**Title**: The relationship between an avian seed disperser and seed availability in the Southwest USA is lagged and dependent on climate and habitat context

# 2 Materials and Methods

### 2.1 Pinyon jay biology and study site

Pinyon jays are a widespread species that occurs across the western US (California, Nevada, Utah, Arizona, Colorado, New Mexico). Pinyon jays are generalist seed- and insect-eaters and are part of an important mutualistic relationship with piñon pines (CITE Balda 2020, Ligon 1978). Birds cache large amounts of piñon pine seeds when they are available and by burying seeds they increase the potential for seedlings to germinate (CITE: Floyd 1982). Pinyon jays often flock and nest in large numbers and are highly social (Bednekoff and Balda 1996, Stotz and Balda 1995, Marzluff and Balda 1988, Marzluff 1988). Pinyon jays are in decline across much of their range for a variety of potential reasons, including changes in habitat quality and food availability (CITE Boone, VanLannen). Piñon pine is a masting species [Andreas can you fill in some cool ecology mirroring above for piñon?] Concurrently, piñon pines have seen shifts in their range, age and size structure, and reproduction [Andreas - provide some details here about decline/change for pinyon pine, matching above for the bird side].

Although pinyon jays occur across the western US, in this study, we focused on the intermountain and southwest regions of the US because of available information on cone production (as a proxy for seed production) for single-needle pinyon (e.g., data from Utah, Colorado, New Mexico, and Arizona; Wion et al. 2025). [Any other justifications for including only data from this region WRT climate/habitat? Please add tree guys]. Pinyon jays occur in piñon ecosystems of other subspecies of piñon pine (CITE), but no masting information is available at a broad scale for these species (Andreas - TRUE?)

### 2.2 Covariate data

Because we were interested in relationships between pinyon jays and piñon pine seed production, we built a model that included cone production (as a proxy for piñon seed production; Wion et al. 2025) and other mediating factors. We used the cone production dataset both as a direct measure of piñon seed availability and a proxy for the availability of other seeds eaten by pinyon jay (e.g., acorns, juniper berries) that respond to the same climate signals as piñon pine, but with different temporal lags (1-year lag for acorns and juniper berries, 2-year lag for piñon pine; Parmenter paper). The covariates we included as mediating factors included site-level yearly climate and pinyon basal area data as well as site level average monsoonal moisture (precipitation occurring in July-September). We compiled all covariates at a 4x4 km raster grid scale since this is a scale at which these data are available (e.g., PRISM). Cone production data came from published models estimating yearly cone production across the region [Andreas, please fill in here]. [Andreas, can you provide citations linking cone to seed availability?]. We derived yearly basal area by [Kyle please fill in here]. In addition to these covariates, we also gathered gridded climate data relevant to bird biology (physiology and relative food availability; Ligon 1978, Stotz and Balda, McDermott et al., Skagen et al.) from PRISM, including monthly maximum temperature, monthly cumulative precipitation, and the percent of the 30-year monthly normal precipitation falling in the months of July-September to quantify the proportion of moisture experienced in a grid during the monsoon season (Wion et al. 2025 OTHER “monsoon” citations?).

### 2.3 eBird filtering and subsampling

To estimate pinyon jay abundance in relation to recent (2010-2022) seed abundance, we used data from eBird [cite] from the four states of interest (Arizona, Colorado, New Mexico, and Utah) and further filtered pinyon jay observations to areas with pinyon pine vegetation cover (CITE where Kyle got his pinyon pine layers from). We chose to use eBird data because it has broad spatial and temporal coverage within the region of interest and equivalent spatial scale to compare bird abundance and our covariates (4 x 4 km grid cells). This covariate scale is too fine for many long-term, broad-scale standardized monitoring protocols, such as Breeding Bird Survey (25-mile transects; CITE BBS).

eBird is participatory science data, meaning that members from the public collect bird observation data and contribute to the eBird database. Thus, these data often require subsetting based on data completeness. We followed filtering methods from the eBird database curators and similar studies (Strimas-Mackey et al. 2023, Robinson et al. 2018). Specifically, we selected data that represented only complete checklists (e.g., observers recorded all birds of all species observed, not just individuals of species of interest). We selected checklists from the nesting season (late February to early May; CITE Wiggens 2005) and included only observations of “stationary” and “traveling” protocols. We considered only the nesting season because this is a time of year in which the mechanisms for local bird abundance are easy to evaluate because of communal nesting behavior and because nesting often occurs near areas where birds have cached seeds from the previous year (Ligon 1978, Wiggens 2005). We only included checklists that were collected in under five hours, 0-5 km in distance, and with ten or fewer observers.

eBird data are unstructured data not collected under standardized protocols with systematic sampling designs for sampling location, effort, and sampling methods, so we also performed spatial and temporal sub-sampling (Robinson et al. 2018). To do this, we divided the landscape into 4x4 km grid cells to match the spatial scale of the covariates. We spatially sub-sampled eBird observations to include a set of checklists in each grid cell that represented relatively equal sampling effort across time and space. Given that we observed the fewest number of checklists per year in 2010 (average 4.2 checklists per grid cell), we subsampled to this level ( 4 checklists per cell). After filtering to this level, we still had fewer checklists in early years due to fewer cells having checklists in them. Thus, we further reduced temporal bias by subsampling checklists again so that we had equal sample sizes across years based on the year with the fewest checklists (*n* = 1715 checklists in 2010). We performed this temporal sampling randomly (e.g., not by 4x4 km grid cell). We also removed all checklists within 10 km of the exterior edges of the four states for which we had data (Arizona, Colorado, New Mexico, and Utah) since we had masked some other covariate layers to these states and observations near edges may have misrepresented values for pinyon basal area and cone production. Of 833,834 checklists for the region in the nesting seasons between 2010-2022, we ran our model on 22,295 checklists (average: 1.4 (SE: 0.01) checklists per grid cell in a specific year). We downloaded eBird data in May (Colorado and New Mexico, CITE) and June (Arizona and Utah, CITE) 2024.

### 2.4 Modeling relationships between cones and birds with eBird

To explore relationships between pinyon jay abundance and environmental covariates, we used a two-part model with an observation process for eBird and a biological process model that estimates latent “abundance” values (CITE N-mixture models; Royle 2004). The observation process models include covariates that describe variation in detection probabilities (e.g., survey conditions). The biological process includes a likelihood that describes the biological drivers of latent abundance. These biological covariates include a stochastic antecedent modeling structure that allows seed abundance to have both lead and lagged effects (e.g., bird abundance can respond before or after cone and seed maturation) and other yearly climate variables to have lagged effects (Ogle et al. 2015).

Replicate surveys within a closed time and location are a requirement of these types of models that incorporate an observation process (e.g., occupancy and N-mixture models CITE). To use unstructured or semi-structured eBird data in this way (e.g., as a space-for-time substitution), checklists must be collected at locations that describe the range of potential covariate values (e.g., are not biased toward certain values of covariates of interest, Hochachka et al. 2023). We verified that checklists within our filtered dataset (Section 2.3) met this assumption by comparing values for all covariates at locations with checklists to all gridded data available for those covariates (described more in Section 2.2). Thus, we verified that we could use multiple checklists within a location and year as replicate “surveys” with which to estimate detection probability.

### 2.4.1 Observation process for eBird

In our model, observed count data from eBird checklists were used to estimate a “true” latent abundance, , which is dependent on space and time (more in section NAME). eBird data were provided in an observation process model that incorporates covariates that impact bird detection. Observed count data were binomially distributed based on a detection probability for for birds in a year and the “true” latent abundance for the space and time in which those checklists occured, , being the total possible “trials”.

eBird checklists are recorded with a geographical point location and also often include observers walking some distance (< 5km in our dataset after filtering). Thus, there is spatial uncertainty arising from two sources: 1) uncertainty in the location in which the survey was conducted (e.g., was the location recorded at the beginning, middle, or end of a survey and at what level of precision) and 2) uncertainty in the location along a survey at which pinyon jays were observed (e.g., were birds observed at the beginning, middle, or end of a walking survey). To account for this spatial uncertainty, we created buffers around each eBird checklist based on the sampling distance for that checklist (1/2 of distance traveled for “traveling” checklists; 0.5 km for “stationary” checklists). Then, we created a merged polygon (“blob”) of all the checklists within a 4x4 grid cell in a given year by merging the buffers of all checklists within a grid cell. Thus, the “blob” and not the underlying 4x4 km grid cell, is the spatial unit at which we estimated abundance, *N*.

eBird observed count data for checklist, , in blob, , and year, , were binomially distributed based on a detection probability, *p*, and the latent true abundance of birds in the checklist:

The detection probability for eBird data included a regression with the detection covariates of start time, duration of survey, speed, and the number of observers for each checklist. See above (\_\_\_) for how we filtered these variables initially based on best practices so that checklists represented similar effort. These covariates all depended on checklist, blob, and year. Thus, the detection component of the model has a likelihood of the form:

where is the underlying detection probability (intercept) and each is the slope of the relationship between detection probability (on the logit scale) and a set of *J* covariates (j = 1, 2, … 5).

### 2.4.2 Biological process of latent pinyon jay abundance with covariates with lead and lagged effects

Latent true abundance of jays in a checklist in a year, was Poisson distributed around an expected rate parameter, , for the blob in which the checklist was located multiplied by the area of the checklist, as follows:

Furthermore, was dependent on a regression of environmental covariates:

In this formulation, covariates include only concurrent effects. These variables in the model included monsoonality and pinyon pine basal area (yearly). Conversely, values for each comprised a stochastic antecedent structure (Ogle et al. 2015) dependent on the covariate, where each was the weighted average value of a set of seasonal or yearly values for that covariate across a defined set of seasons or years. The weights for each covariate across all seasons or years, *l*, , had a Dirichlet prior and sum to 1. The weight, or importance weight, for each season or year indicated how important that value was to the overall effect (larger values of the weight mean more importance of that season or year to that covariate effect). The covariates for which we calculated antecedent weights were seed (cone) production, maximum temperature, and precipitation. Seed (cone) production was considered on a yearly scale (cones are only produced up to one time in a year) and included both lead and lagged effects (e.g., jays “predict” seed years or “respond” to good seed years), and included one prior year and the current year (both “predictive” time periods), and three years after the current year (“lagged” time periods). Lead (“predictive”) years represented years in which “quicker” resources are available (e.g., acorns, juniper berries) and/or birds observe cones developing on trees (Parmenter, “Predict” paper for jays). Lag years represented both the year in which birds cached seeds and years following in which individuals that hatched during the caching year nesting season enter the breeding population (~1.5-2 years CITE; Figure 1). Both temperature and precipitation were at a seasonal scale and only consider lagged effects. The seasons were based on important seasonal biology for pinyon jays, including: breeding (February-April), feeding dependent young (May-June), summer (July), fall and winter foraging, potentially with irruption (August - January) (Wiggins 2005). For seasonal covariates, we considered 12 seasons into the past. We also considered interactions between seed (cone) abundance and all other covariates (pinyon basal area, monsoonality, maximum temperature, and precipitation) to determine how the relationship between cones and birds is mediated by other covariates related to piñon seed reliability and alternative food resources (e.g., juniper berries, acorns, insects).

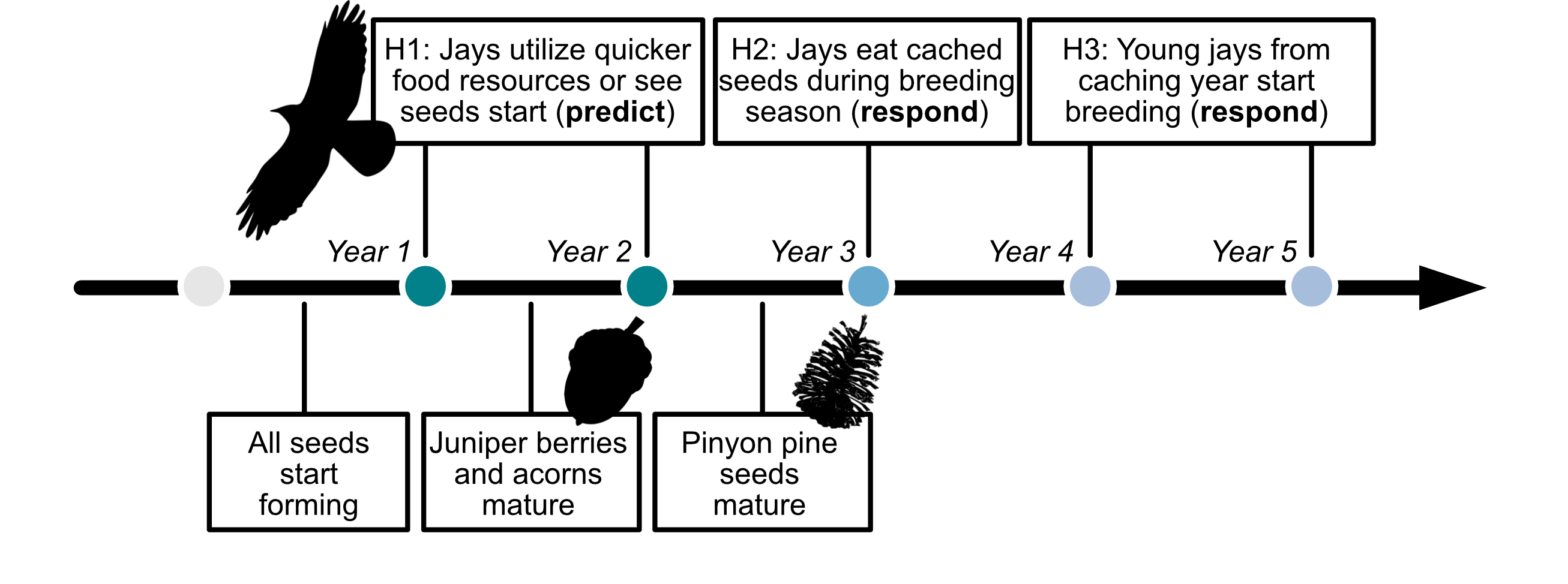


Figure 1: A timeline of hypothesized pinyon jay responses (top) and the progression of seed and cone development (bottom) following a favorable climate year for cone development. We outline three hypotheses driving the temporal relationships between pinyon jay and seed availability, including H1: jays ‘predict’ seed availability through utilization of other foods (e.g., acorns, juniper berries) or because they see cones developing in an area; H2: jays ‘respond’ to seed availability in the nesting season after they cache seeds and H3: jays ‘respond’ to seed availability with a delay as the young from the caching year enter the breeding population 1.5-2 years later.

Because blobs sometimes covered more than one 4x4 km grid cell, we determined the percent of a blob covering each grid cell and then calculated weighted average covariate values based on the covariate values of all overlapped cells times the percent of the blob overlapping that cell. On average, blobs covered 2.13 [SE = 0.01] 4x 4 km grid cells, with a minimum of one grid and a maximum of nine for a single blob.

### 2.5 Model implementation, convergence, and diagnostics

We implemented models in the Bayesian software JAGS (CITE) using R (CITE) and the jagsUI wrapper package (CITE). We prepared data using the here (CITE), tidyverse (CITE), sf (CITE), terra (CITE), readxl (CITE), exactextractr (CITE), spatialEco (CITE), nngeo (CITE), sp (CITE), auk (CITE), lubridate (CITE), prism (CITE), and data.table (CITE) packages.

We initially ran models to generate 4000 MCMC samples with which to determine how many iterations would be needed to reach convergence (CITE). We then re-ran models with more iterations and initial values sufficient to reach convergence based on the raftery.diag() function in the coda package (CITE). We assessed model convergence by evaluating trace, history, and autocorrelation plots generated with the mcmcplots package (). We quantified model convergence by evaluating whether Gelman-Ruman statistics, , for all root nodes in the model were < 1.1 (CITE). We assessed model goodness-of-fit by comparing observed count data from eBird to predicted count data based on the model mean value. We verified the lack of spatial autocorrelation in observed data by examining the residuals using [give methods here SI Figure X].

### 2.6 Out of sample model validation

To evaluate model fit, we performed the same set of steps for spatially and temporally stratifying available checklists from eBird (section \_\_ above), ensuring that we selected checklists that were not used in model fitting. We had 20,380 checklists that represented “out of sample” data. We predicted out of sample model fit using ~1000 posterior samples for covariate effects and intercepts from the test dataset and model. We used RMSE to compare the test and out-of-sample datasets for model fit.

# Results

[How often are jays observed in eBird data? How abundant are they in a given checklist?]

Of the total number of available eBird checklists in the years of interest (2010-2022), pinyon jays were observed in 5,658 of them (0.7%). In our in-sample dataset (n = 22,295 checklists), pinyon jays were observed in 600 (2.7%). The maximum number of birds observed in one survey in this dataset was 350 (mean: 0.42; SE: 0.04). All parameters in the model converged with an 1.1 (SI Figure 1).

[What is the magnitude and time scales of the relationship between birds and piñon pine seed availability?]

Pinyon jay abundance was positively correlated with cone availability ( = 0.95 [0.92, 0.98]). The importance weights of the years considered (two years before cones; three years after cones) indicate the greatest weight the year after cones ( = 0.64 [0.60, 0.67]) and some weight to two and three years after cones ( = 0.13 [0.07, 0.18] and 0.21 [0.17, 0.25], for two and three years after cones, respectively), though negligible weight for the years before cones ( = 0.01 [0, 0.03] and 0.01 [0, 0.06] for two and one years before cones, respectively).

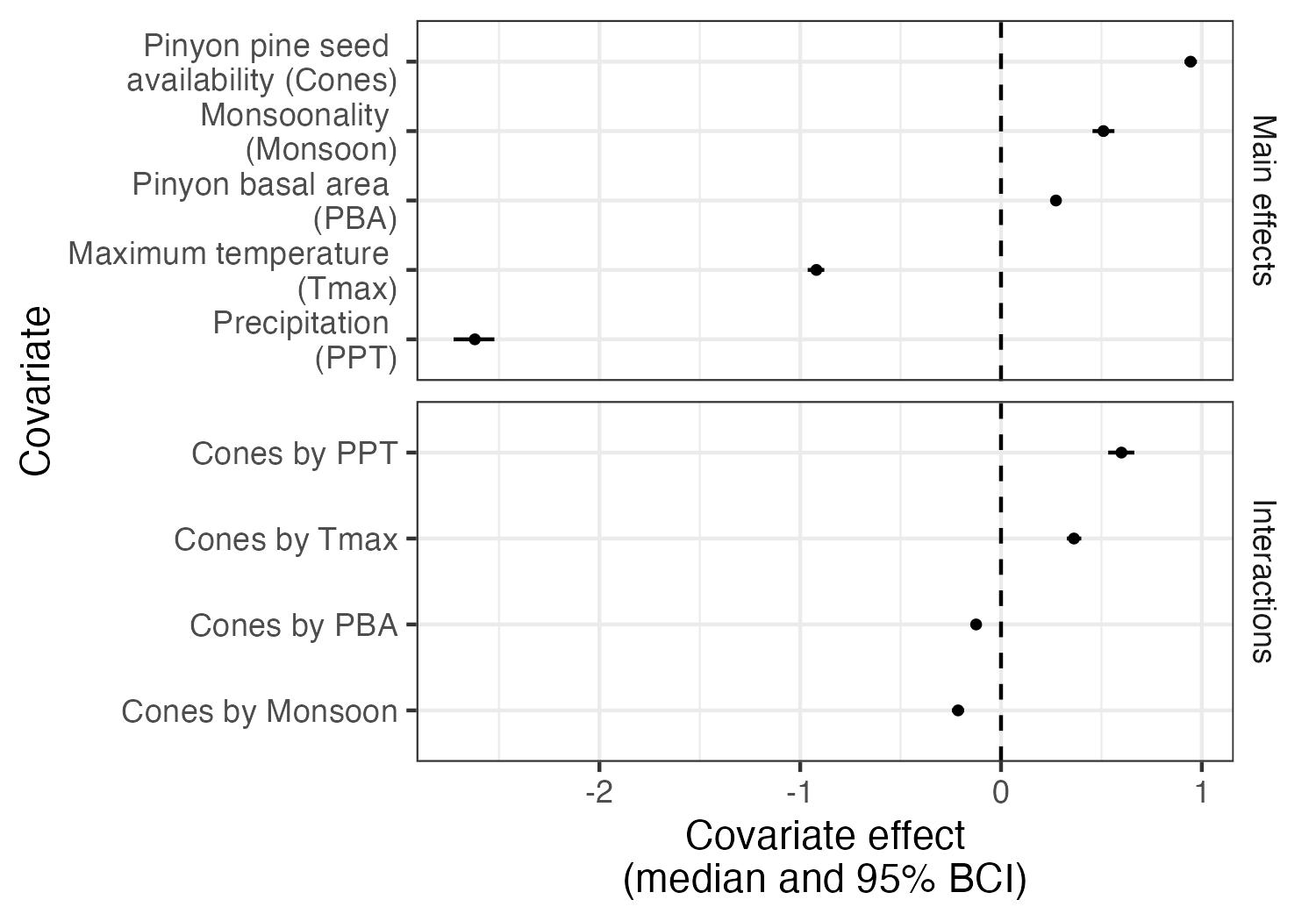


Figure 2: Covariate effects for all main covariates and their interactions with cones. All covariates were scaled and so their relative magnitudes of effect can be compared. All values represent posterior median and 95% Bayesian Credible Intervals and any covariate ranges that cross the dashed line at zero have no clear effect on pinyon jay abundance.

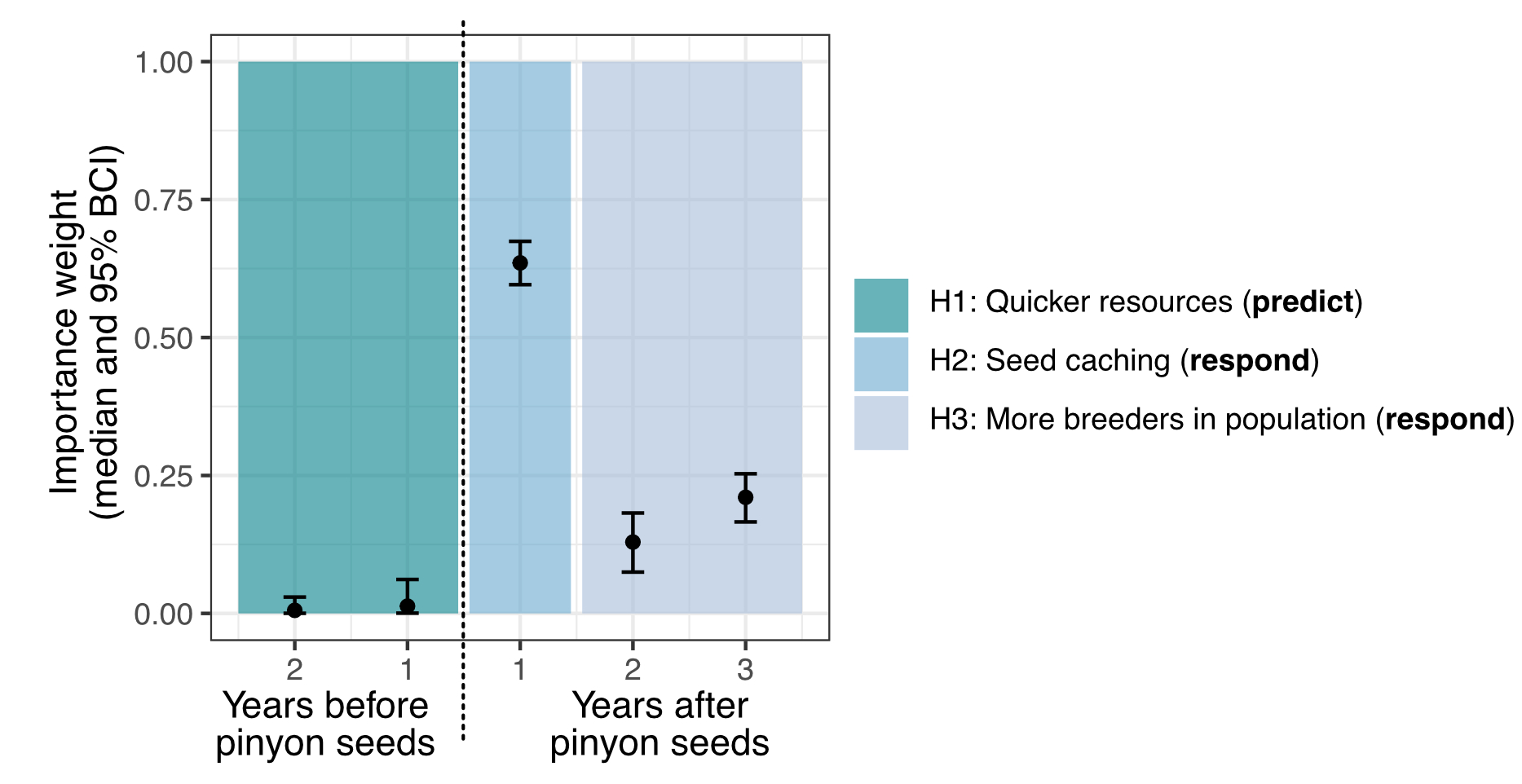


Figure 3: Importance weights for the years considered for the relationship between pinyon jays and pinyon cone (seed) availability. Colored boxes indicate the three ecological hypotheses about the time period(s) of the relationship between jays and cones. The dashed line indicates the distinction between the years before cones (left) and years after cones (right).

[How is this relationship mediated by habitat and climate variables that vary by site and/or year?]

All other covariates and interactions in the model had non-zero effects on pinyon jayabundance. Those main effects that were negatively related to pinyon jay abundance were: precipitation ( = -2.62 [-2.72, -2.53]) and maximum temperature ( = -0.92 [-0.96, -0.88]). Main effects that were positively related to pinyon jay abundance were: Monsoonality ( = 0.51 [0.46, 0.56]) and pinyon basal area ( = 0.27 [0.25, 0.29]). Positive interactions included the interactions between cones and: precipitation ( = 0.60 [0.56, 0.66]) and maximum temperature ( = 0.37 [0.33, 0.40]). Negative interactions included the interactions between cones and: pinyon basal area ( = -0.12 [-0.14, -0.11]) and monsoonality ( = -0.21 [-0.25, -0.19]). All covariates were scaled in the model, so the relative magnitudes of these effects can be compared.

Important seasons for precipitation based on importance weights from the SAM model included, in order of importance weights: two years ago during fledging, three years ago during breeding, two years ago during breeding, and one year ago during fledging ( = 0.37 [0.35, 0.39]; 0.20 [0.18, 0.21]; 0.16 [0.14, 0.18]; 0.12 [0.10, 0.14], for each of those seasons, respectively). All other weights were not important (SI Figure 2). Important seasons for maximum temperature based on the importance weights from the SAM model included: the breeding and fledging seasons two years ago ( = 0.60 [0.50, 0.69] and 0.34 [0.25, 0.44], for the breeding and fledging seasons, respectively). All other weights were not important (SI Figure 2).

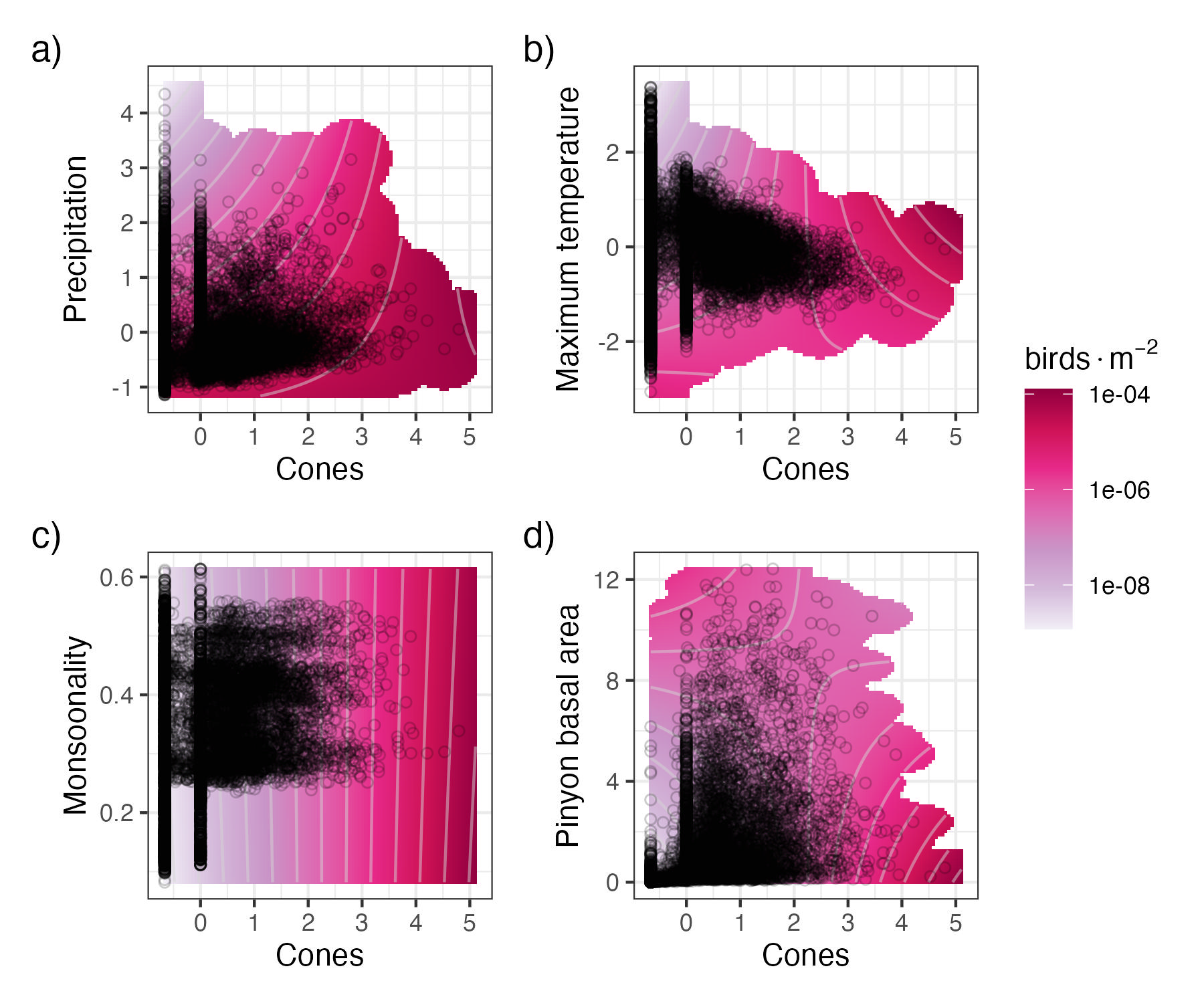
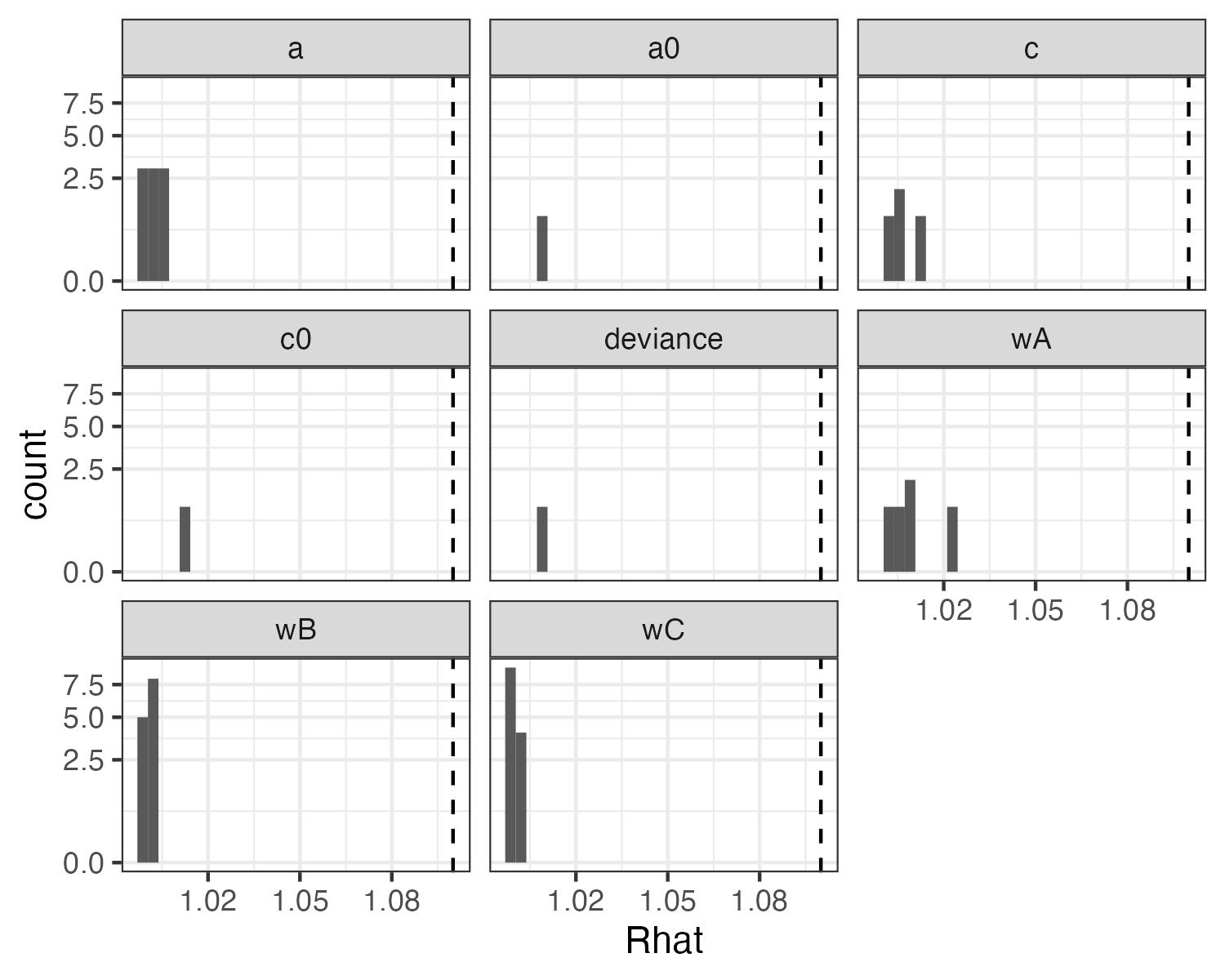


Figure 4: Interaction plots for the four covariates with significant interactive effects with cone (seed) availability, including: a) precipitation, b) maximum temperature, c) monsoonality, and d) pinyon basal area. Darker pink/red color corresponds to a greater number of birds in that combination of covariates. Black circles represent the distribution of the observed data. The predicted number of birds, lambda, is presented on the log scale to better highlight differences.

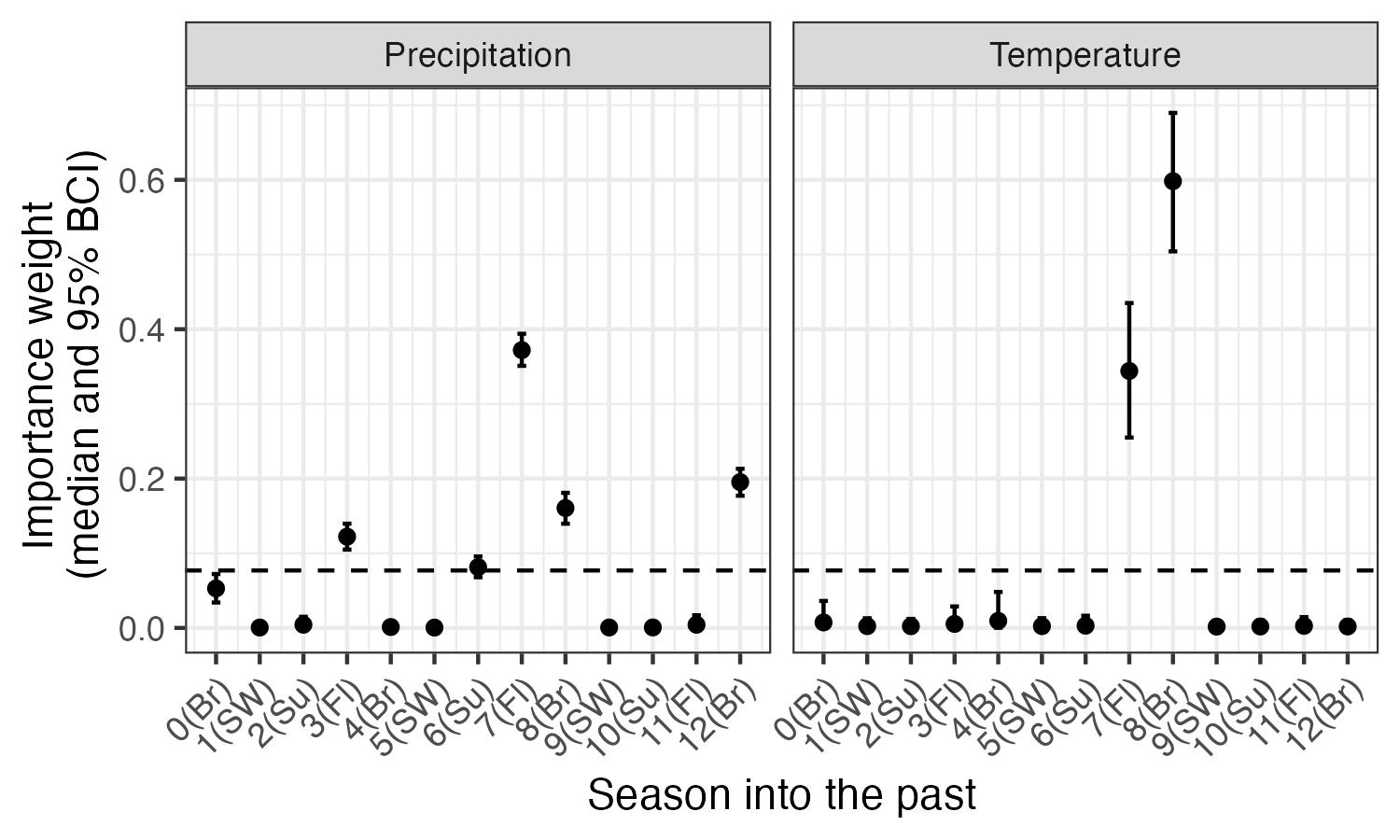
[Model fit and oos evaluation]

As a goodness-of-fit metric, the relationship between our observed data and replicated data produced by the model had a mean = 0.74 (SE: 0.0003). In general, the model under-estimates higher counts (SI Figure 3). The RMSE for our in-sample (“test”) dataset had an average value of 5.10 (SE: 0.0002) and the out-of-sample dataset had a value of 3.25 (SE: 0.001; SI Figure 4). Thus, our model predicts slightly better to our in-sample (“test”) data than out-of-sample data, but neither are large on the scale of the count data (counts per checklist ranged from 0 - 350), suggesting good model fit and predictive capacity. Because of the spatial nature of the data in this project, we also verified that there was no spatial autocorrelation in residuals (SI Figure 5).

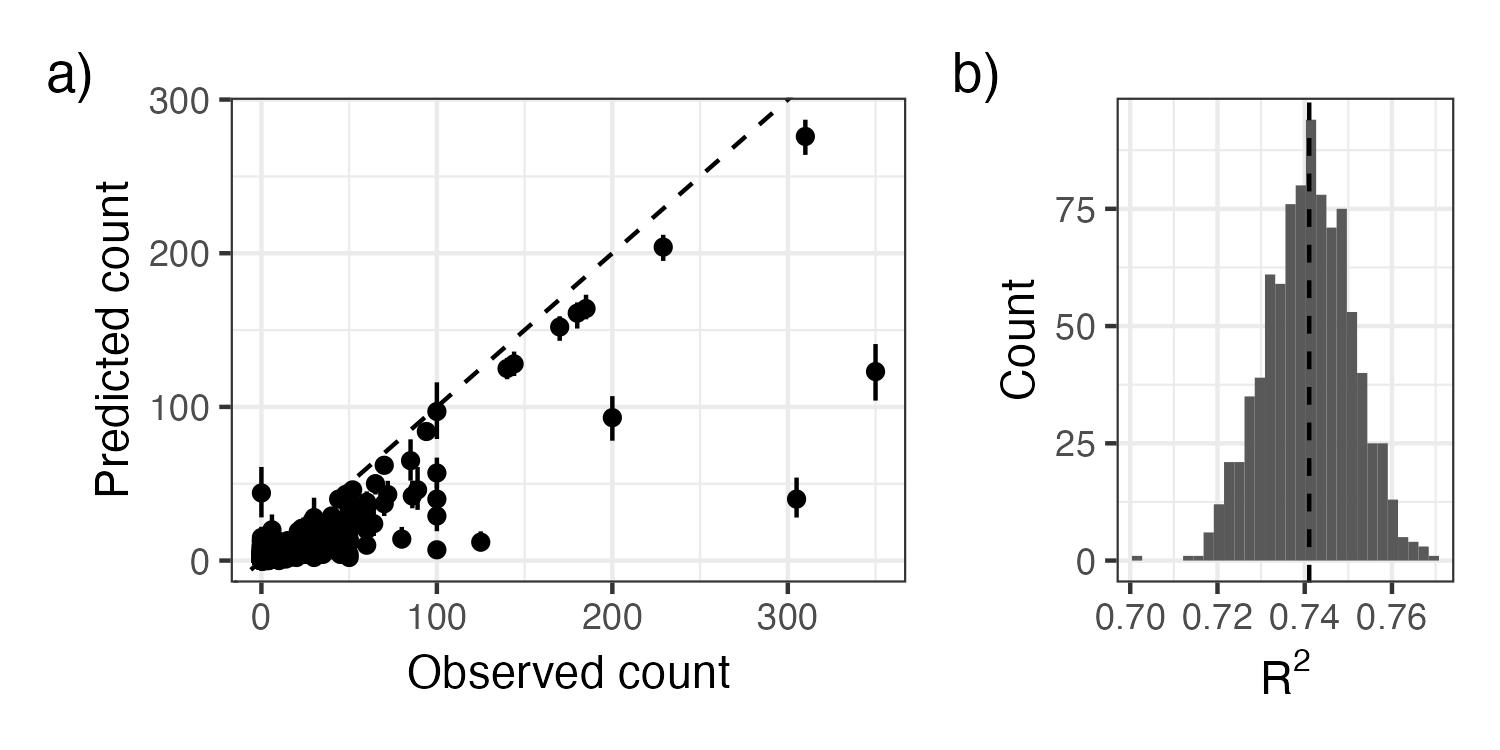
# SI Figures



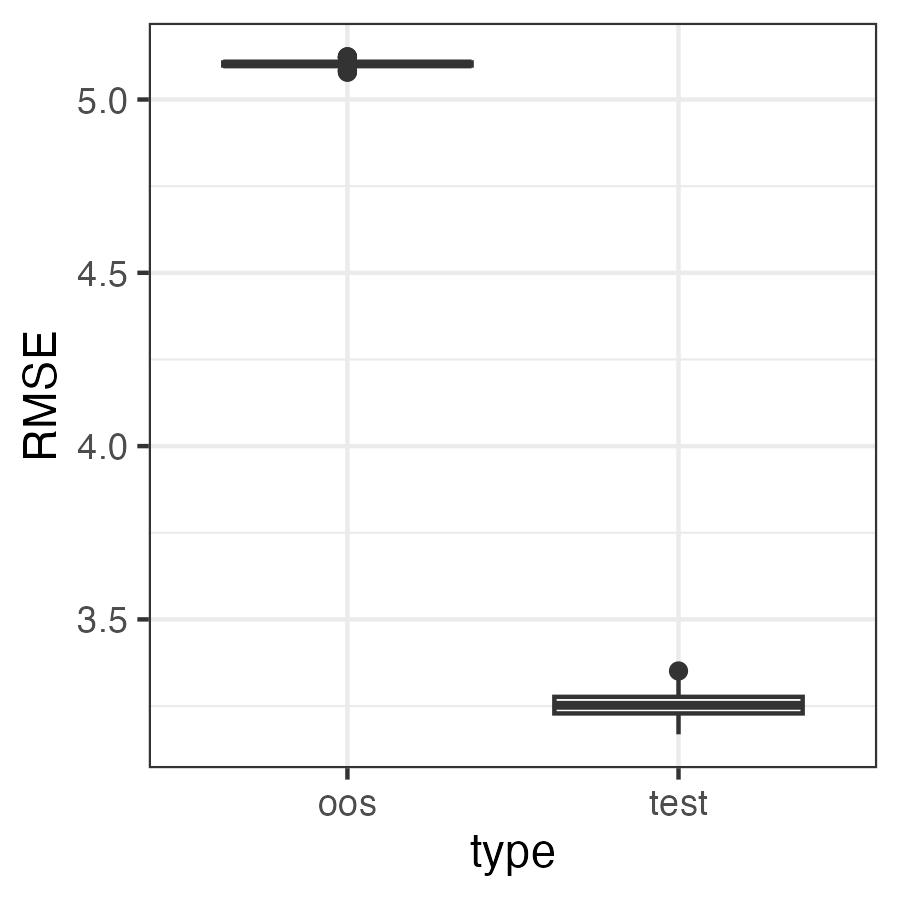
SI Figure 1: Convergence statistics for biological (a and a0) and detection (c and c0) covariate effects, deviance, and the cone and climate weights (wA, wB, and wC). All covariate effects and parameters converged with Rhat values <= 1.1 (dashed line).



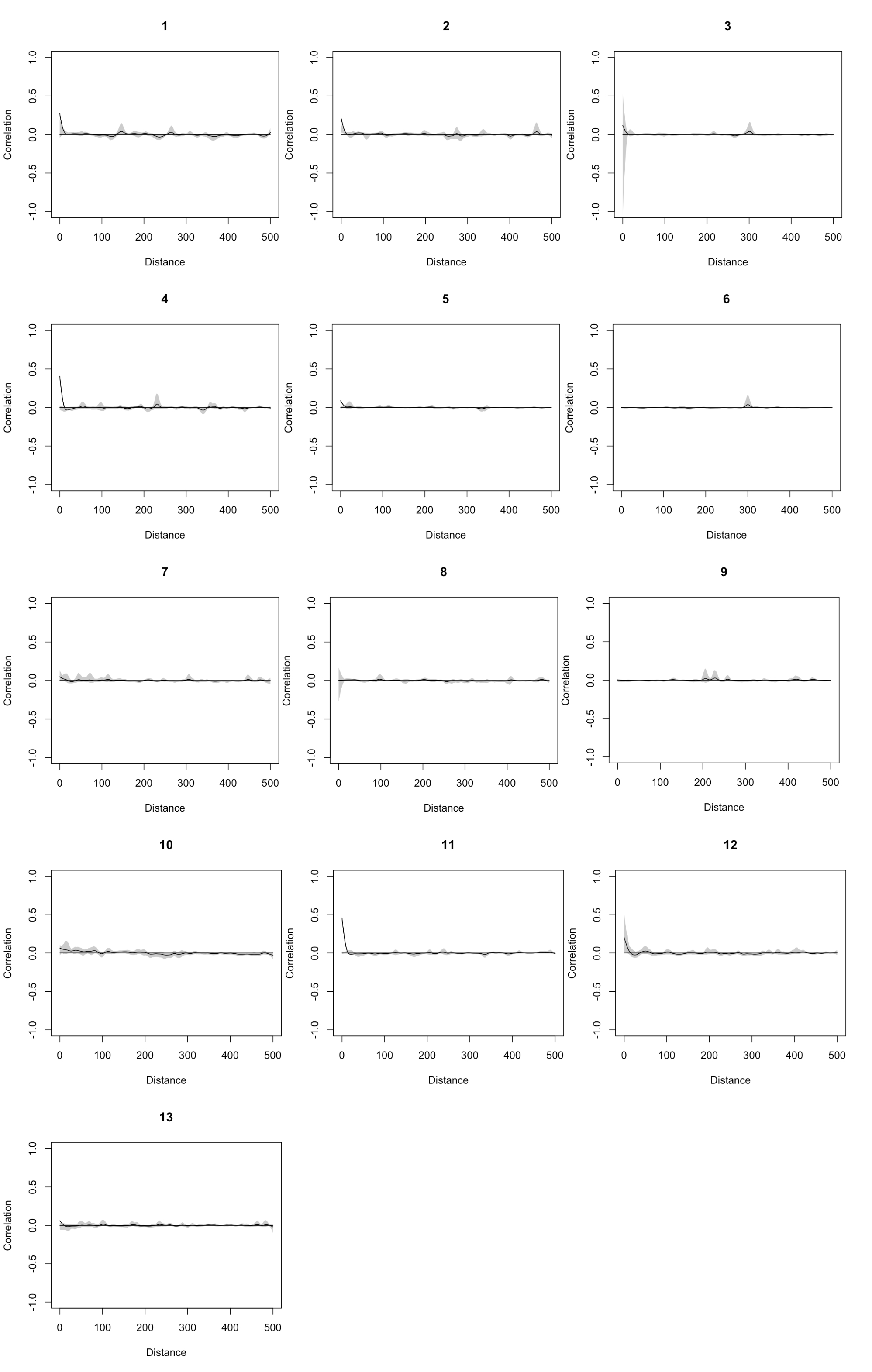
SI Figure 2: Seasonal weights for climate (precipitation and maximum temperature) in the model. The dashed line indicates the ‘null’ weight in which all time periods would have equal weights. Any weights that extend clearly above this line have significant importance to the climate variable. Seasons are based on bird biology and include Br: breeding (February - May), Fl: fledging (May - June), Su: summer localized foraging (July), SW: summer and winter dispersed foraging (‘irruption’; August - January).



SI Figure 3: As an evaluation of goodness-of-fit, we examined a) the relationship between observed pinyon jay counts and counts predicted from the model (median and 95% Bayesian Credible Interval) and b) the R2 of this relationship from ~1000 posterior samples. In a), the dashed line represents the 1:1 line where predictions exactly match. In b) the dashed vertical line represents the mean R2.



SI Figure 4: RMSE for the test and out of sample (oos) datasets. Values for RMSE are on the scale of the data (counts of pinyon jays). They are relatively small compared to the range of possible counts in a checklist (0 - 350 birds in our test dataset).



SI Figure 5: Evaluation of spatial autocorrelation (Pearson correlation on the y-axis) for residuals in each year (1 = 2010; 13 = 2022). There was no consistent evidence of spatial autocorrelation of residuals.