**Title**: Linking pinyon jay population abundance to pinyon cone production and climate in the SW United States

**Author list**: Ana Miller-ter Kuile1,2, Andreas Wion3, Kyle Rodman4, Jamie S. Sanderlin2

**Author affiliations**:

1 Northern Arizona University School of Informatics, Computing, and Cyber Systems, Flagstaff, Arizona, USA

2 USDA Forest Service Rocky Mountain Research Station, Flagstaff, Arizona, USA

3 USGS Santa Fe, New Mexico, USA

4 Ecological Restoration Institute, Northern Arizona University, Flagstaff, Arizona, USA

**Keywords**:

# Data availability

# Abstract

# 1. Introduction

In the SW, there is a species of jay that has been declining over the last several decades. This jay is co-evolved with pinyon pine trees, whose seeds it caches across the landscape to eat later or to forget later and help the tree disperse to new areas. Pinyon pines have also seen dramatic shifts in range (expansions, die offs) and reproduction.

Given the complex co-evolution of the interaction between the pinyon pine and the pinyon jay, it is hard to tease apart the cause and effect of these patterns or how changes in resource availability alter jay abundances. This has not been evaluated to date. It is likely that, given the broad geographical range of both species, that a different combination of forces may be playing out across the species’ range, and that these effects may not be concurrent. Pinyon pine cone production occurs ~2 years following favorable climate years. Other food resources for jays in these systems may respond more quickly to the same climate conditions (e.g., insects, junipers, and gambel oak). Thus, it may be that jay populations increase \*prior\* to good cone years for pinyon pines as they are either signaled to “get ready” by the same climate conditions as trees or because their populations get a boost prior to a good cone year from other food resources. The result of jays “predicting” good cone years is that more jays will be in the landscape to distribute seeds to more new places. In addition or conversely, jays could also respond with an immediate or lagged effect to cone production. Teasing these patterns out and how other climatic conditions mediate these effects could be imperative for understanding the link between patterns in resources and birds and help in conservation of the species.

On a more methodological note, further muddying the waters is the fact that while we have several broad-scale monitoring programs that capture pinyon jays, they all have limitations both spatially and temporally. Ideally, we would like to combine them into one model that helps us predict jay-pine relationships.

In this study, we pulled two data sources for pinyon jays (BBS and eBIRD) into the same statistical model and combining them into a joint likelihood – where both datasets contribute to estimates of population numbers. (this method is called “data integration”). We then used SAM models to explore the timescales at which jays either predict or respond to pinyon pine cone production in the context of a suite of other broad-scale climatic mediates (temperature, precip, monsoonality).

# 2. Methods

### 2.1 Study site and pinyon jay biology

Pinyon jays are a widely-spread species that occurs across the western US (California, Nevada, Utah, Arizona, Colorado, New Mexico). The core of their range is in the Great Basin region of Nevada. Pinyon jays are generalist seed- and insect-eaters, but have co-evolved with pinyon pine. They cache [give some info on caching behavior and co-evolution here]. Pinyon jays often occur in flocks. According to several studies [cite here], pinyon jays have been in decline across their range, likely for a variety of reasons [provide some possibilities here]. Concurrently, pinyon pines have seen shifts in their range and reproduction [provide some details here].

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 (e.g., data from Utah, Colorado, New Mexico, and Arizona; CITE ANDREAS HERE). [Any other justifications for including only data from this region WRT climate/habitat? Please add tree guys].

### 2.2 Bird dataset descriptions

To estimate pinyon jay abundance, we used data from both Breeding Bird Survey (BBS) [cite] and eBIRD [cite] in combination in this study since eBIRD has broader spatial coverage with the environmental covariates we were interested in linking to pinyon jay abundances. For both datasets, we filtered pinyon jay observations to areas with pinyon pine vegetation cover (where did Andreas get the data to mask cone production? CITE).

BBS surveys have occurred since 1966 with the goal of determining long-term trends in bird populations in North America. BBS data are collected each year in June at a set of roadside sampling transects. Each transect is roughly 24.5 miles long with a set of fifty stops roughly 0.5 miles apart. At each stop, observers conduct a 3-minute point count for all birds seen and heard within a 0.25 mile radius. All surveys occur within five hours of local sunrise (CITE). We selected BBS data for pinyon jays from four states (Colorado, Utah, New Mexico, and Arizona) and then masked observations to include only those with starting points within pinyon pine habitat. We also only include data from the first 10 stops (~5 miles; ~8km) in our models because these represent stops within a spatial resolution that is fine enough to mesh with 4x4 km resolution covariate data (described below SECTION ##).

eBIRD is community science data, meaning that non-scientist birders can collect eBIRD data and provide it to the eBIRD database. We followed filtering methods from the eBIRD database curators (Cornell Lab CITE/LINK HERE). 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 a similar timeframe as BBS data (June and July) and included only observations of “stationary” and “traveling” protocols. We only included checklists that were collected in under five hours, 0-5 km in distance, and with ten or fewer observers.

eBIRD data is unstructured data without a regular sampling design, so we had to spatially sub-sample eBIRD data to limit spatial bias. To do this, we implemented a modified version of a previously-used method that links checklists with and without observations of a species of interest within a certain spatial scale as repeat measurements at that spatial scale (e.g., raster grid cell) (Schindler et al., 2022). We adapted the method in that study; however, to not explicitly select checklists within specific raster grid cells. Rather, we linked observed-not observed checklists based on spatial proximity (1km distance or less). We then spatially subset these pairs along with all checklists (observed and not observed) such that within a given year, all checklists (or checklist pairs) were at least 4km from another checklist (or pair).

### 2.3 Exploring relationships between cone production and jay abundances using BBS from 1960-2023

[Kyle please fill this part in]

### 2.4 Examining recent (2010-2023) relationships between pinyon jay abundance and cone production

To explore relationships between pinyon jay abundance and environmental covariates, especially cone production, we implemented a process of data integration (Miller et al., 2019) using a joint-likelihood approach (Schindler et al., 2022) combining recent (2010-2023) data from both BBS and eBIRD. We chose to do only recent relationships since this is the period in which there are greater amounts of data with which to explore environmental relationships.

For this modeling approach, we created a two-part model with an observation process separately for each of the two bird datasets (BBS and eBIRD) and a biological process model that combines information from both bird dataset observation processes with equal weights to estimate latent “abundance” values. The observation process models include covariates that describe variation in detection probabilities (e.g., observer experience and/or survey conditions). The biological process includes a likelihood that describes the biological drivers of latent abundance. These covariates include a stochastic antecedent modeling structure that allows cone abundance to have both lead and lagged effects (e.g., birds can respond before, concurrent, or after high cone years) and other climate variables to have lagged effects (Ogle et al., 2015). We estimated latent abundance at a 4x4 km scale because this is the scale of the spatial covariates. We describe how we account for spatial uncertainty in the bird observation datasets (i.e. “downscale”) in section \_\_\_ below.

### 2.4.1 Covariate datasets

[explanations of covariate datasets]

### 2.4.2 Observation process for BBS and eBIRD

(ALSO include methods on accounting for spatial uncertainty in bird observation datasets)

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

[explanations of the biological process model and of SAM component]

| Covariate or interaction | Explanation |
| --- | --- |
| Pinyon pine cone abundance | Modeled abundance of cone availability (can have lead and lagged effects) |
| Maximum temperature | A proxy for physiological limits and/or insect resource availability (can have lagged effects) |
| Cumulative precipitation | A proxy for insect resource availability (can have lagged effects) |
| Monsoonality | important in conjunction with cone interaction (see below for interactions) |
| Pinyon basal area | important in conjunction with cone interaction (see below for interactions) |
| Cone abundance x Temperature | Is the importance of cone abundance for jays mediated by temperature? |
| Cone abundance x Precipiation | Is the importance of cone abundance for jays mediated by precipitation (e.g., more/less availability of insect prey) |
| Cone abundance x Monsoonality | Is the importance of cone abundance for jays different across a range of monsoonal conditions that alter cone reliability? |
| Cone abundance x Pinyon basal area | A combined metric of actual seed availability - is jay abundance linked to availability of cone abundance on the landscape |

### 2.4.4 Model implementation, convergence, and diagnostics

# Results

# Discussion

# Tables and Figures

# References

Miller, D. A. W., Pacifici, K., Sanderlin, J. S., & Reich, B. J. (2019). The recent past and promising future for data integration methods to estimate species’ distributions. *Methods in Ecology and Evolution*, *10*(1), 22–37. <https://doi.org/10.1111/2041-210X.13110>

Ogle, K., Barber, J. J., Barron-Gafford, G. A., Bentley, L. P., Young, J. M., Huxman, T. E., Loik, M. E., & Tissue, D. T. (2015). Quantifying ecological memory in plant and ecosystem processes. *Ecology Letters*, *18*(3), 221–235. <https://doi.org/10.1111/ele.12399>

Schindler, A. R., Cunningham, S. A., Schafer, T. L. J., Sinnott, E. A., Clements, S. J., DiDonato, F. M., Mosloff, A. R., Walters, C. M., Shipley, A. A., Weegman, M. D., & Zhao, Q. (2022). Joint analysis of structured and semi-structured community science data improves precision of relative abundance but not trends in birds. *Scientific Reports*, *12*(1), 20289. <https://doi.org/10.1038/s41598-022-23603-0>