L. L., Thoma, D., Daley, R., & Legg, K. (2019). Multistate occupancy modeling improves understanding of amphibian breeding dynamics in the greater yellowstone area. *Ecological Applications*, 29(1), e01825. <https://doi.org/10.1002/eap.1825>
- Guillera-Aroita, G., Lahoz-Monfort, J. J., van Rooyen, A. R., Weeks, A. R., & Tingley, R. (2017). Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods in Ecology and Evolution*, 8(9), 1081–1091. <https://doi.org/10.1111/2041-210X.12743>
- Hanski, I., & Gilpin, M. (1991). Metapopulation dynamics: Brief history and conceptual domain. *Biological Journal of the Linnean Society*, 42(1–2), 3–16.
- Hooten, M. B., & Hobbs, N. T. (2015). A guide to Bayesian model selection for ecologists. *Ecological Monographs*, 85(1), 3–28.
- Huntly, N. J., Smith, A. T., & Ivins, B. L. (1986). Foraging behavior of the Pika (*Ochotona princeps*), with comparisons of grazing versus hay-ing. *Journal of Mammalogy*, 67(1), 139–148.
- Jeffress, M. R., Garrett, L. K., & Rodhouse, T. (2010). *Pika monitoring at Crater Lake National Park, Craters of the Moon National Monument and Preserve, Lassen Volcanic National Park, and Lava Beds National Monument 2010 Annual Monitoring Report*. US Department of the Interior, National Park Service.
- Jeffress, M. R., Rodhouse, T. J., Ray, C., Wolff, S., & Epps, C. W. (2013). The idiosyncrasies of place: Geographic variation in the climate-distribution relationships of the American pika. *Ecological Applications*, 23(4), 864–878. <https://doi.org/10.1890/12-0979.1>
- Karanth, K. U., Gopalaswamy, A. M., Kumar, N. S., Vaidyanathan, S., Nichols, J. D., & Mackenzie, D. I. (2011). Monitoring carnivore populations at the landscape scale: Occupancy modelling of tigers from sign surveys. *Journal of Applied Ecology*, 48, 1048–1056. <https://doi.org/10.1111/j.1365-2664.2011.02002.x>
- Kellner, K. (2021). *Jagsui: A wrapper around 'rjags' to streamline 'jags' analyses* [R package version 1.5.2]. <https://CRAN.R-project.org/package=jagsUI>
- Kéry, M. (2010). *Introduction to WinBUGS for ecologists: Bayesian approach to regression, ANOVA, mixed models and related analyses*. Elsevier Science.
- Kéry, M., & Royle, J. (2021). *Applied hierarchical modeling in ecology—Modeling distribution, abundance and species richness using R and bugs. Volume 2: Dynamic and advanced models*. Academic Press.
- Kreuzer, M. P., & Huntly, N. J. (2003). Habitat-specific demography: Evidence for source-sink population structure in a mammal, the pika. *Oecologia*, 134(3), 343–349.

- MacArthur, R., & Wang, L. (1973). Physiology of thermoregulation in the pika, *Ochotona princeps*. *Canadian Journal of Zoology*, 51(1), 11–16. <https://doi.org/10.1139/z73-002>
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8), 2248–2255.
- MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G., & Franklin, A. B. (2003). Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology*, 84, 2200–2207. <https://doi.org/10.1890/02-3090>
- McClintock, B. T., Bailey, L. L., Pollock, K. H., & Simons, T. R. (2010). Unmodeled observation error induces bias when inferring patterns and dynamics of species occurrence via aural detections. *Ecology*, 91(8), 2446–2454. <https://doi.org/10.1890/09-1287.1>
- Millar, J. S., & Zwickel, F. C. (1972). Characteristics and ecological significance of hay piles of pikas. *Mammalia*, 36(4), 657–667. <https://doi.org/10.1515/mamm.1972.36.4.657>
- Miller, D. A., Nichols, J. D., Gude, J. A., Rich, L. N., & Podrutzny, K. M. (2013). Determining occurrence dynamics when false positives occur: Estimating the range dynamics of wolves from public survey data. *PLoS One*, 8, 65808. <https://doi.org/10.1371/journal.pone.0065808>
- 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>
- Moilanen, A., Smith, A. T., & Hanski, I. (1998). Long-term dynamics in a metapopulation of the American pika. *American Naturalist*, 152, 530–542.
- Nichols, J. D., Bailey, L. L., O'Connell, A. F., Jr., Talancy, N. W., Campbell Grant, E. H., Gilbert, A. T., Annand, E. M., Husband, T. P., & Hines, J. E. (2008). Multi-scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology*, 45(5), 1321–1329. <https://doi.org/10.1111/j.1365-2664.2008.01509.x>
- Nichols, L. B. (2010). Fecal pellets of American pikas (*Ochotona princeps*) provide a crude chronometer for dating patch occupancy. *Western North American Naturalist*, 70, 500–507. <https://doi.org/10.3398/064.070.0410>
- Peacock, M. M., & Smith, A. T. (1997). The effect of habitat fragmentation on dispersal patterns, mating behavior, and genetic variation in a Pika (*Ochotona princeps*) metapopulation. *Oecologia*, 112(4), 524–533.
- Plummer, M. (2003, March). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, No. 125.10, pp. 1–10).
- Plummer, M. (2021). *Rjags: Bayesian graphical models using MCMC*. [R package version 4–12]. <https://CRAN.R-project.org/package=rjags>
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ray, C., Beever, E., & Loarie, S. (2012). 12. Retreat of the American pika: Up the mountain or into the void? In J. F. Brodie, E. S. Post, & D. F. Doak (Eds.), *Wildlife conservation in a changing climate* (pp. 245–270). University of Chicago Press. <https://doi.org/10.7208/9780226074641-012>
- Ray, C., Beever, E. A., & Rodhouse, T. J. (2016). Distribution of a climate-sensitive species at an interior range margin. *Ecosphere*, 7(6), e01379.
- Rodhouse, T. J., Beever, E. A., Garrett, L. K., Irvine, K. M., Jeffress, M. R., Munts, M., & Ray, C. (2010). Distribution of American pikas in a low-elevation lava landscape: Conservation implications from the range periphery. *Journal of Mammalogy*, 91(5), 1287–1299.
- Rodhouse, T. J., Jeffress, M. R., Sherrill, K. R., Mohren, S. R., Nordensten, N. J., Magnuson, M. L., Schwalm, D., Castillo, J. A., Shinderman, M., & Epps, C. W. (2018). Geographical variation in the influence of habitat and climate on site occupancy turnover in American pika (*Ochotona princeps*). *Diversity and Distributions*, 24(11), 1506–1520.
- Royle, J. A. (2004). Modeling abundance index data from anuran calling surveys. *Conservation Biology*, 18(5), 1378–1385.
- Royle, J. A. (2006). Site occupancy models with heterogeneous detection probabilities. *Biometrics*, 62(1), 97–102.
- Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities*. Elsevier.
- Royle, J. A., & Link, W. A. (2006). Generalized site occupancy models allowing for false positive and false negative errors. *Ecology*, 87(4), 835–841.
- Ruiz-Gutierrez, V., Hooten, M. B., & Campbell Grant, E. H. (2016). Uncertainty in biological monitoring: A framework for data collection and analysis to account for multiple sources of sampling bias. *Methods in Ecology and Evolution*, 7(8), 900–909.
- Ruiz-Gutierrez, V., Zipkin, E. F., & Dhondt, A. A. (2010). Occupancy dynamics in a tropical bird community: Unexpectedly high forest use by birds classified as non-forest species. *Journal of Applied Ecology*, 47(3), 621–630.
- Schmidt, B. R., Kery, M., Ursenbacher, S., Hyman, O. J., & Collins, J. P. (2013). Site occupancy models in the analysis of environmental DNA presence/absence surveys: A case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution*, 4(7), 646–653. <https://doi.org/10.1111/2041-210X.12052>
- Schwalm, D., Epps, C. W., Rodhouse, T. J., Monahan, W. B., Castillo, J. A., Ray, C., & Jeffress, M. R. (2016). Habitat availability and gene flow influence diverging local population trajectories under scenarios of climate change: A place-based approach. *Global Change Biology*, 22(4), 1572–1584.
- Shinderman, M. (2015). American pika in a low-elevation lava landscape: Expanding the known distribution of a temperature-sensitive species. *Ecology and Evolution*, 5(17), 3666–3676.
- Simpson, W. G. (2009). American pikas inhabit low-elevation sites outside the species' previously described bioclimatic envelope. *Western North American Naturalist*, 69(2), 243–250.
- Smith, A. T. (1974). The distribution and dispersal of Pikas: Influences of behavior and climate. *Ecology*, 55(6), 1368–1376.
- Smith, A. T., & Weston, M. L. (1990). *Ochotona princeps*. *Mammalian Species*, 1–8. <https://doi.org/10.2307/3504319>
- Wilkening, J. L., & Ray, C. (2016). Characterizing predictors of survival in the American pika (*Ochotona princeps*). *Journal of Mammalogy*, 97(5), 1366–1375.
- Wolf, S., Nowicki, B., & Siegel, K. (2007). *Petition to list the American pika (Ochotona princeps) as threatened under the California endangered species act*. Center for Biological Diversity publication.

## SUPPORTING INFORMATION

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

**How to cite this article:** Goldman, M. R., Shinderman, M., Jeffress, M. R., Rodhouse, T. J., & Shoemaker, K. T. (2023). Integrating multiple sign types to improve occupancy estimation for inconspicuous species. *Ecology and Evolution*, 13, e10019. <https://doi.org/10.1002/ece3.10019>